

**IN SILICO SCREENING AND MOLECULAR DYNAMIC SIMULATION STUDIES  
OF POTENTIAL SMALL MOLECULE IMMUNO-MODULATORS OF THE  
KIR2DS2 RECEPTOR**

Adekunle Babajide Rowaiye<sup>1</sup>, Jide Olubiyi<sup>2,3</sup>, Doofan Bur<sup>1\*</sup>, Ikemefuna Chijioke Uzochukwu<sup>4</sup>, Alex Akpa<sup>1</sup>, Charles Okechukwu Esimone<sup>5</sup>

<sup>1</sup>Department of Medical Biotechnology, National Biotechnology Development Agency, Abuja, Nigeria.

<sup>2</sup> Department of Pharmaceutical Chemistry, Faculty of Pharmaceutical sciences, Obafemi Awolowo University, Ile-Ife, Nigeria.

<sup>3</sup> Institute of Complex Systems 6, Forschungszentrum Julich, 52428, Germany.

<sup>4</sup>Department of Pharmaceutical and Medicinal Chemistry, Faculty of Pharmaceutical sciences, Nnamdi Azikiwe University, Anambra State, Nigeria

<sup>5</sup>Department of Pharmaceutical Microbiology and Biotechnology, Faculty of Pharmaceutical Sciences, Nnamdi Azikiwe University, Anambra State, Nigeria

## ABSTRACT

Cancer is reported to be one of the most common causes of global deaths and it accounted for an estimated 9.6 million fatalities in 2018. Immunotherapy destroys cancer cells with mechanisms of action considered to be safer and with greater precision than chemotherapy or radiotherapy. The killer cell immunoglobulin-like receptor 2DS2 (KIR2DS2) expressed on Natural Killer (NK) cells are involved in signal transduction processes that produce pro-inflammatory cytokines and directly destroy cancer and virally infected cells. The aim of this study is to identify small molecules from natural products that have strong binding affinity to KIR2DS2 and to predict possible nuclear receptor ligand bioactivity. A library of small molecule natural compounds obtained from edible African plants was used for *in silico* molecular docking simulations of the human natural killer cell activator receptor KIR2DS2 (PDBID: 1m4k) using PyRx. Only compounds with binding affinity stronger than that elicited by the target and the standard, N-acetyl-D-glucosamine were selected. Screening for drug-likeness and ligand efficiency was based on the molecular descriptors of the compounds as provided by PubChem. Further screening for saturation, molar refractivity, promiscuity, pharmacokinetic properties, and bioactivity was done using SWISSADME, PKCSM, and Molinspiration webserver respectively. The molecular dynamic simulation and analyses was done using the Galaxy webserver which uses the GROMACS software. Analyses of molecular dynamic simulation were done using Galaxy and MDWEB webserver. Gibberellin A20 and A29 were obtained as the lead compounds, and they show better promise as drug candidates for KIR2DS2 than the standard. It is recommended that the immuno-stimulatory effect of the lead compounds on KIR2DS2 be further investigated.

**Keywords:** KIR2DS2, NK Cells, ligands, Molecular Docking, Molecular Dynamics Simulation

**Correspondence:** \*Doofan Bur Tel: +234-7038333990; Email: doughfaniyorza@gmail.com

## Introduction

The Killer cell immunoglobulin-like receptors (KIRs) are usually located on a subset of T cells and NK cells. These proteins are coded for by the polymorphic KIR genes, which translate into both activating (aKIR) and inhibiting (iKIR) receptors [1]. The KIR proteins are categorized according to the length of the cytoplasmic domains (short or long) they have or the quantity of extracellular domains (2D or 3D) they possess. The long cytoplasmic domains trigger inhibitory signals through the immune tyrosine-based inhibition motif (ITIM) as compared with the short cytoplasmic domains which trigger activating signals through the immune tyrosine-based activation motif (ITAM) [2].

Like other aKIRs, KIR2DS2 has two functional domains namely: Ig-like C2 type 1 and Ig-like C2 type 2 domains spanning from residues 42-107 and 142-205 respectively. Within these functional domains, KIR2DS2 differs from KIR2DL2 (an inhibitory receptor) with only three residues Tyrosine 45, Arginine 148, and Threonine 200; while it differs from KIR2DL3 (an inhibitory receptor) at only one residue: Tyrosine 45. Tyrosine 45 and Glutamine 71 prevent the recognition of HLA-C ligands by KIR2DS2. KIR2DS2 transduce signals through the TYRO protein tyrosine kinase binding protein (DAP12), Zeta-chain associated protein kinase (ZAP70) and Spleen tyrosine kinase (syk) molecules [3, 4].

Like other Killer cell immunoglobulin-like receptors (KIRs), KIR2DS2 regulates NK cells function and this defines their involvement in conditions such as cancers, inflammations, autoimmune disorders, and infectious diseases [1]. In autoimmunity, KIR2DS2 has been shown to influence alloreactivity and increase the number of EB6 (KIR2DL1/S1)-expressing NK clones [5]. As in other KIRs, KIR2DS2 also affects the susceptibility to various diseases such as ulcerative colitis [6], cytomegalovirus infection [7], and hepatitis C [8].

Though largely unidentified, the ligands for KIR can be found in the subsets of Human Leucocyte Antigen (HLA) class I molecules. Whereas HLA–ligand specificities are well established for iKIRs, they are not for aKIRs [9, 10]. In spite of a high structural similarity, KIR2DS2 does not bind with HLA-C2 while KIR2DL2 and KIR2DL3 do [11, 12].

However, a successful crystallization of KIR2DS2 in complex with HLA-A\*11:01 has been reported [13]. The Tyr 45 residue of KIR2DS2 binds to the NH group of Thr80 of the  $\alpha$ -1 helix of HLA-A\*11:01 forming a hydrogen bond. This is remarkably different from iKIRs that use residue 44 to bind to Asn80 in HLA-C1 or Lys80 in HLA-C2 [13]. Beyond this, KIR2DS2 also recognizes a  $\beta$ 2-microglobulin-independent ligand expressed on cancer cells [14].

N-acetyl-D-glucosamine (NAG) has been known to bind effectively with the activation receptors of the rat natural killer cells [15].

Similar to NAG, researchers on this project were saddled with the responsibility of prospecting for other small molecules that have an immuno-stimulatory effect on the KIR2DS2 receptor. A library of compounds from fruits and vegetables was used with the hope that the leads generated from this research would be used to produce nutraceuticals.

## Methodology

### Materials

The ligand and protein databases used include Protein Databank, Uniprot, and PubChem. MolProbity, ExPasy, Chiron, pkCSM, SwissADME, Molinspiration, Clustal Omega, Protein-Ligand Interaction Profiler (PLIP), Galaxy and MDWEB were used as webservers. The software used were PyMol, Python prescription (PyRx) 0.8 and Discovery studio 2017.

### Methods

*Preparation, analysis and validation of target protein structure:* The crystal structure of the human natural killer cell activator receptor KIR2DS2 (PDBID: 1m4k) was retrieved from the Protein Data Bank. The visualization tool, PyMol was used to remove water molecules and ligand from the protein structure [16]. The online server, Chiron was used for energy minimization to reduce the steric clashes of the protein structure [17]. The Volume, Area, Dihedral Angle Reporter (VADAR 1.8) webserver was used to reveal the architecture of the target protein while it was analyzed and validated using the Ramachandran and hydrophobicity plots obtained from the

MolProbity and EXPasy web servers respectively [18,19].

*Molecular docking:* A library of 1,647 natural compounds was built from 79 edible African plants which are predominantly fruits, vegetables and spices. 1m4k was docked against compounds in the library using the PyRx software (0.8 version) and these compounds were screened for total compliance with Lipinski and Veber rules using their molecular descriptors obtained from the PubChem database [20]. All ligands were converted from the SDF to pdbqt format. For stable conformation, the conjugate gradient descent was used as optimization algorithm and the Universal Force Field (UFF) was used as the energy minimization parameter [20].

### *Cheminformatics-based characterization of compounds:*

A reference compound, N-acetyl-D-Glucosamine (Pubchem ID: 439174) was also docked against the target protein and the score obtained was used as a cut off. For an effective management of data, further screening was performed with a cut off of -7.0 kcal/mol. All the docked complexes were further evaluated for the efficiency of the ligand interaction to the target using the Ligand Efficiency Metrics (LEM). The LEM used are the Ligand Efficiency (LE), Ligand-Efficiency-dependent Lipophilicity, (LELP) and the Ligand-Lipophilicity Efficiency (LLE). Further screening for molar refractivity, saturation, promiscuity, pharmacokinetic properties, and bioactivity were performed using the SWISSADME, pkCSM, and Molinspiration webservers respectively [21].

**Analysis of protein-ligand interactions:** After molecular docking, the poses of the selected ligands as they interact with the receptor were saved on PyRx and viewed on PyMol. Also using PyMol, the protein structure was superimposed on the ligand poses and saved in the pdb format. The resultant structure was uploaded into the Protein-Ligand Interaction Profiler (PLIP) webserver to reveal the bonds and interactions between the target and the ligands [22, 16]. Like NAG, only compounds that have hydrogen bond interaction with TYR45 were selected. However, a front runner compound, Monocrotaline which had hydrophobic interaction at TYR 45 was removed because of its established toxicity profile. Though computationally predicted to be safe, Monocrotaline is used to induce pulmonary arterial hypertension in experimental animals [23]

**Molecular Dynamics Simulations:** The molecular dynamics simulation (MDS) of the native protein and all its ligand complexes was performed using the Galaxy (versions 2019.1 and 2019.1.4) supercomputing server which uses the Groningen Machine for Chemical Simulations (GROMACS) software [24] GROMACS-compatible Gro and topology files for small molecules were generated using LigParGen with OPLS-AA/1.14\*CM1A as force field parameters [25]

The parameters for the initial simulation set up include: SPC/E for water model; OPLS/AA as the force field; box type is rectangular with all

equal sides each measuring 1.0nm; and Hydrogen was ignored. GROMACS solvation and addition of ions was done using the SPC water model. Energy minimization was conducted using the steepest descent algorithm for 50,000 steps. 1.0 was used for Distance cut- offs include for short range van der Waals, Coulomb, and the short-range neighbor list respectively. Other parameters for minimization include Fast smooth Particle-Mesh Ewald electrostatics and EM tolerance of 1000. Equilibration was performed using NVT and NPT ensembles with the following parameters: leap-frog integrator, constrained Bonds with H-atoms, 300k temperature, 0.002 ps step length and 5,000 steps between saving data points. Finally, a 1ns molecular dynamics simulation was carried out for the Apo and Holo proteins with 500,000 steps.

Using the BIO3D tool on the Galaxy supercomputing platform, the Root Mean Square Deviation of atomic positions (RMSD), Root Mean Square Fluctuation (RMSF) of protein backbone, Dynamical Cross-Correlation Matrix (DCCM), Principal Component Analysis (PCA), were determined [26]. The Radius of Gyration (RoG) and B Factor were obtained using the MDWEB webserver [27]. To differences between the crystal structure of 1m4k (freed from water and co-crystallized ligand) and the apo structure of 1m4k that had undergone MDS were observed [28].

## Discussion of Results

*Structural analysis, validation and preparation of KIR2DS2 (PDB ID: 1m4k):* Figure 1 reveals

the Kyte & Doolittle hydrophobicity plot of the full protein structure as found in the UniProt database (P43631). The structure consists of signal peptides (1-21), extracellular domains (22-245), helical domain (246-265), and cytoplasmic domain (266-304). Within the extracellular domain are the Ig-like C2-type 1 (42-107) and Ig-like C2-type 2 (142-205) sub-domains. From Figure 2, the signal peptide is hydrophilic while Ig-like C2-type 1 and Ig-like C2-type 2 domains are predominantly hydrophobic.

The Apo structure, 1m4k is the extracellular domain of KIR2DS2 (Figure 1). It has 193 amino acids with the following constituent secondary structures:  $\alpha$  helix 0%; beta sheets 72%; coil 27%; and turns 10%. The total accessible solvent area (ASA) is  $9074.7(\text{\AA})^2$ . X ray diffraction study revealed resolution is  $2.3\text{\AA}$  and unit cell crystal dimensions are  $a = 97.535\text{\AA}$ ,  $b = 97.535\text{\AA}$ , and  $c = 54.375\text{\AA}$  for  $\alpha$  ( $90^\circ$ ),  $\beta$  ( $90^\circ$ ) and  $\gamma$  ( $120^\circ$ ) angles respectively. The R-Value is 0.221 and Rfree value is 0.247. The residues of 1m4k show no helical propensity and this is because it is not a transmembrane protein which is usually rich in alpha helices [29]. The R value is an evaluation of the quality of a refined protein crystal structure. As the R value tends to zero, the quality improves. With an R value of 0.221 and a resolution of  $2.3\text{\AA}$ , 1m4k is a high quality crystal structure and this would yield high quality results during molecular docking [30].

From Figure 3, Ramachandran analysis of the candidate structure shows that 95.29% of residues are within the favoured region and

2.09% outliers. Rotamer analysis shows 91.93% of rotamers are within the favoured region and 4.97% poor rotamers. Further protein geometry reveals no  $C\beta$  deviations, 0.13% bad bonds and 0.24% bad angles. The VDW repulsion energy of 1m4k was minimized by Chiron from 56.15 Kcal/mol (clash ratio of 0.023) to 35.85 Kcal/mol (clash ratio of 0.015). The results confirm the high stereochemical quality of 1m4k as seen in the high percentage of residues and rotamers in the favoured region, and the very low clash ratio. These factors would enhance good quality docking results and reduce steric clash between atoms of protein and ligand during interaction [18, 17].

#### ***Molecular docking analyses of ligands against***

***KIR2DS2***: Ligand Efficiency Metrics (LEM) was used to evaluate the ligand efficiency of the compounds based on their size and lipophilicity [31]. The LEM used include Lipophilic Efficiency (LE), Ligand-Efficiency dependent Lipophilicity (LEDL), and Ligand Lipophilicity Efficiency (LLE). As seen in Table 3, GA20, GA29 and the standard met the cut-off LE ( $\geq 0.3$ ), LELP (-10 to 10), and LLE (greater than 5) [32]. The standard and the lead compounds met all the cut-offs for ligand efficiency.

#### ***Cheminformatics profile of ligands (Figure 4,***

***Table 1)***: Drug-like properties are important for bioavailability and cellular uptake within the body. To meet these criteria, all the molecular descriptors of the front runner compounds should not violate the Ghose, Lipinski (RO5),

and Veber rules. Put together these rules state that the molecular weight should be  $\leq 500$ g/mol; hydrogen bond acceptors should be  $\leq 10$ ; hydrogen bond donors should be  $\leq 5$ ; Log P should be  $\leq 5$ ; the polar surface area should be  $\leq 140$ Å<sup>2</sup>; number of rotatable bonds should be  $< 10$ ; and molar refractivity should be between 40-130 cm<sup>3</sup> [33,34,35].

From the results, the standard and the lead compounds did not violate the Ghose, Lipinski (RO5), and Veber rules suggesting that they have good drug permeability. Specifically, the standard is predicted to have the greatest absorbability across lipid membranes because it has the highest lipophilicity of all the compounds. GA 20 will permeate the blood-brain barrier; this is because its TPSA value is less than 90 angstroms squared. GA 29 and the standard will not permeate the blood-brain barrier. [36].

The molecular complexity of a compound can be evaluated by its saturation. The standard and the two lead compounds have their fraction of carbons in the sp<sup>3</sup> hybridization more than 0.25 (Table 1). The results suggest that the standard has the highest saturation [37]. In biological assays, promiscuous or PAINS (pan assay interference) compounds yield false positive response because they have undesirable molecular moieties. The standard and the two lead compounds had no PAIN alerts [38].

Furthermore, the standard, GA 20, and GA29 showed lead-like properties by having good bioactivity prediction scores against the major classes of drug targets such as ion channels, G-

Coupled Protein Receptors (GCPR), nuclear receptors, proteases, kinases, and enzymes. The standard and the two lead compounds are predicted to have moderate kinase inhibition activity. For all the targets except enzymes, GA20 and GA29 are predicted to have greater activity than the standard. The standard is predicted to have highest best enzyme inhibition activity when compared with the two lead compounds [39]. Overall, the results show that the standard, GA20, and GA29 are good drug candidates (Table 1).

**Pharmacokinetic properties of ligands:** Drugs should have good pharmacokinetic properties. The Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) properties of the front-runner compounds were predicted *in silico* using graph-based signatures (Table 2) [40]. Absorption parameters such as water solubility (insoluble:  $-4.0$  Log mol/L), Caco2 permeability (high:  $> 0.9$ ), human intestinal absorption (poor:  $<30\%$ ), and skin permeability (low: LogKp  $>-2.5$ ) are indicators of the therapeutic potential of chemical compounds as these ensure their penetration to reach the target molecule. All values obtained for GA20, GA29 and the standard are within pharmacological range and none of them are inhibitors of p-glycoprotein [40,41].

Comparing to standard values, distribution properties such as fraction unbound (not less than 0.1), volume of distribution (Low: Log VD<sub>ss</sub>  $<- 0.15$ ; High: Log VD<sub>ss</sub>  $> 0.45$ ), the Blood Brain Barrier (permeable: Log BBB  $> 0.3$ ; poor  $<: \text{Log BBB} <-1$ ), and Central

Nervous System (permeable Log PS > -2; poor Log PS < -3) permeability were evaluated. All values obtained for GA20, GA29 and the standard for volume of distribution and the Blood Brain Barrier are within pharmacological range. However, the standard has a poor CNS permeability [40,42].

The inhibition of some isoforms of cytochrome P450 could lead to potential drug-drug interaction that might have toxic consequences. The predicted metabolic behavior of GA20, GA29, and the standard shows no inhibition of CYP1A2, CYP2C19, CYP2C9, CYP2D6 and CYP3A4 enzymes. However, GA20 and GA29 are substrates of CYP3A4. This means that their dose would be affected either by induction or the inhibition of CYP3A4 [43].

The predicted excretion values for Total Clearance for GA20, GA29 and the standard are within pharmacological range [44]. GA20, GA29 and the standard are not substrates of Renal Organic Cation Transporter 2 (OCT2). This implies that Renal OCT2 will eliminate none of them from the blood into the proximal tubular cell [45].

The maximum recommended tolerated dose determines the dose to be administered in the phase 1 of clinical trials. GA20 and GA29 have low maximum recommended tolerated dose (less than 0.477 log mg/kg/day) while the standard has a high value (more than 0.477 log mg/kg/day). GA29 is the most potent [40].

The toxicity required to inhibit 50% of the growth of *T.pyriformis* (IGC50), protozoan

bacteria is very important parameter measured in drug discovery. A compound is considered toxic when the pIGC50 (negative logarithm of the concentration required to inhibit 50% growth in log ug/L) value is greater than -0.5 log ug/L. From the results, GA20, GA29, and the standard are predicted to be toxic against *T.pyriformis*. This suggests an antibacterial effect and might not be toxic to human cells [40]. Similarly, log LC50 is the log of a compound to cause death of 50% of flathead Minnows. A log LC50 value less than 0.3 log mM signifies high acute toxicity. From the results, GA20, GA29 and the standard are not toxic to Minnows [40].

HerG inhibition and AMES toxicity are critical toxicity parameters as it reveals the cardio-toxic properties and mutagenic behavior of the compounds respectively. GA20, GA29 and the standard did not show HerG inhibition, or AMES toxicity. They all also did not show dermatotoxic and hepatotoxic properties [40].

*Analyses of protein-ligand interactions:* Hydrogen bond determines the specificity of ligand binding and the analysis of this bond captures two key factors which are the length and orientation of the bond [46]. From Figures 5 & 6 and Table 4, the standard and lead compounds all have hydrogen bonding with residues within the functional domains of the target which are the Ig-like c2 type 1 (42-107) and the Ig-like c2 type 2 (142-205) domains. NAG had the highest number of hydrogen bonds within these domains.

Notably, the standard and the lead compounds form hydrogen bonds on LYS44 and TYR45. Only NAG forms hydrogen bond at GLN 71. TYR45A and LYS44 are important in determining the signals of the KIR2DS2 (activating) and KIR2DL2 (inhibitory) respectively. GLN 71 is important in determining whether the target binds to HLA-C. This suggests that these compounds can both trigger the two receptors [3, 4, 13]. In terms of the donor to acceptor distance, NAG forms a moderate (2.5-3.2 Å) hydrogen bond at TYR 45 while GA 20 and GA29 form weak (3.2-4.0 Å) hydrogen bonds at the same residue. In spite of the different strata of hydrogen bond strength, only a slight difference exists between the standard (3.15 Å) and GA 20 and GA29 (3.40 and 3.42Å respectively). In terms of bond angle formed at TYR45, all compounds (standard and the leads) form weak (less than 130°) hydrogen bonds [47, 48].

From Figure 5 and Table 5, all the compounds form other kinds of interactions at LYS44. The standard forms a salt bridge at this residue while the lead compounds form both salt bridges and hydrophobic interactions. The presence of salt bridges in all the protein-ligand interactions of the standard and lead compounds enhances the strength and stability of the complexes [49]. The strength of the salt bridge is affected by the distances between the atoms of ligand and those of residues. The distance required is less than 4 Å [50]. This suggests that the salt bridge formed by NAG at LYS44 is a weak one.

The number of hydrophobic interactions is three times higher in GA20 and GA29 as compared with the standard. These optimal interactions suggest that GA 20 and GA29 are having more atom efficient binding than NAG [51]. Only NAG had a hydrophobic interaction at TYR 45, while the lead compounds had at GLN 71 [13, 3, 4].

### **Molecular Dynamic Simulation Analyses**

*Root Mean Square Deviation of Atomic Positions (RMSD):* RMSD is the most common quantitative measure of similarity between two superimposed protein structures (the reference and target structures). It measures the variations in the distances between atoms in two superimposed structures and an RMSD value of 0.0 indicates a perfect overlap [52].

From Figure 8 and Table 6, the RMSD of the simulated Apo protein relative to the crystal structure suggests a gradual increase as the production time increased. It peaked at Frame 43 (2.51 Å) and thereafter stabilized. The global RMSD for the Apo structure is 180.94 Å while the average is 1.79 Å. The RMSD for the simulated KIR2DS2 - NAG complex suggests a steady rise until it peaked at Frame 70 (2.25 Å) and thereafter stabilized. The global RMSD for this holo structure is 174.91 Å while the average is 1.73 Å. The RMSD for the simulated KIR2DS2-GA20 complex suggests a steady rise until it peaked at Frame 39 (2.26 Å) and thereafter stabilized. The global RMSD for this holo structure is 166.41 Å while the average is 1.65 Å. Also, the RMSD for the simulated KIR2DS2-GA29 complex suggests a steady



rise until it peaked at Frame 65 (2.5 Å) and thereafter stabilized. The global RMSD for this holo structure is 165.22 Å while the average is 1.64 Å. From Figure 9 and Table 6, the RMSD histogram suggests that peak distribution is more tilted to the left in the KIR2DS2 – GA20 and KIR2DS2 – GA29 complexes than the KIR2DS2 - NAG complex. Specifically, KIR2DS2 – GA29 had the highest number of peaks within the 1.0 to 1.5 Å range.

Put together, 98.7% all the RMS distances between alpha carbon atoms of the three holo structures fall within the resolution of the crystal structure of the KIR2DS2 (2.3 Å) [53]. KIR2DS2 – GA20 and KIR2DS2 – GA29 complexes showed better stability than the than the KIR2DS2 - NAG complex. The KIR2DS2 – GA29 complex showed the best stability as seen with the least total and average RMSD and the highest number of peaks within the 1.0 to 1.5 Å range (greatest left shift).

**Root Mean Square Fluctuations (RMSF):** At a wide range of time scales, proteins undergo structural fluctuations under physiological and pathological conditions. The RMSF reveals the dynamic behaviour of various residues by measuring the lowest energy modes of fluctuations around their equilibrium conformations [54].

From Figure 10 and Table 6, the total and average global RMSF is smaller in the KIR2DS2 – GA20 and the KIR2DS2 – GA29 complexes than the KIR2DS2 - NAG complex showing greater stability. The KIR2DS2–

GA20 complex shows slightly lesser fluctuations globally than the KIR2DS2 – GA29 complex. In a similar manner, the total and average fluctuations within the Ig-like c2 type 1 functional domain (residues 42-107), is smaller in the KIR2DS2 – GA20 and KIR2DS2–GA29 complexes than the KIR2DS2-NAG complex showing that the lead compounds induce greater stability than the standard. In this domain, the KIR2DS2 – GA29 complex shows a marginally lesser fluctuation than the KIR2DS2 – GA20 complex. Furthermore, on TYR45, both lead compounds induced less fluctuation as compared with the standard. The KIR2DS2 – GA29 complex had the least fluctuation value at this important residue.

Overall, while evolutionary conservation of backbone fluctuations is seen in KIR2DS2, greater number of conformations due to ligand binding reveals that KIR2DS2-GA29 complex is the most stable holo structure. Though KIR2DS2 – GA20 complex had the least global fluctuations, KIR2DS2-GA29 complex had the least regional and local (TYR 45) fluctuations.

**Radius of gyration (RoG):** The RoG is an indicator of the compactness of the secondary structures within the 3D structure of the protein. Measured from the centre of mass of the molecule, a high RoG suggests loose packing while a low RoG suggests a tight packing of the protein [55]. After an MDS of 1ns, both Apo and holo structures showed relative compactness in terms of the degree of folding due to the hydrophobic effect and the

architecture of target protein is largely conserved. This means that ligand induced conformation by NAG, GA20 and GA29 had little effect in unfolding the target protein. This is revealed in the marginal differences in their percentages of gyration. The KIR2DS2-GA29 is the most compact of all the structures. This is as suggested by the lowest average gyration over the trajectory, the lowest range of gyration and the lowest percentage gyration. Predictably, both lead compounds have less impact on the unfolding of KIR2DS2 than the standard (Figure 11 and Table 6).

*B-Factor:* This describes the thermostability of a protein molecule by identifying regions of flexibility, rigidity and internal motions [56]. From Figure 12 and Table 6, the values of the global average, regional (residues 42-107) average and local B-factors are less in the KIR2DS2-GA20 and KIR2DS2-GA29 complexes than in the KIR2DS2-NAG complex. This suggests that the lead compounds-induced conformations are more stable than the conformation induced by the standard. While KIR2DS2-GA29 complex is most stable at the global and regional level, KIR2DS2-GA20 complex is most stable at TYR45.

*Principal components Analysis:* Principal component analysis (PCA) can be used to determine the relationship between statistically meaningful conformations (major global motions) sampled during the trajectory [57]. From Figure 13 and Table 6, the global motions were least in the KIR2DS2 – GA29 complex.

The KIR2DS2–GA20 complex showed greater global motion than the KIR2DS2 - NAG complex. In a similar manner, the average motion of the first three principal components (PC1, PC2 & PC3) at residue 45 shows that the KIR2DS2–GA29 complex is the most stable holo structure while the KIR2DS2– NAG complex is more stable than the KIR2DS2–GA20 complex.

Based on the least motions, the best global conformations are PC1 of the Apo protein, PC2 of KIR2DS2–NAG complex, PC3 of KIR2DS2-GA20 complex, and PC3 of the KIR2DS2 – GA29 complex. Similarly, the best conformations that produced the least motions at TYR 45 are PC1 and PC3 of the Apo protein, PC3 of KIR2DS2-NAG complex, PC3 of KIR2DS2-GA20 complex, and PC3 of the KIR2DS2–GA29 complex. The cosine contents of the principal components reveal the quality of the sampling and to what direction the MD simulation is converging (Table 6).

*Dynamic Cross Correlation Map (DCCM):* The DCCM reveals the heat map of cross correlation of residual fluctuations. The pairwise graph reveals the positive and negative correlation effects of the atomic displacements in the residues as they correlate with one another [58].

Figure 14 and Table 6, reveals a complex pattern of correlated, non-correlated and anti-correlated motions in Apo and Holo structures. Comparative results reveal that though more intensified, the atomic motions in the

KIR2DS2-NAG complex resemble more closely that of the Apo structure. Similar to the Apo structure, approximately residues 1-50 (which includes TYR45) of the KIR2DS2-NAG complex shows correlated motion. The Ig-like c2 type 1 domain (42-107) of the same holo structure predominantly shows anti-correlated atomic motions. The map for Ig-like c2 type 2 domain (142-205) suggests that most residues predominantly showed non-correlated atoms. The heat maps of the KIR2DS2-GA20 and KIR2DS2-GA29 complexes closely resembled. They showed less anti-correlated and correlated motions, and more non-correlated motions than the KIR2DS2-NAG complex in all regions of the protein.

The standard and lead compounds: Glucosamine can be obtained from the shells of crabs, lobsters and shrimps. Orally administered NAG has been used in the treatment of Inflammatory Bowel Disease (IBD) and osteoarthritis [59]. The anti-cancer effect of NAG has also been well documented. By activating Death Receptor 5, NAG as an adjuvant promotes TRAIL-induced apoptosis in Non-Small Cell Lung Cancer Cells [60]. Glucosamine also to exhibited antitumor activity in lung cancer cells by suppressing Forkhead box O proteins (FOXO) phosphorylation. Through a structural modification, Glucosamine specifically inhibited the translocation of FOXO1 and FOXO3 proteins which downstream signaling molecules for the PI3K/AKT and MAPK/ERK pathways [61]. Glucosamine has also been known to induce cell death in prostate cancer

cells by the inhibition of proteasomal activity [62].

Gibberellins are plant hormones that regulate its various developmental processes and they are found abundantly in *Abelmoschus esculentus* (okra) and *Pisum sativum* (green peas) [63,64]. They are known to be involved in plant innate immunity by mostly by regulating SA-JA-ET signaling systems [65]. A Gibberellin derivative, GA-13315 has shown great promise as a chemotherapeutic agent against lung cancer [66]. Some gibberellin-based molecules which show strong anticancer activities have also been designed and synthesized [67]. Specifically, the cytotoxic or immunomodulatory effect of GA20 and GA29 are yet to be ascertained.

**Summary of comparative analyses:** Following a thorough evaluation of the bioavailability, pharmacokinetic properties and binding site analyses, NAG, GA20 and GA29 are predicted to be good drug candidates and have an immuno-stimulatory effect on KIR2DS2. However, GA29 would have the greatest immuno-stimulatory effect because it has the best binding affinity (-8.2 Kcal/mol) to the KIR2DS2 receptor and a greater nuclear receptor ligand activity prediction than NAG. GA20 would also has a better binding affinity (-7.4 Kcal/mol) and greater nuclear receptor ligand activity than the standard. Unlike, NAG, GA20 and GA29 are substrates of the CYP3A4 and this suggests that they should be administered with an inhibitor of that enzyme.

Studying the time-resolved motions of Apo and Holo macromolecules, GA20 and GA29 are predicted to have better pharmacodynamics on the target receptor than NAG. Specifically, of the two lead compounds, GA29 is predicted to have a better efficacy. This is because GA29 has shown better structural stability as seen with the values from the RMSD (values and distribution of peaks), RMSF (at TYR45), RoG (percentage gyration, range and average), B Factor (global and regional), and PCA (global and local motions). Furthermore, GA 29 has the strongest binding affinity to KIR2DS2 (-8.2 kcal/mol).

### Conclusion

This study was aimed at identifying compounds from natural products that have strong binding affinity to KIR2DS2 and to predict possible nuclear receptor ligand bioactivity. After molecular docking, virtual screening, and molecular dynamics simulations, two lead compounds emerged namely, GA29 and GA20. GA29 is predicted to have a better efficacy. It is recommended that the immuno-stimulatory effect of Gibberellins A20 and A29, and NAG on KIR2DS2 be evaluated using *in vivo* and *in vitro* experiments.

**Conflict of interest:** None

### References

[1] S. Kulkarni, M.P. Martin & M. Carrington (2008). The Ying and Yang of HLA and KIR in human disease. *Seminars in Immunology*. 20:343–352

[2] L.B. Ivashkiv, (2011). How ITAMs Inhibit Signaling. *Science Signaling*. 4(169): 20.

[3] A.K. Moesta and P. Parham (2012). Diverse functionality among human NK cell receptors for the C1 epitope of HLA-C: KIR2DS2, KIR2DL2, and KIR2DL3. *Front Immunology* 3: 336.

[4] X. Saulquin, L.N. Gastinel & E. Vivier (2003). Crystal structure of the human natural killer cell activating receptor KIR2DS2 (CD158j). *Journal of Experimental Medicine* 197(7): 933-8.

[5] J.H. Chewning, C.N. Gudme, G. Heller & B. Dupont (2007). The Activating KIR2DS2 Gene Influences NK Alloreactivity and NK Repertoire. *Blood* 110:313.

[6] D.C. Jones, R.S. Edgar, T. Ahmad, J.R.F. Cummings, D.P. Jewell, J. Trowsdale & N.T. Young (2006). Killer Ig-like receptor (KIR) genotype and HLA ligand combinations in ulcerative colitis susceptibility. *Genes and immunity*, 7(7), 576.

[7] V. Béziat, L.L. Liu, J.A. Malmberg, M.A. Ivarsson, E. Sohlberg, A.T. Björklund, & M. Schaffer (2013). NK cell responses to cytomegalovirus infection lead to stable imprints in the human KIR repertoire and involve activating KIRs. *Blood*, 121(14), 2678-2688.

[8] M. Askar, R. Avery, R. Corey, R. Lopez, D. Thomas, D. Pidwell & N.N. Zein (2009). Lack of killer immunoglobulin-like receptor 2DS2 (KIR2DS2) and KIR2DL2 is associated with poor responses to therapy of recurrent hepatitis C virus in liver transplant recipients. *Liver Transplantation*, 15(11), 1557-1563.

- [9] M.N. Naiyer, S.A. Cassidy, A. Magri, V. Cowton, K. Chen, S. Mansour, H. Kranidioti, B. Mbiribindi, P. Rettman, S. Harris, L.J. Fanning, A. Mulder, F.H.J. Claas, A.D. Davidson, A.H. Patel, M.A. Purbhoo & S.I. Khakoo (2017). KIR2DS2 recognizes conserved peptides derived from viral helicases in the context of HLA-C. *Science Immunology* 2: 15.
- [10] C.M. Ayo, P.G. Reis, M.M.dO Dalalio, J.E.L. Visentainer, C.dF. Oliveira, S.M. de Araújo, et al. (2015) Killer Cell Immunoglobulin-like Receptors and Their HLA Ligands are Related with the Immunopathology of Chagas Disease. *PLOS Neglected Tropical Diseases* 9(5): e0003753.
- [11] C.C. Winter, J.E. Gumperz, P. Parham, E.O. Long & N. Wagtmann (1998) Direct binding and functional transfer of NK cell inhibitory receptors reveal novel patterns of HLA-C allotype recognition. *Journal of Immunology*. 161(2):571–577.
- [12] A.M. Older Aguilar, L.A. Guethlein, L. Abi-Rached & P. Parham (2011). Natural variation at position 45 in the D1 domain of lineage III killer cell immunoglobulin-like receptors (KIR) has major effects on the avidity and specificity for MHC class I. *Immunogenetics* 63(8):543–547.
- [13] J. Liu, X. Ziwei, K. Hui Ling, S. Meixin, & R. Ee Chee (2014). Activating killer cell immunoglobulin-like receptor 2DS2 binds to HLA-A\*11. *PNAS* 111 (7) 2662-2667
- [14] L. Thiruchelvam-Kyle, S.E. Hoelsbrekken, P.C. Saether, E.G. Bjørnsen, D. Pende, S. Fossum, M.R. Daws & E. Dissen (2017). The Activating Human NK Cell Receptor KIR2DS2 Recognizes a  $\beta$ 2-Microglobulin-Independent Ligand on Cancer Cells. *Journal of Immunology* 198(7): 2556-2567.
- [15] K. Krenek, M. Kuldová, K. Hulíková, I. Stibor, P. Lhoták, M. Dudic, J. Budka, H. Pelantová, K. Bezouska, A. Fiserová & V. Kren (2007). N-acetyl-D-glucosamine substituted calix arenes as stimulators of NK cell-mediated antitumor immune response. *Carbohydrate Research* 342(12-13): 1781-92.
- [16] W.L. DeLano (2002). Pymol: An open-source molecular graphics tool. *CCP4 Newsletter on Protein Crystallography*. 40: 82-92.
- [17] S. Ramachandran, P. Kota, F. Ding and N.V. Dokholyan (2011) *PROTEINS: Structure, Function and Bioinformatics*, 79: 261-270
- [18] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel & A. Bairoch (2005). Protein Identification and Analysis Tools on the ExPASy Server; (In) John M. Walker (ed): *The Proteomics Protocols Handbook*, Humana Press pp. 571-607
- [19] I.W. Davis, A. Leaver-Fay, V.B. Chen, J.N. Block, G.J. Kapral, X. Wang, L.W. Murray, W.B. Arendall, J. Snoeyink, J III. Richardson, J.S. and Richardson, D.C. (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research* 35: W375-W383.
- [20] S. Dallakyan & A.J. Olson (2015) Small-Molecule Library Screening by Docking with PyRx. *Methods in Molecular Biology*. 1263:243-50.

- [21] A. Daina, O. Michielin & V. Zoete (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*. 7: 42717.
- [22] S. Salentin, S. Schreiber, V.J. Haupt, M.F. Adasme & M. Shroeder (2015) PLIP: Fully automated protein-ligand interaction profiler. *Nucleic Acids Research* 1(43):443-447
- [23] M. Rey, P. Hess & M. Clozel (2009) Monocrotaline-induced pulmonary hypertension in Wistar rats. *Current Protocols in Pharmacology* 5(5):56.
- [24] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith et al., 2015 GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1-2: 19–25.
- [25] L.S. Dodda, I. Cabeza de Vaca, J. Tirado-Rives & W.L. Jorgensen (2017). LigParGen web server: An automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Research*, 45 (1):331-336
- [26] E. Afgan, D. Baker, M. van den Beek, D. Blankenberg, D. Bouvier, M. Cech, J. Chilton, D. Clements, N. Coraor, C. Eberhard, B. Gruning, A. Guerler, J. Hillman-Jackson, G. Von Kuster, E. Rasche, N. Soranzo, N. Turaga, J. Taylor, A. Nekrutenko & J. Goecks (2016) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: *Nucleic Acids Research* 44, W3–W10.
- [27] H. Adam, A. Pau, F. Carles, C. Damjan, O. Modesto & J.L. Gelpí (2012). MDWeb and MDMoby: an integrated web-based platform for molecular dynamics simulations. *Bioinformatics*, 28(9):1278-1279.
- [28] R Maiti, G. H. Van Domselaar, H. Zhang, and D. S. Wishart (2004). SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Res.* 32(Web Server issue): W590–W594.
- [29] L. Liu and C. M. Deber (1998). Uncoupling hydrophobicity and helicity in transmembrane segments.  $\alpha$ -helical propensities of the amino acids in non-polar environments. *Communication*. 273 (37): 23645-23648.
- [30] J. Wang (2015). Estimation of the quality of refined protein crystal structures. *Protein Sci.* 24(5): 661–669.
- [31] P.W. Kenny, L. Andrei & C.A. Montanari (2014). Ligand efficiency metrics considered harmful. *Journal of Computer-Aided Molecular Design* 28(7)699-710.
- [32] S.A. Egieyeh, S. James, F.M. Sarel & C. Alan (2016) Prioritization of anti-malarial hits from nature: chemo-informatic profiling of natural products with in vitro antiplasmodial activities and currently registered anti-malarial drugs. *Malaria Journal* 15:50
- [33] C.A. Lipinski, F. Lombardo, B.W. Dominy & P.J. Feeney (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Rev.* 46, 3–26.
- [34] D.F. Veber, S.R. Johnson, H.Y. Cheng, B.R. Smith, K.W. Ward & K.D. Kopple (2002). Molecular properties that influence the oral bioavailability of drug candidates. *Journal of Medicinal Chemistry*, 45(12), 2615-2623.

- [35] A.K. Ghose, V.N. Viswanadhan & J.J. Wendoloski (1999). A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. A qualitative and quantitative characterization of known drug databases. *Journal of Combinational Chemistry*. 1(1):55-68.
- [36] S.A. Hitchcock & L.D. Pennington (2006). "Structure-Brain Exposure Relationships". *Journal of Medicinal Chemistry* 49 (26): 7559–7583. in *Clinical Challenges in Therapeutic Drug Monitoring*, 2016.
- [37] P.A. Clemons, N.E. Bodycombe, H.A. Carrinski, J.A. Wilson, F.S. Alykhan, B.K. Wagner, A.N. Koehler & S.L. Schreiber (2010). Small molecules of different origins have distinct distributions of structural complexity that correlate with protein-binding profiles. *PNAS*. 107(44) 18787-18792
- [38] S.J. Capuzzi, E.N. Muratov & A. Tropsha (2017). Phantom PAINS: Problems with the Utility of Alerts for Pan-Assay INterference CompoundS. *Journal of Chemical Information & Modelling*. 57(3): 417–427.
- [39] T. Khan, S. Dixit, R. Ahmad, S. Raza, I. Azad, S. Joshi & A. Khan (2017). Molecular docking, PASS analysis, bioactivity score prediction, synthesis, characterization and biological activity evaluation of a functionalized 2-butanone thiosemicarbazone ligand and its complexes. *Journal of Chemical Biology*. 10(3): 91–104
- [40] D.E.V. Pires, T.L. Blundell & D.B. Ascher (2015). pkCSM: predicting small-molecule pharmacokinetic properties using graph-based signatures. *Journal of Medicinal Chemistry*, 58 (9), 4066–4072.
- [41] K.T. Savjani, A.J. Gajjar. & K.J. Savjani (2012). *Drug Solubility: Importance and Enhancement Techniques*. ISRN Pharmaceutics. 2012: 195727.
- [42] K.L. Johnson-Davis & A. Dasgupta (2016). *Special Issues in Therapeutic Drug Monitoring in Patients with Uremia, Liver Disease, and in Critically Ill Patients in Clinical Challenges in Therapeutic Drug Monitoring*, 2016.
- [43] S.F. Zhou (2008), Drugs behave as substrates, inhibitors and inducers of human cytochrome P450 3A4. *Current Drug Metabolism*. 9(4):310-22.
- [44] S.W. Paine, P. Barton, J. Bird, R. Denton, K. Menochet, A. Smith, N. P. Tomkinson, K. K. Chohan (2010) A rapid computational filter for predicting the rate of human renal clearance. *Journal of Molecular Graphics & Modelling*. 29:529-37
- [45] K. Hacker, R. Maas, J. Kornhuber, M.F. Fromm & O. Zolk (2015). Substrate-Dependent Inhibition of the Human Organic Cation Transporter OCT2: A Comparison of Metformin with Experimental Substrates. *PLoS One*. 10(9): e0136451.
- [46] R.C. Wade & P.J Goodford (1989). The role of hydrogen-bonds in drug binding. *Progress in clinical and biological research*, 289:433-44.
- [47] G.A. Jeffrey (1987). *An introduction to hydrogen bonding*, Oxford University Press.
- [48] G.R. Desiraju and T. Steiner (2001) *The Weak Hydrogen Bond in Structural*

Chemistry and Biology. Oxford: Oxford University Press.

[49] T. Stockner, H.J. Vogel & D. Peter Tieleman (2005). A Salt-Bridge Motif Involved in Ligand Binding and Large-Scale Domain Motions of the Maltose-Binding Protein. *CellPress* 89 (5): 3362-3371.

[50] S. Kumar & R. Nussinov (July 2002). "Close-range electrostatic interactions in proteins". *ChemBioChem*. 3 (7): 604–17.

[51] P.W. Kenny (2019) The nature of ligand efficiency. *Journal of Cheminformatics* 11(1):8.

[52] I. Kufareva & R. Abagyan (2012). Methods of protein structure comparison. *Methods in Molecular Biology* 857: 231–257.

[53] O. Carugo (2003). How root-mean-square distance (r.m.s.d.) values depend on the resolution of protein structures that are compared. *Journal of Applied Crystallization* 36, 125-128.

[54] E. Fuglebakk, J. Echave & N. Reuter (2012). Measuring and comparing structural fluctuation patterns in large protein datasets. *Bioinformatics*, 28 (19): 2431–2440.

[55] L. Mlu, N.S. Bogatyreva & O.V. Galzitskaia (2008). Radius of gyration is indicator of compactness of protein structure. *Molekuliarnaia Biologiia (Mosk)*. 42(4):701-6.

[56] Z. Sun, L. Qian, Q. Ge, F. Yan & T.R. Manfred (2019). Utility of B-Factors in Protein Science: Interpreting Rigidity, Flexibility, and Internal Motion and Engineering Thermostability. *Chemical Reviews* 2019, 119, 3, 1626-1665.

[57] C.C. David and J.D. Jacobs (2014). Principal Component Analysis: A Method for

Determining the Essential Dynamics of Proteins. *Methods in Molecular Biology* 1084: 193–226.

[58] A.J. McCammon and S.C. Harvey (1986) *Dynamics of Proteins and Nucleic Acids*, Cambridge University Press, Cambridge.

[59] R. Peluso, F. Manguso, M. Vitiello, S. Iervolino & M. Nicola Dario Di Minno (2015) Management of arthropathy in inflammatory bowel diseases. *Therapeutic Advances in Chronic Diseases*. 6(2): 65–77.

[60] Y. Liang, W. Xu, S. Liu, J. Chi, J. Zhang, A. Sui, L. Wang, Z. Liang, D. Li, Y. Chen & H. Niu (2018) N-Acetyl-Glucosamine Sensitizes Non-Small Cell Lung Cancer Cells to TRAIL-Induced Apoptosis by Activating Death Receptor 5. *Cell Physiological Biochemistry*. 45:2054–2070.

[61] Y. Zhanwu, J. Yinghua & L. Hongxu (2017). Anti-lung cancer effect of glucosamine by suppressing the phosphorylation of FOXO. *Molecular Medicine Reports*. Volume 16 Issue 3.

[62] B. Liu, M. Xin, L. Chao, G. Yan-Yan, L. Ning, N. Xiao-Fang, G. Yifu & W. Hua-Qin (2011). Glucosamine induces cell death via proteasome inhibition in human ALVA41 prostate cancer cell. *Experimental & Molecular Medicine* 43: 487–493

[63] M. Koshioka, T. Nishijima & H. Yamazaki (1996). Endogenous gibberellins in the immature seeds of okra (*Abelmoschus esculentus*) *Journal of Plant Physiology*. 149 (1–2): 129-132

[64] V.A. Smith (1992). Gibberellin A1 Biosynthesis in *Pisum sativum* L. II. Biological



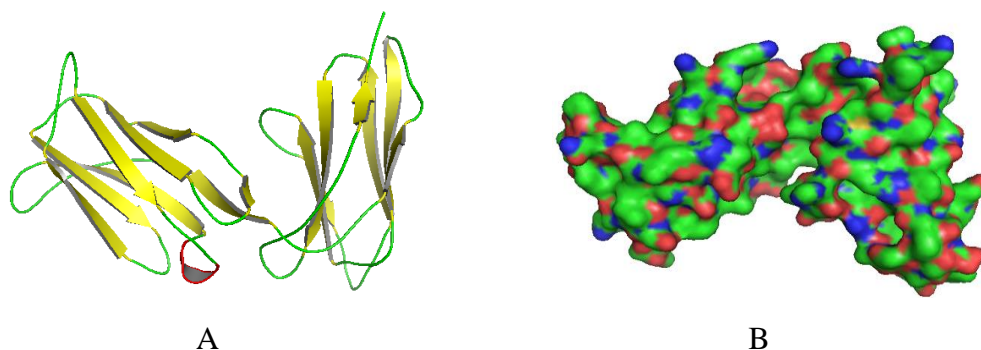
and Biochemical Consequences of the *le* Mutation. *Plant Physiology* 99(2): 372–377

[65] P. Vidhyasekaran (2014). Gibberellin Signaling in Plant Innate Immunity. *Plant Hormone Signaling Systems in Plant Innate Immunity* pp 383-401

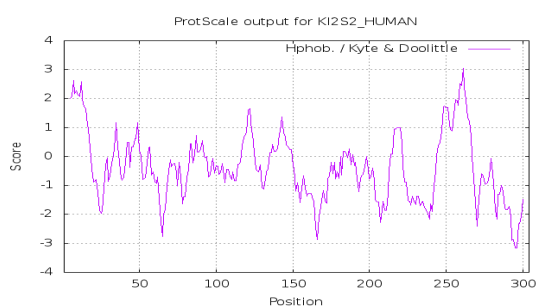
[66] X. Lin, C. Yajuan, C. Jingbo, Z. Hongbin, L. Yedan, Z. Yonghong, Z. Ling & Q. Chen (2019). Anti-tumor effects and

mechanism of GA-13315, a novel gibberellin derivative, in human lung adenocarcinoma: an in vitro and in vivo study. *Cellular & Molecular Biology Letters* 24 Article number: 6

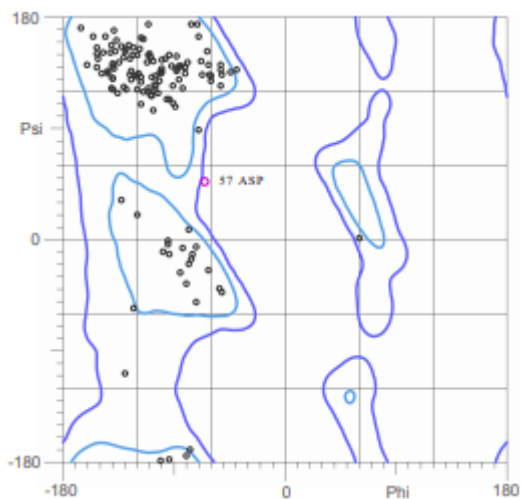
[67] C. Jingbo, Z. Sun, Y. Zhang, X. Zeng, C. Qing, J. Liu, L. Li & H. Zhang (2009). Synthesis of gibberellin derivatives with anti-tumor bioactivities. *Bioorganic & Medicinal Chemistry Letters*. 19 (18): 5496-5499.



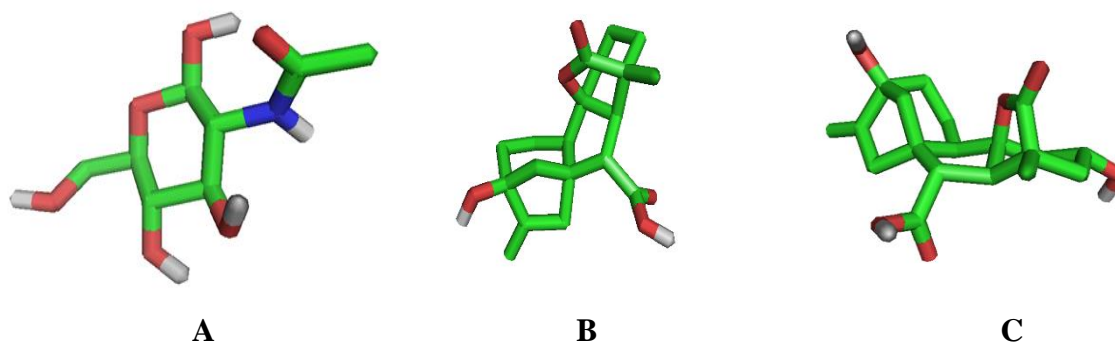
**Figure 1:** a: Cartoon model of the crystal structure of KIR2DS2 (PDBID: 1m4k): Beta sheets (yellow), Alpha helix (red) and Loops (green) b: Surface representations.



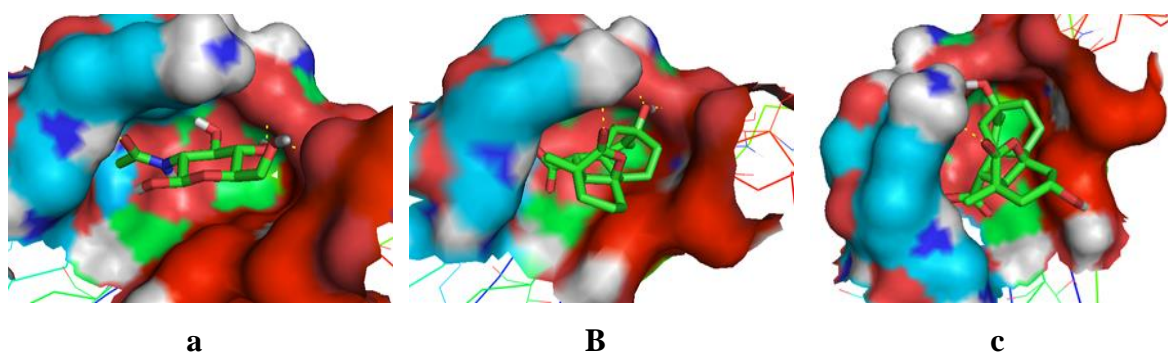
**Figure 2:** Kyte & Doolittle hydrophobicity plot of KIR2DS2 (PDBID: 1m4k)



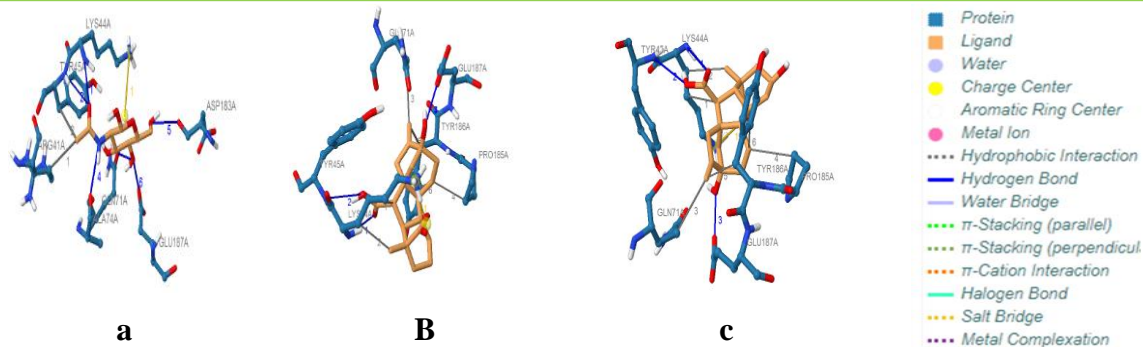
**Figure 3:** Ramachandran plot for KIR2DS2 (PDBID: 1m4k)



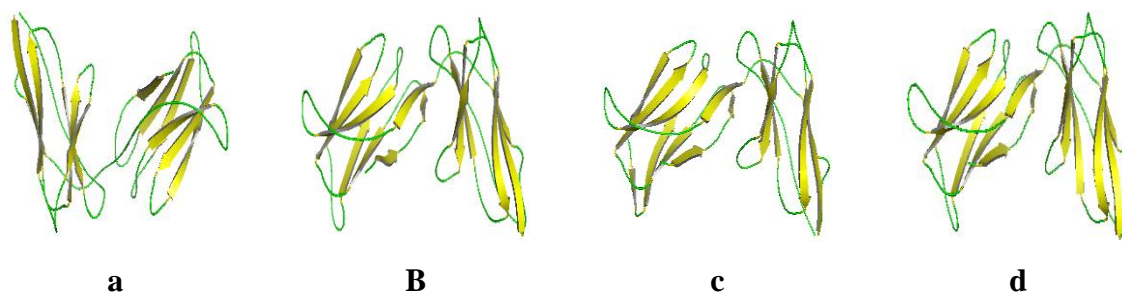
**Figure 4:** The 3D chemical structures (stick model) of standard and lead compound a: NAG b: GA20 c: GA29



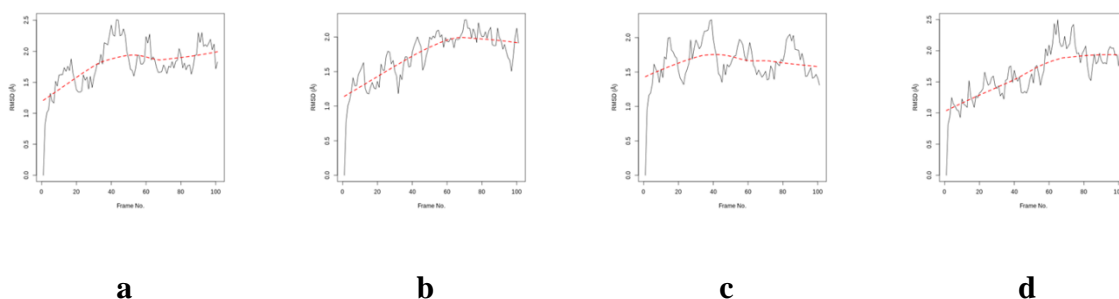
**Figure 5:** Binding site of KIR2DS2 interacting with standard and lead compounds a: KIR2DS2-NAG b: KIR2DS2-GA20 c: KIR2DS2-GA29.



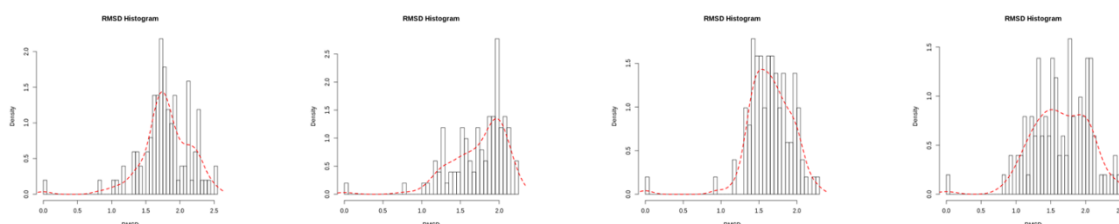
**Figure 6:** Protein-Ligand interactions of KIR2DS2 with standard and lead compound a: KIR2DS2-NAG complex b: KIR2DS2-GA20 complex c: KIR2DS2-GA29 complex



**Figure 7:** Cartoon model of the crystal structure of KIR2DS2 Apo and Holo structures (without water and ions) after molecular dynamics simulation. Beta sheets (yellow), Alpha helix (red) and Loops (green) a: KIR2DS2; b: KIR2DS2-NAG c: KIR2DS2 -GA20 d: KIR2DS2-GA29

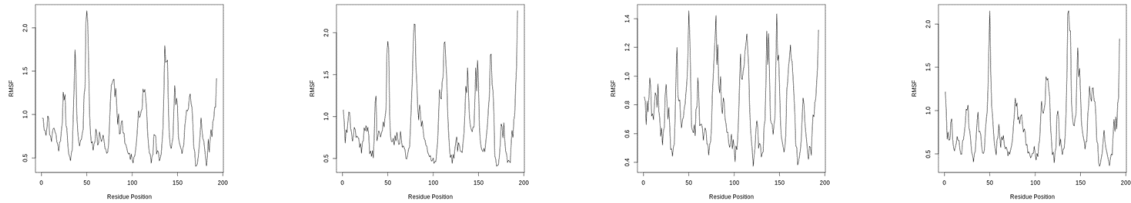


**Figure 8:** RMSD for Apo and Holo structures (a). KIR2DS2 (b) KIR2DS2-NAG (c) KIR2DS2-GA20 (d) KIR2DS2-GA29



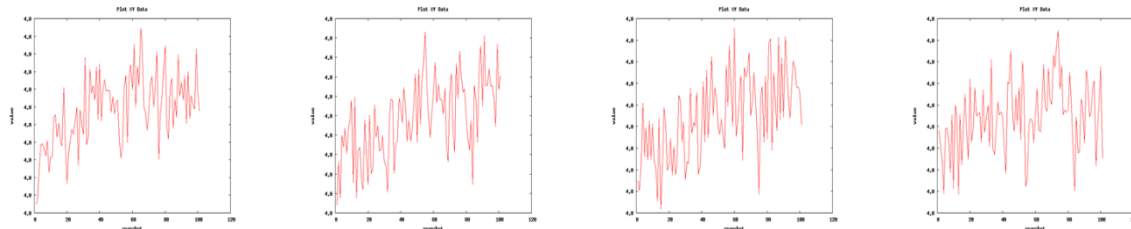
**A** **B** **C** **D**

**Figure 9:** RMSD histogram for Apo and Holo structures (a). KIR2DS2 (b) KIR2DS2-NAG (c) KIR2DS2- GA20 (d) KIR2DS2-GA29



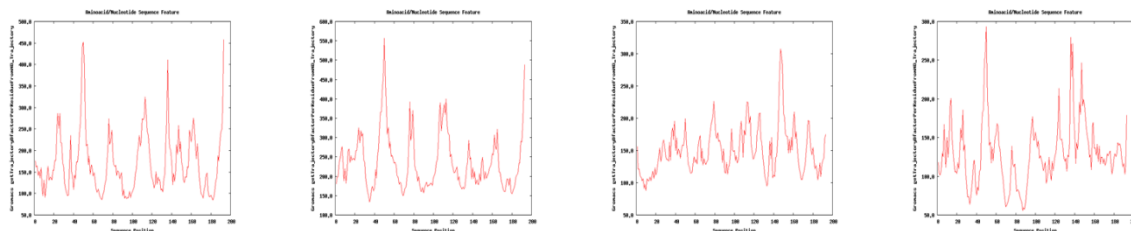
**A** **B** **C** **D**

**Figure 10:** Per-residue RMSF for Apo and Holo structures (a). KIR2DS2 (b) KIR2DS2-NAG (c) KIR2DS2- GA20 (d) KIR2DS2-GA29



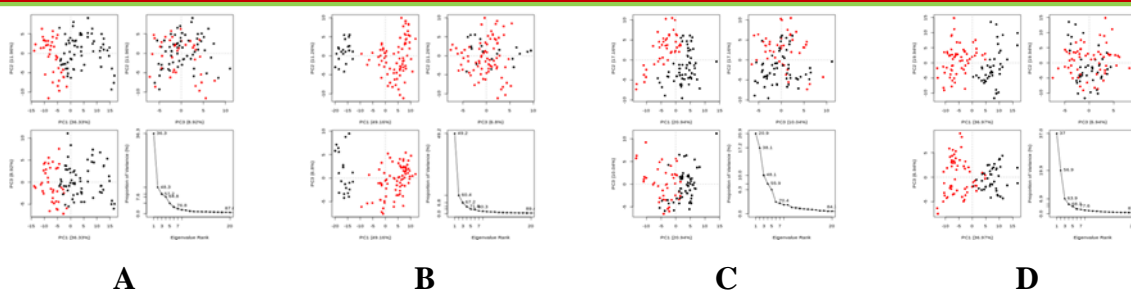
**A** **B** **C** **D**

**Figure 11:** Radius of gyration for Apo and Holo structures (a). KIR2DS2 (b) KIR2DS2-NAG (c) KIR2DS2- GA20 (d) KIR2DS2-GA29

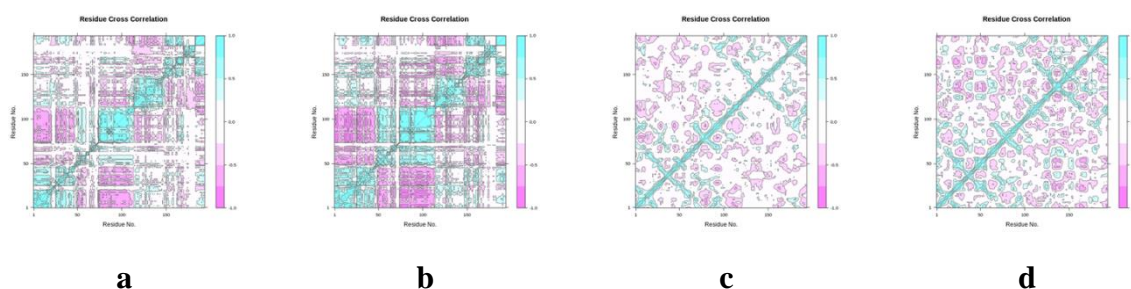


**A** **B** **C** **D**

**Figure 12:** B Factor for Apo and Holo structures (a). KIR2DS2 (b) KIR2DS2-NAG (c) KIR2DS2- GA20 (d) KIR2DS2-GA29



**Figure 13:** Principal component analysis cluster plot of Apo and Holo structures. The projection of trajectory onto 1st few eigenvectors for: (a). KIR2DS2 (b) KIR2DS2-NAG (c) KIR2DS2- GA20 (d) KIR2DS2-GA29



**Figure 14:** Dynamic cross correlation map Apo and Holo structures of 1m4k Purple represents anti-correlated, dark cyan represents fully correlated while white and cyan represents moderately and uncorrelated respectively. 1.0= correlated; 0 is non-correlated; and -1 is anti-correlated. (a). KIR2DS2 (b) KIR2DS2-NAG (c) KIR2DS2-GA20 (d) KIR2DS2-GA29

**Table I: Chemo-informatic properties of standard and lead compounds**

	<b>NAG (standard)</b>	<b>GA20</b>	<b>GA29</b>
Molecular Formula	C8H15NO6	C19H24O5	C19H24O6
Molecular Weight (g/mol)	221.21	332.4	348.4
Log P	-1.7	1.2	0.2
Hydrogen Bond Acceptors	5	5	6
Hydrogen Bond Donors	5	2	3
# heavy atoms	15	24	25
# rotatable bonds	2	1	1
TPSA (A <sup>a</sup> )	119	83.8	104
Molar Refractivity	47.19	86.18	87.34

Saturation (fraction csp <sup>3</sup> )	0.88	0.79	0.79
PAIN Alert	No	No	No
GCPR ligand	-0.47	0.22	0.24
Ion channel modulator	-0.30	0.23	0.20
Kinase Inhibitor	-0.71	-0.21	-0.24
Nuclear receptor ligand	-0.74	0.49	0.60
Protease inhibitor	0.07	0.09	0.19
Enzyme inhibitor	0.46	0.30	0.42

**Table II: Pharmacokinetic properties of lead compounds**

	<b>NAG</b>	<b>GA20</b>	<b>GA29</b>
Water solubility (log mol/L)	-1.38	-2.636	-2.659
Caco2 permeability (log Papp in 10 <sup>-6</sup> cm/s)	-0.219	1.186	0.691
Human Intestinal absorption (% Absorbed)	31.963	98.911	71.42
Skin Permeability (log Kp)	-3.234	-2.735	-2.735
P-glycoprotein substrate (Yes/No)	No	No	No
P-glycoprotein I inhibitor (Yes/No)	No	No	No
P-glycoprotein II inhibitor (Yes/No)	No	No	No
VDss (human) (log L/kg)	0.041	-0.829	-0.82

Fraction unbound (human) (Fu)	0.856	0.418	0.466
BBB permeability (log BB)	-0.618	-0.209	-0.601
CNS permeability (log PS)	-3.694	-2.998	-3.112
CYP2D6 substrate (Yes/No)	No	No	No
CYP3A4 substrate (Yes/No)	No	Yes	Yes
CYP1A2 inhibitor (Yes/No)	No	No	No
CYP2C19 inhibitor (Yes/No)	No	No	No
CYP2C9 inhibitor (Yes/No)	No	No	No
CYP2D6 inhibitor (Yes/No)	No	No	No
CYP3A4 inhibitor (Yes/No)	No	No	No
Total Clearance (log ml/min/kg)	0.711	0.417	0.42
Renal OCT2 substrate (Yes/No)	No	No	No
AMES toxicity (Yes/No)	No	No	No
Max. Tolerated dose (human) (log mg/kg/day)	1.944	0.371	0.262
hERG I inhibitor (Yes/No)	No	No	No
hERG II inhibitor (Yes/No)	No	No	No
Oral Rat Acute Toxicity (LD50) (mol/kg)	1.547	2.051	2.097
Oral Rat Chronic Toxicity (log mg/kg_bw/day)	3.406	2.135	2.498
Hepatotoxicity (Yes/No)	No	No	No

Skin Sensitization (Yes/No)	No	No	No
<i>T.Pyriformis</i> toxicity (log ug/L)	0.285	0.285	0.285
Minnow toxicity (log mM)	4.705	1.958	2.694

**Table III: Molecular docking scores and Ligand Efficiency Metrics of ligands against KIR2DS2**

Ligand	Binding (Kcal/mol) affinity	LE	LLEP	LLE
NAG	-5.70	0.38	-4.47	7.4
GA20	-7.40	0.31	3.87	6.2
GA29	-8.20	0.33	0.61	8.0

**Table IV: Hydrogen bond analysis**

Compound	No. of bonds	Residues	Distance (H-A)	Distance (D-A)	Bond angle
NAG	6	LYS 44	2.32	2.88	113.53
		TYR45	2.65	3.15	110.26
		GLN 71	2.37	2.94	115.97
		ALA74	2.15	3.14	161.14
		ASP183	2.29	3.09	136.96
		GLU187	2.01	2.70	129.79
GA 20	4	LYS 44	3.14	3.93	135.58
		TYR45	2.88	3.40	112.2
		GLU187	2.69	3.10	106.90
		GLU187	2.19	3.10	151.21
GA29	3	LYS44	3.15	3.95	136.17
		TYR45	2.90	3.42	111.94



		GLU187	2.71	3.10	105.81
--	--	--------	------	------	--------

**Table V: Other Protein-ligand interactions**

	Hydrophobic Int.		Salt bridge	
	Residue	Distance	Residue	Distance
NAG	ARG41	3.78	LYS44	5.33
	TYR 45	3.62		
GA20	LYS44	3.93	LYS44	3.67
	LYS44	3.47		
	GLN71	3.81		
	PRO185	3.57		
	TYR186	3.71		
GA29	TYR186	3.24		
	LYS44	3.95	LYS44	3.71
	LYS44	3.54		
	GLN71	3.76		
	PRO185	3.57		
	TYR186	3.75		
	TYR186	3.24		

**Table VI: Summary of data from Molecular Dynamic Simulation of Apo and Holo structures**

MDS Parameters	KIR2DS2	KIR2DS2+NA G	KIR2DS2+GA2 0	KIR2DS2+G A29
<b>RMSD</b>				
Total RMSD	180.94	174.91	166.41	165.22
Average RMSD	1.79	1.73	1.65	1.64
Lowest RMSD	0.00	0.00	0.00	0.00
Highest RMSD	2.51	2.25	2.26	2.50
Timeframe of Highest RMSD	43	70	39	65
Timeframe of Lowest RMSD	1	1	1	1
<b>RMSD Peak Distribution</b>				
0.00 – 0.49A	1	1	1	1
0.50 – 0.99A	1	1	1	3
1.00 – 1.49A	12	20	27	33
1.50 – 1.99A	60	55	62	41
2.00 – 2.49A	25	24	10	22
2.50 – 2.99A	2	0	0	1

<b>RMSF</b>				
Total global RMSF	165.32	170.97	147.41	157.00
Average global RMSF	0.86	0.89	0.76	0.81
Total regional (Res 42-107) RMSF	57.35	58.35	50.30	49.52
Average regional (Res 42-107) RMSF	0.87	0.88	0.76	0.75
Local RMSF (Res 45)	0.78	0.94	0.78	0.75
Least fluctuation	0.40	0.41	0.37	0.36
Highest fluctuation	2.20	2.26	1.46	2.15
Range of RMSF	1.80	1.85	1.09	1.79
<b>Radius of Gyration</b>				
Average Gyration	4.0377	4.0377	4.0349	4.0329
Minimum Gyration	4.0270	4.0248	4.0264	4.02575
Maximum Gyration	4.0469	4.0426	4.0431	4.04089
Range of Gyration	0.0199	0.0178	0.0167	0.01514
% of Gyration	0.49	0.44	0.41	0.38
Timeframe of maximum Gyration	65	55	60	74
Timeframe of minimum Gyration	1	1	15	13
<b>B-Factor</b>				
Global average B-Factor	170.12	236.89	150.26	130.13
Regional (Res 42-107) average B-Factor	166.78	248.35	150.38	123.78
Local B-Factor (Res 45)	191.44	315.63	138.38	154.77
<b>PCA</b>				
Total motions PCA1	11.70	12.01	12.48	11.43
Average motions PCA1	0.06	0.06	0.06	0.06
Best global conformation	PC1	PC2	PC3	PC3
Best local conformation (Res. 45)	PC1 & PC3	PC3	PC3	PC3
Average motions at Res. 45	0.06	0.06	0.07	0.05
PC1 Eigenvalue	36.30%	49.16%	20.94	36.97
PC2 Eigenvalue	12.00%	11.26%	17.16	19.94
PC3 Eigenvalue	8.90%	6.80%	10.04	6.94
Total	57.21%	67.22%	48.14	63.85
PC1 Cosine content	0.91	0.07	0.85	0.53
PC2 Cosine content	0.84	0.06	0.69	0.31
PC3 Cosine content	0.47	0.12	0.03	0.29

- A preliminary version of this work is available on the bioRxiv pre-print server.