

# Virtual screening, molecular dynamics and binding energy-MM-PBSA studies of natural compounds to identify potential EcR inhibitors against *Bemisia tabaci* Gennadius

Harmilan Kaur Mangat

Punjab Agricultural University

Manisha Rani

Punjab Agricultural University

Rajesh Kumar Pathak (✉ [rkpathakbt@gmail.com](mailto:rkpathakbt@gmail.com))

Punjab Agricultural University <https://orcid.org/0000-0003-3835-8430>

Inderjit Singh Yadav

Punjab Agricultural University

Divya Utreja

Punjab Agricultural University

Pardeep Kumar Chhuneja

Punjab Agricultural University

Parveen Chhuneja

Punjab Agricultural University

---

## Research Article

**Keywords:** Bemisia tabaci, insect growth regulators (IGRs), molecular dynamics simulation, natural compound, pest control, virtual screening, whitefly.

**Posted Date:** May 18th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-269641/v2>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

**Version of Record:** A version of this preprint was published at PLOS ONE on January 21st, 2022. See the published version at <https://doi.org/10.1371/journal.pone.0261545>.

# Abstract

Whitefly (*Bemisia tabaci* Gennadius) is a homopteran phyto polyphagous sucking insect pest which is an important pest of cotton that causes economic losses to the crop by reducing its yield and quality. Ecdysteroids like 20-hydroxy ecdysone (20-E), have significant role in larval moulting, development, and reproduction in pterygota insects. Intending to obstruct these fundamental, developmental physiological processes, the receptor of 20-E, the Ecdysone Receptor (BtEcR) of *Bemisia tabaci* has been targeted. To identify potent inhibitors of BtEcR, 98,072 natural compounds were retrieved from ZINC database. A structure-based virtual screening of these compounds was conducted for their binding to BtEcR and the top two compounds (ZINC08952607 and ZINC04264850) were selected based on minimum binding energy. 50 ns molecular dynamics simulation (MDS) study was then performed for the prediction of dynamics and stability of BtEcR and top-scoring ligand-BtEcR complexes. Besides, g\_mmpbsa tool was used to calculate and analyse binding free energy of BtEcR-ligand complexes. The results revealed that ZINC08952607 and ZINC04264850 had binding free energy of – 170.156 kJ/mol and – 200.349 kJ/mol, respectively. Thus, these ligands can be utilized as lead compounds for the development of environmentally safe insecticides against the whitefly.

## 1. Introduction

Cotton (*Gossypium spp.*), popularly known as 'White Gold', is one of the world's most valuable commercial fibre and oil crop. It has a dominating role in the industrial and agricultural economy. In India, it plays important role in the national economy and employs more than eight million people. The main bottleneck in cotton cultivation is the yield losses due to biotic stresses like insect-pests' attack, diseases, and crop-weed competition. Major insect-pests of cotton include bollworms, foliage feeders and sucking pests, which cause serious damage to different parts of the plant all through the cropping season. With the introduction of Bt cultivars of cotton having resistance against cotton bollworms, sucking pests especially the whitefly, have emerged as serious problems (<https://atariz1.icar.gov.in/pdf/Report-KVKs-at-Farmers-Doorsteps.pdf>).

The whitefly (*Bemisia tabaci* Gennadius) is homopteran insect belonging to family *Aleyrodidae*. It is a polyphagous insect that feeds on more than 500 plant species including field crops, vegetables, fruits, oil seed crops, ornamental plants, etc. It is a highly destructive pest that causes huge economic losses every year by reducing crops yields and quality. This pest also acts as a vector for several viral pathogens in cultivated plants (Byrne 1990; Schuster et al. 1995; Oetting and Buntin 1996; Kumar et al. 2020).

On cotton crop, the damage caused by whitefly is due to direct feeding and through lint contamination by the development of sooty mould honeydew on its excrete (Kedar et al. 2014). The whitefly is also the vector of Cotton leaf curl virus (CLCuV). Due to lack of effective genetic resistance, farmers occasionally had to rely on insecticides for efficient management of whitefly. However, in certain situations, highly effective integrated pest management (IPM) strategy has been developed, disseminated and ensured its adoption that has given effective control of the pest. Supervised use of chemical insecticides using economic threshold level (ETL) is one component under IPM (Kumar et al. 2020).

Several insecticides are approved and recommended for this pest but *B. tabaci* has enormous potential to develop resistance against the insecticides. Besides, these chemical insecticides are also hazardous for the environment and injurious to other non-target living organisms (Kumar et al. 2020). So there is need to search for natural molecules with novel selective insecticidal action against *B. tabaci*.

In whitefly and other arthropods, the external cuticle is rigid and inexpandable. Therefore, these animals have evolved a moulting process that facilitates growth and required morphological changes. To enable the co-ordination of growth with environmental signals, moulting is under the regulation of steroid hormones like the ecdysteroids - $\alpha$ -ecdysone ( $\alpha$ E) and its

biologically active form, 20-hydroxy ecdysone (20-E). Receptor for ecdysone, the Ecdysone Receptor (EcR) regulates larval moulting, metamorphosis and reproduction (Auwerx et al. 1999; Riddiford et al. 2000).

Ecdysone receptor is a Ligand-dependent transcription factor which belongs to the superfamily of Nuclear Receptors (Koelle et al. 1991; Mangelsdorf et al. 1995; Freedman 1997). Ecdysone receptor is found only in invertebrates and serves as a key regulator of the gene expression during the development and reproduction of most invertebrates (Riddiford et al. 2000). The functional unit of the receptor is a noncovalent heterodimeric complex of two proteins Ecdysone Receptor (EcR) and another nuclear receptor family member, Ultraspiracle Protein (USP) (Auwerx et al. 1999; Yao et al. 1992; 1993; de Oliveira et al. 2019). EcR and USP have a similar modular domain structure as that of all other nuclear receptors, constituting a well-conserved DNA binding domain (DBD) and a somewhat conserved ligand-binding domain (LBD) (Koelle et al. 1991). The Ecdysone binding site is present in the LBD of the EcR subunit, but USP must be dimerised with EcR for high-affinity binding of ligand. EcRs heterodimerize with USPs to attain transcriptional activity in the presence of ecdysteroids (Nakagawa et al. 2009). Based on literature study as its role in pathophysiology, it can be used as a potential target for the identification of environmentally safe insecticides to specifically obstruct the molting process of insects, thus controlling insect damage.

Recognition of compounds that can act as significant inhibitors of target pathogenic/insect proteins, specific to that pest and optimization of these inhibitors to enhance their affinity, efficiency, selectivity, bioavailability, and metabolic stability are the steps involved in the modern drug/agrochemical discovery programme.

High-throughput screening (HTS) is a tedious process due to the availability of a large number of compounds (Shukla et al. 2018). Therefore, a faster equivalent i.e. structure-based virtual screening can be used for screening of potential inhibitors that can interact with EcR receptor of *B. tabaci* and obstruct its development and reproduction process. In our study, we have performed structure based virtual screening, molecular dynamics simulation (MDS), and binding free energy analysis of natural compounds available at the ZINC database to identify potential inhibitors against *Bemisia tabaci* EcR (Irwin et al. 2012).

## 2. Materials And Methods

### 2.1. Retrieval and preparation of target protein structure

The structure of *B. tabaci* ecdysone receptor (BtEcR) (PDB ID: 1Z5X) was retrieved from RCSB-Protein Data Bank. Using Autodock tools (Carmichael et al. 2005; Morris et al. 2009), the target protein was prepared. Hydrogen atoms were added and a partial atomic charge (Kollman charge) was assigned to the BtEcR structure. The resultant structure was saved in .pdbqt format.

### 2.2. Retrieval and preparation of ligand structures

The natural compounds' subset was downloaded from the ZINC database. The 98,072 natural compounds were retrieved in the .sdf file format. Openbabel was used to convert all sdf file to pdb file format. Further, python script was used to prepare pdbqt file of all selected natural compounds for their interaction study with BtEcR through virtual screening by AutoDock vina (Trott et al. 2010).

### 2.3. Structure based virtual screening

Virtual screening is a lead discovery technique which plays an important role in the discovery of potential compounds. It can screen large ligand databases or compound collections using the residues present in the active site of an enzyme. It enables the screening of large data sets to find the structures possessing the tendency to bind with the target. AutoDock vina is a molecular docking tool which determines the preferred relative orientation of ligand during its docking on protein, to form a stable complex with minimum binding energy. AutoDock vina is used to calculate the interaction energies between ligands and protein (Trott et al. 2010). In this study, the natural compounds were screened against the BtEcR

structure and the complexes with favourable energy were selected. PyMol (<https://pymol.org/2/>) was used to generate these ligand-protein complex structures and Ligplot was used to visualise them (Laskowski et al. 2011).

#### 2.4. Physiochemical properties prediction

For evaluation of drug-likeness, physico-chemical properties of the top 10 natural ligands were retrieved from ZINC database and the remaining unknown properties were calculated using MarvinSketch software (<http://www.chemaxon.com/products/marvin/marvinsketch>). Total eight parameters were estimated, these include Molecular weight (MW), logP, Hydrogen-Bond donor, Hydrogen-Bond acceptor, Polar Surface Area (2D), Polarizability, Van der Waal's Surface Area (3D), and Refractivity (Pathak et al. 2017).

#### 2.5. Molecular dynamics simulation (MDS) and Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) calculation

The MDS studies were carried out using GROMACS 2018.1 for the prediction of the stability of BtEcR and top-scoring ligand-BtEcR complexes. Three systems were generated and subjected to 50ns MDS studies, one for estimating the stability of the BtEcR and other two for BtEcR-ligand complexes. The stability of the selected protein ligand complex is predicted in the presence of a solvent. All three systems were solvated in a cubic box using a simple point charge model. Ligand topology was generated using the ProDRG server (Schüttelkopf et al. 2004). Protein topology was generated using the GROMOS 9653a6 force field (Oostenbrink et al. 2004). The systems were neutralized by adding 8 Na<sup>+</sup> ions. Steepest energy minimization was carried out for all the systems to obtain the maximum force below 1000 kJ/mol/nm to eliminate the steric hinderances. Particle Mesh Ewald method was used to calculate long-range electrostatic interactions (Darden et al. 1993). 1.0 nm radius cut-off was used for the computation of Lennard-Jones, and Coulomb interactions. For constraining the H-bond lengths, the LINCS algorithm was used (Hess et al. 1997). The time step was maintained at 2 fs for the simulation. 10 Å cut-off distance was used to predict the short-range non-bonded interaction, while for long-range electrostatics, the PME method was used with 1.6 Å Fourier grid spacing. All bonds including H-bond were fixed by the Shake algorithm (Ryckaert et al. 1977). The systems were equilibrated after energy minimization. Then position restraint simulation was carried out under NVT and NPT conditions for maintaining the volume, temperature and pressure. Finally, all systems were subjected to a 50 ns molecular dynamics simulation. The 2 fs interval was given for saving the coordinates. Root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), radius of gyration (Rg) and H-bonds were calculated by gmx rms, gmx rmsf, gmx gyrate and gmx hbond. Principal component analysis (PCA) was done by gmx covar and gmx anaeig. Binding free energy was calculated using the g\_mmpbsa tool (Kumari et al. 2014). Finally, xmgrace was used for the generation and visualization of the plots (Turner 2005).

### 3. Results And Discussion

#### 3.1. Generation of grid box for virtual screening

The grid box of BtEcR was generated with centre X = 25.114, Y = 72.57, Z = 4.445 and size X = 46, Y = 48, Z = 44 by AutoDock tool and define in configuration file, besides; the energy range and exhaustiveness were kept 4 and 8, respectively. The generated grid box size defined in configuration file was utilized to carry out interaction studies of BtEcR with natural compounds through virtual screening.

#### 3.2. Structure-based virtual screening

For structure-based virtual screening, a subset of natural compounds (n = 98,072) was downloaded from the ZINC database and the structure of target protein- BtEcR from RCSB PDB. All the ligands and target were prepared using Autodock tools. The pdbqt file of BtEcR and all the selected natural ligands were subjected to virtual screening by AutoDock vina using python script available at vina website. The top 10 ligands were selected based on minimum binding energy which ranged between - 13.4 Kcal/mol and - 12.9 Kcal/mol.

#### 3.3. Visualization and analysis of top 10 docked complexes

Various interactions such as H-bond, hydrophobic interactions and the amino acid residues involved in H-bonding of these complexes were then visualized using Ligplot and are given in Table 1. Both top two natural ligands ZINC08952607 and ZINC04264850, have binding affinity of -13.4 kJ/mol. The ZINC08952607 forms four hydrogen bonds with amino acid residues of EcR, three with Thr231 having bond length 2.84 Å, 2.92 Å, and 2.57 Å and one with Ile227 having bond length 3.18 Å. However, ZINC04264850 forms two hydrogen bonds with EcR amino acid residues one with each Asn390 and Glu199, having bond lengths 3.1 Å and 3.94 Å, respectively (Fig. 1). From these results, it can be concluded that both these ligands form stable complexes with the target protein.

Table 1

Binding free energy, number of H-bonds formed, amino acid residues involved in H-bonding, and bond lengths for the top 10 docked complexes.

Sr.No.	Ligand	Binding Energy (kcal/mol)	No. of H-Bonds	Amino acid residues involved in H-bonding	Bond Length (Å)
1.	ZINC08952607	-13.4	4	Thr231-O Thr231-O Thr231-O Ile227-O	2.84 2.92 2.57 3.18
2.	ZINC04264850	-13.4	2	Asn390-N Glu199-O	3.1 2.94
3.	ZINC03844966	-13.2	3	Ala286-N Thr393-O Thr304-O	3.0 3.08 2.83
4.	ZINC04027828	-13.1	3	Arg275-N Leu284-O Glu199-O	2.91 3.94 2.91
5.	ZINC08952608	-13.1	2	Thr231-O Thr231-O	2.95 2.77
6.	ZINC68589458	-13.0	3	Thr234-O Thr234-O Thr231-O	2.9 2.83 2.99
7.	ZINC08878289	-13.0	3	Thr234-O Thr234-O Thr231-O	2.91 2.79 3.03
8.	ZINC33927923	-13.0	2	Thr393-O Thr304-O	2.99 3.1
9.	ZINC04082366	-13.0	1	Asn390-N	3.04
10.	ZINC19938609	-12.9	3	Asn390-N Thr231-O  Ala283-N	2.99 2.71  3.26

### 3.4 Calculation and analysis of physicochemical properties

Lipinski's rule of five states that to exhibit good ADME (absorption, distribution, metabolism and excretion) properties, a drug should have molecular weight less than 500 daltons, logP value less than 5, Hydrogen bond donors less than 5, and Hydrogen bond acceptors less than 10 (Lipinski et al. 1997). The top two selected natural ligands ZINC08952607 and ZINC04264850 have molecular weight 448.607 and 430.614, logP value 4.97 and 4.253, number of Hydrogen-bond donors 2 and 2, and number of Hydrogen-bond acceptors 3 and 4, respectively. Also, the polar surface area less than 140 Å is an indication of good cell membrane permeability of a drug and refractivity ranging from 40 to 130 indicates a better drug or agrochemical properties. ZINC08952607 and ZINC04264850 show polar surface area as 66.65 and 78.76, Polarizability 51.55 and 47.10, Van der Waal's surface area 643.52 and 570.78, and refractivity 127.34 and 118.55, respectively (Table.2). These results show that both the ligands ZINC08952607 and ZINC04264850 possess the physico-chemical properties of a good agrochemical and can be used as potential insecticide candidates against whitefly and other related insects (Fig. 2).

Table 2

Physicochemical properties i.e. Molecular Weight, LogP, Hydrogen bond donor, Hydrogen bond acceptor, Polar surface area (2D), Polarizability, Van der Waal's surface area (3D) and Refractivity of top 10 docked ligands.

Sr. No.	Ligand	Molecular weight (g/mol)	log P	H-bond donor	H-bond acceptor	Polar surface area (2D)(Å)	Polarizability	Van der Waal's surface area (3D)(Å)	Refractivity
1.	ZINC08952607	448.607	4.97	2	3	66.65	51.55	643.52	127.34
2.	ZINC04264850	430.614	4.253	2	4	78.76	47.10	570.78	118.55
3.	ZINC03844966	418.368	3.085	0	6	139.22	41.70	498.20	134.59
4.	ZINC04027828	396.502	4.593	2	3	49.69	43.08	551.83	110.58
5.	ZINC08952608	448.607	4.97	2	3	66.65	51.55	652.86	127.34
6.	ZINC68589458	408.457	3.639	0	3	57.69	47.28	519.18	117.16
7.	ZINC08878289	448.519	3.293	0	6	76.15	47.37	569.44	76.15
8.	ZINC33927923	391.559	5.376	1	3	54.35	44.83	614.96	116.69
9.	ZINC04082366	429.645	5.386	1	3	46.61	49.18	683.76	123.23
10.	ZINC19938609	407.466	4.311	1	5	68.23	43.86	548.94	113.78

### 3.5. Molecular dynamics simulations

All-atom time-dependent MDS is a technique that allows us to unravel the structural dynamics, conformational behavior, and stability of the protein and protein-ligand complexes. The field of structure-based drug discovery has been revolutionized with the advent of MDS. To understand the dynamic perturbations taking place in BtEcR during ligand binding, the MDS study was conducted. It also enabled us to predict the stability of the protein-ligand complexes (BtEcR-ZINC08952607, and BtEcR- ZINC04264850). In this study, we calculated the RMSD, RMSF, Rg, and hydrogen bonds of the protein-ligand complexes. Also, the principal component analysis and binding free energy analysis were also performed. RMSD analysis revealed that the entire trajectory was equilibrated after 30 ns. Therefore, all the parameters mentioned above were calculated for the last 20ns trajectory.

#### 3.5.1. Root mean square deviation (RMSD)

RMSD is the measure of the deviation of the protein backbone from its initial structural conformation to its final conformation. Deviations produced throughout the simulation determine the stability of the protein concerning its structural conformation. A stable protein structure shows smaller deviations in the protein backbone and vice versa. RMSD value for the C $\alpha$  backbone of all three systems was calculated for the last 20ns simulation. Figure 3(A) shows the plot of RMSD (nm) vs. time (ps) for BtEcR, BtEcR-ZINC08952607, and BtEcR-ZINC04264850. The average RMSD values for BtEcR, BtEcR-ZINC08952607, and BtEcR-ZINC04264850 were found to be 0.376, 0.281, and 0.315nm respectively. Both BtEcR-ZINC08952607 and BtEcR-ZINC04264850 showed lower RMSD value as compared to BtEcR, suggesting greater stability of BtEcR-ligand complexes than BtEcR.

### 3.5.2. Root mean square fluctuation (RMSF)

RMSF analysis describes the flexible regions of the protein-ligand complexes. In proteins, higher RMS fluctuation is shown by loosely organized structures like loops, turns, and coils while well-structured regions such as  $\alpha$ -helix and  $\beta$ -sheets show low RMS fluctuation. We have calculated the RMSF value for prophesying the structural changes induced in protein structure by ligand binding. Figure 3(B) shows the RMSF plots for BtEcR, BtEcR-ZINC08952607, and BtEcR-ZINC04264850. The average RMSF value for BtEcR, BtEcR-ZINC08952607, and BtEcR-ZINC04264850 were recorded as 0.109, 0.104, and 0.114 nm respectively. BtEcR-ZINC08952607 showed less RMSF value than BtEcR, which implies that upon ligand binding (ZINC08952607), the flexibility of catalytically significant residues of BtEcR is reduced. Hence, it can be concluded that ZINC08953607 can be a potent inhibitor of the catalytic activity of BtEcR.

### 3.5.3. Radius of gyration (Rg)

The radius of gyration indicates the compactness level of the protein structure before and after the binding of the ligand. Figure 3(C) shows the Rg plot for BtEcR, BtEcR-ZINC08952607 and BtEcR-ZINC04264850 with average Rg values as 1.777, 1.813, 1.800 nm, respectively. BtEcR-ZINC04264850 showed slightly lesser Rg as compared to the other complex, indicating that on binding with BtEcR, it formed a relatively stable complex.

### 3.5.4. Hydrogen bonds analysis

The hydrogen bond interactions are transitory and play an important role in stabilization of protein-ligand complex. Figure 3 (D) shows the H-bonding plot for BtEcR-ZINC08952607, and BtEcR-ZINC04264850. The complex BtEcR-ZINC08952607 formed an average of 0–5 hydrogen bonds while complex BtEcR-ZINC04264850 formed an average of 0–3 hydrogen bonds. The overall result of hydrogen bonds implies that both the ligands form stable complexes with the protein. The ligand ZINC0852607 forms more H-bonds with BtEcR, indicating that it forms a more stable complex.

### 3.5.5. Principal component analysis (PCA)

PCA was carried out for the prediction of large concerted motions during ligand binding. It is known that the first few eigenvectors describe the overall motions of the protein. So we used the first 50 eigenvectors to calculate the significantly correlated motions for the last 20 ns simulations. Eigenvalues were predicted after diagonalisation of the covariance matrix of atomic fluctuations. Figure 3(A) shows the plot of eigenvalues in descending order vs the corresponding eigenvector for BtEcR, BtEcR-ZINC08952607, and BtEcR-ZINC04264850. The first 10 eigenvectors exhibit 77.06, 70.85 and 75.67 percent motions for BtEcR, BtEcR-ZINC08952607 and BtEcR-ZINC04264850, respectively. This indicates that the protein and protein-ligand complexes show different proportion of motions. Hence, it was inferred that binding of the ligand causes the protein conformation and dynamics to change. BtEcR-ligand complexes show lesser correlated motions as compared to BtEcR, which implies that ligand binding leads to complex stabilization. From the results, it was concluded that the BtEcR-ZINC08952607 complex is exhibiting lesser correlated motions and is more stable than BtEcR-ZINC04264850. From Fig. 4(A), it is clear that the first few eigenvectors play an important role in overall motions. So we used the first two eigenvectors for prediction of 2D projection plot for better representations of the results. The 2D projection plot representing the protein motions in a phase space was plotted for BtEcR, BtEcR-ZINC08952607, and BtEcR-ZINC04264850 (Fig. 4(B)). From the plot, it is evident that EcR-ZINC08952607 forms a more stable cluster as compared to BtEcR-ZINC04264850. Figure 5 shows the Gibbs free energy for PC1 and PC2. The plot shows energy ranging between 0 to



8.43, 0 to 8.66, and 0 to 8.55 kJ/mol for BtEcR, BtEcR-ZINC08952607, and BtEcR-ZINC04264850, respectively. BtEcR-ZINC04264850 complex showed lesser energy as compared to BtEcR-ZINC08952607, which implies that the complex followed energetically more favourable transition from one conformation to another. The BtEcR-ZINC04264850 complex showed a higher blue coloured region indicating lower energy, and hence suggesting that this complex is thermodynamically more stable.

### 3.5.6. MM-PBSA: Calculation and analysis of binding free energy

Binding free energy is the total of all the non-bonded interactions. It was estimated for BtEcR-ZINC08952607 and BtEcR-ZINC04264850, using MM-PBSA method. The interaction energies like, Van der Waal’s energy, polar solvation energy, electrostatic energy, SASA energy, and binding energy were calculated for last 5ns of MD trajectory and are tabulated in Table.3. The binding free energy for BtEcR-ZINC08952607 and BtEcR-ZINC04264850 is – 170.156 kJ/mol and – 200.349 kJ/mol, respectively. From the results, it is evident that these compounds show a significant binding affinity with BtEcR, but ZINC04264850 shows a higher binding affinity as compared to ZINC08952607.

**Table.3.** Van der Waal’s, electrostatic, polar salvation, SASA and binding energy in kJ/mol for top two complexes.

Sr. No.	Compound	Van der waal’s energy (kJ/mol)	Electrostatic energy (kJ/mol)	Polar solvation energy (kJ/mol)	SASA energy (kJ/mol)	Binding energy (kJ/mol)
1	ZINC 08952607	-247.973 +/- 11.225	-26.227 +/- 5.427	127.589 +/- 9.382	-23.545 +/- 1.049	-170.156 +/- 12.658
2	ZINC04264850	-241.724 +/- 10.213	-13.819 +/- 5.831	78.048 +/- 7.531	-22.854 +/- 1.061	-200.349 +/- 10.772

## 4. Conclusion

Cotton whitefly (*Bemisia tabaci*) is a highly destructive polyphagous pest that causes huge economic losses, inter alia in cotton by direct feeding and lint contamination. The repeated use of limited number of effective chemical insecticides would results in insecticide resistance, limiting their use over longer time. In the current study, we have used a series of *in-silico* approaches to investigate efficient natural lead by targeting BtEcR, which is the receptor of moulting hormone, the Ecdysone. We have demonstrated that ZINC08952607 and ZINC04264850 have the potential to be developed as natural insecticides against *B. tabaci*. Based on RMSD, RMSF, Rg, number of H-bonds, PCA, Gibb’s free energy and binding free energy, we have concluded that both compounds may be used to validate their insecticidal activity against *B. tabaci* in laboratory and filed conditions for crop protection.

## Declarations

### Acknowledgements

Authors wish to acknowledge Bioinformatics Centre (Sub DIC under BTISNet programme of the Department of Biotechnology, Govt. of India), School of Agricultural Biotechnology, Punjab Agricultural University, India for providing all the necessary facilities.

### Conflicts of interest

All authors declare that they have no conflict of interest.

## Ethics approval

This article does not contain any studies with human participants or animals performed by the authors.

## Funding

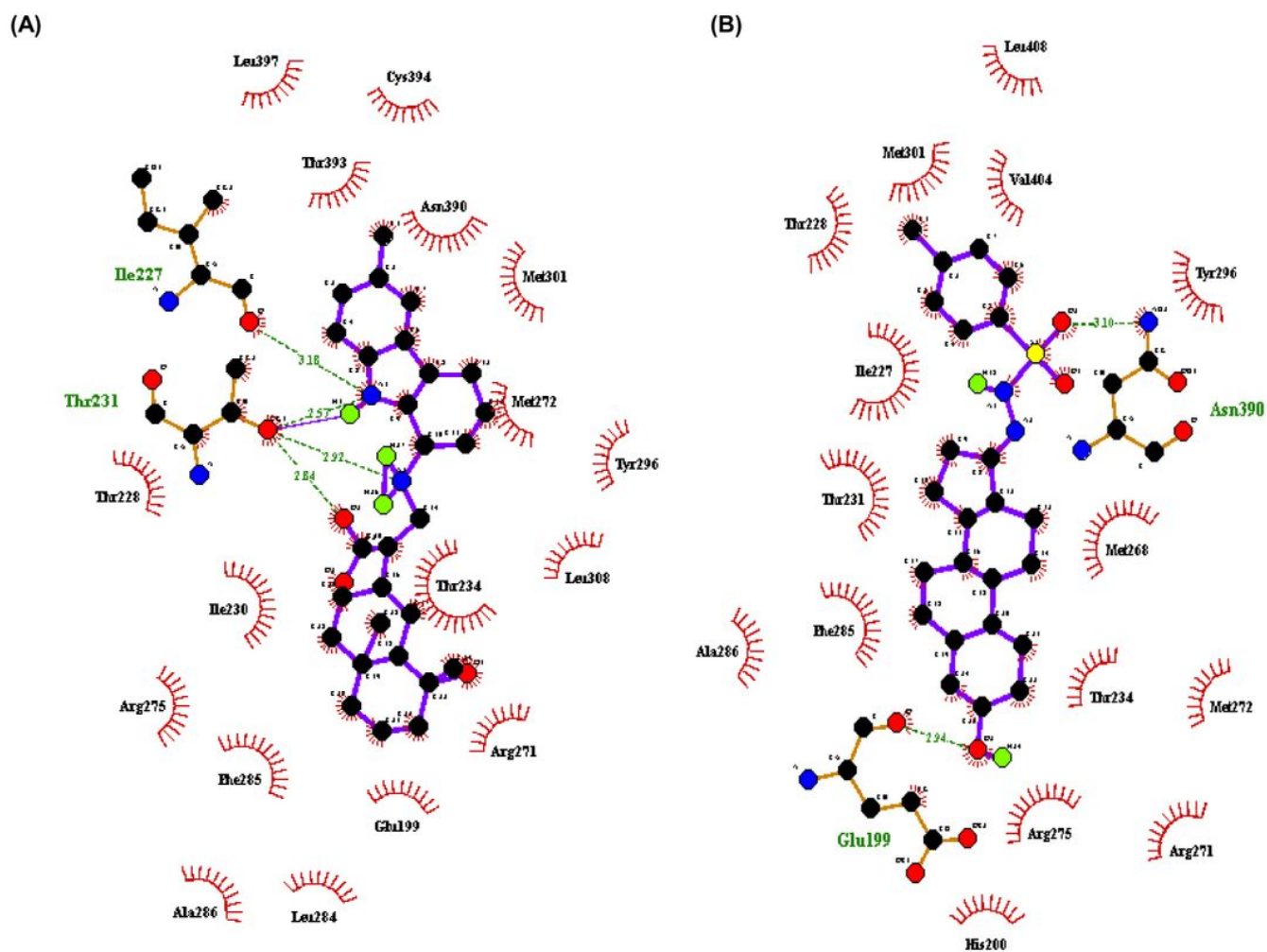
Not applicable

## References

1. Byrne, D. N. Whiteflies in agricultural systems. *Whiteflies: Their bionomics, pest status and management*, 227-261 (1990).
2. Oetting, R. D., & Buntin, G. D. Bemisia damage expression in commercial greenhouse production. *Bemisia: 1995, taxonomy, biology, damage, control and management*. Gerling D and RT Mayer (Eds.). Intercept Ltd., Andover, UK , 702 (1996).
3. Schuster, D. J., Stansly, P. A., & Polston, J. E. Expressions of plant damage by Bemisia. *Bemisia: 1995, taxonomy, biology, damage, control and management*. Gerling, D and RT Mayer (Eds.). Intercept Ltd., U. K. 153-165 (1996).
4. Kumar, V., Kular, J. S., Kumar, R., Sidhu, S. S., & Chhuneja, P. K. Integrated whitefly [*Bemisia tabaci* (Gennadius)] management in Bt-cotton in North India: an agroecosystem-wide community-based approach. *Current Science*, 119(4), 618-624 (2020).
5. Kedar, S. C., Saini, R. K., & Kumaranag, K. M. Biology of cotton whitefly, *Bemisia tabaci* (Hemiptera: Aleyrodidae) on cotton. *Journal of Entomological Research*, 38(2), 135-139 (2014).
6. Auwerx, J., Baulieu, E., Beato, M., Becker-Andre, M., Burbach, P.H., Camerino, G., Chambon, P., Cooney, A., Dejean, A., Dreyer, C. and Evans, R.M A unified nomenclature system for the nuclear receptor superfamily. *Cell*, 97(2), 161-163 (1999).
7. Riddiford, L. M., Cherbas, P., & Truman, J. W. Ecdysone receptors and their biological actions. *Horm.* 60, 1–73 (2000).
8. Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schütz, G., Umesono, K., & Evans, R. M. The nuclear receptor superfamily: the second decade. *Cell*, 83(6), 835 (1995).
9. Freedman, L. P. *Molecular Biology of Steroid and Nuclear Hormone Receptors* Birkhouser, Boston (1997).
10. Koelle, M. R., Talbot, W. S., Segraves, W. A., Bender, M. T., Cherbas, P., & Hogness, D. S. The Drosophila EcR gene encodes an ecdysone receptor, a new member of the steroid receptor superfamily. *Cell*, 67(1), 59-77 (1991).
11. Yao, T. P., Forman, B. M., Jiang, Z., Cherbas, L., Chen, J. D., McKeown, M. & Evans, R. M. Functional ecdysone receptor is the product of EcR and Ultraspiracle genes. *Nature*, 366(6454), 476-479 (1993).
12. Yao, T. P., Segraves, W. A., Oro, A. E., McKeown, M., & Evans, R. M. Drosophila ultraspiracle modulates ecdysone receptor function via heterodimer formation. *Cell*, 71(1), 63-72 (1992).
13. de Oliveira, D. A. B., da Silva, A. V., & dos Santos Niculau, E. Molecular Docking of Azadirachtin in Nuclear Ecdysone Receptor. *Current Physical Chemistry*, 9(1), 50-57 (2019).
14. Nakagawa, Y., & Henrich, V. C. Arthropod nuclear receptors and their role in molting. *The FEBS journal*, 276(21), 6128-6157 (2009).
15. Shukla, R., Shukla, H., Kalita, P., Sonkar, A., Pandey, T., Singh, D. B. & Tripathi, T. Identification of potential inhibitors of Fasciola gigantica thioredoxin1: computational screening, molecular dynamics simulation, and binding free energy studies. *Journal of Biomolecular Structure and Dynamics*, 36(8), 2147-2162 (2018).
16. Irwin, J. J., Sterling, T., Mysinger, M. M., Bolstad, E. S., & Coleman, R. G. ZINC: a free tool to discover chemistry for biology. *Journal of chemical information and modeling*, 52(7), 1757-1768 (2012).

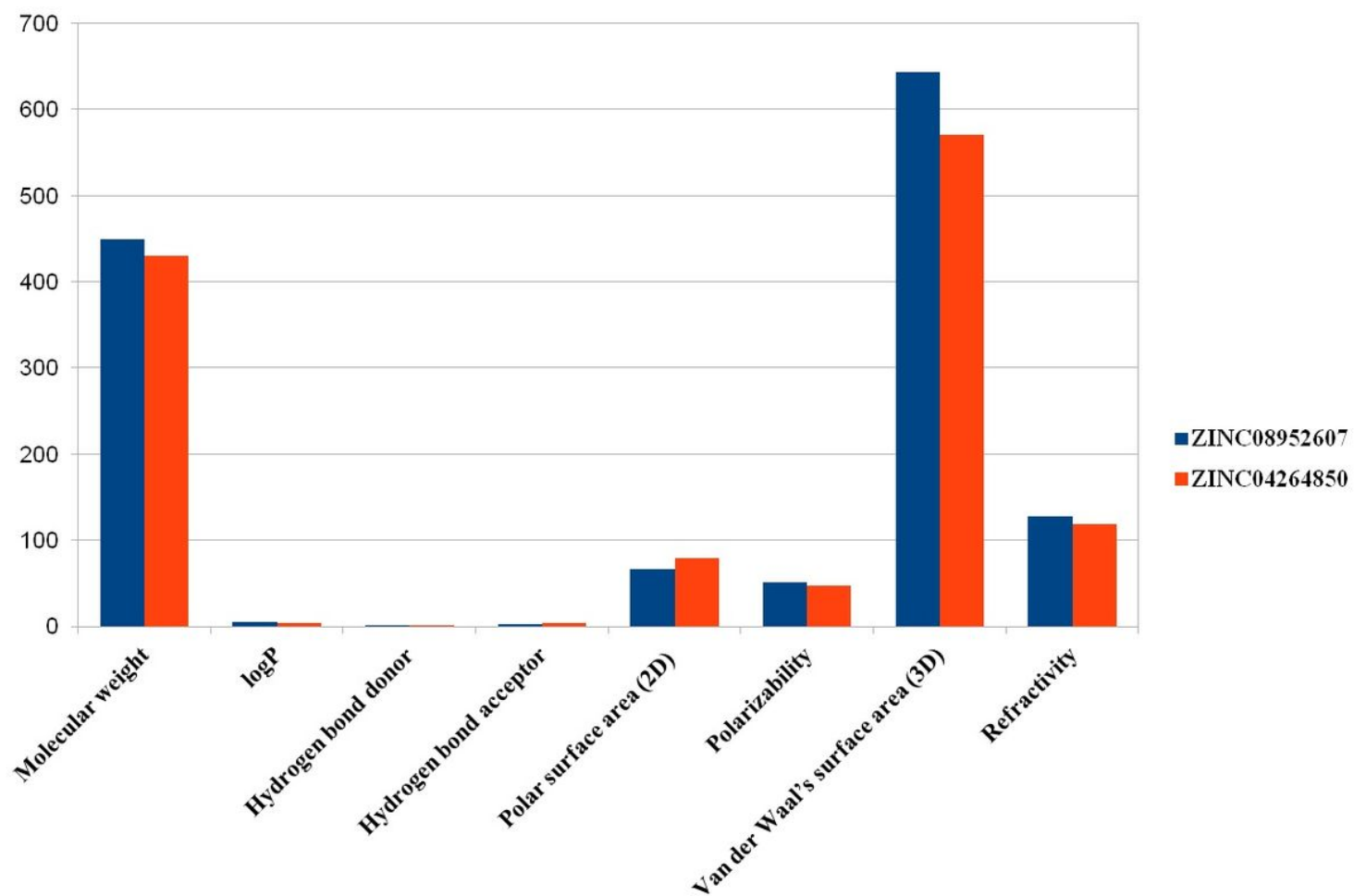
17. Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews*, 23(1-3), 3-25 (1997).
18. Carmichael, J.A., Lawrence, M.C., Graham, L.D., Pilling, P.A., Epa, V.C., Noyce, L., Lovrecz, G., Winkler, D.A., Pawlak-Skrzecz, A., Eaton, R.E. and Hannan, G.N. The X-ray structure of a hemipteran ecdysone receptor ligand-binding domain comparison with a lepidopteran ecdysone receptor ligand-binding domain and implications for insecticide design. *Journal of Biological Chemistry*, 280(23), 22258-22269 (2005).
19. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry*, 30(16), 2785-2791 (2009).
20. Trott, O., & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31(2), 455-461 (2010).
21. Laskowski, R. A., & Swindells, M. B. LigPlot+: multiple ligand–protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.*, **51**, 2778-2786 (2011).
22. Pathak, R. K., Baunthiyal, M., Shukla, R., Pandey, D., Taj, G., & Kumar, A. In silico identification of mimicking molecules as defense inducers triggering jasmonic acid mediated immunity against alternaria blight disease in brassica species. *Frontiers in plant science*, 8, 609 (2017).
23. Schüttelkopf, A. W., & Van Aalten, D. M. PRODRG: a tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallographica Section D: Biological Crystallography*, 60(8), 1355-1363 (2004).
24. Oostenbrink, C., Villa, A., Mark, A. E., & Van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *Journal of computational chemistry*, 25(13), 1656-1676 (2004).
25. Darden, T., York, D., & Pedersen, L. Particle mesh Ewald: An N- log (N) method for Ewald sums in large systems. *The Journal of chemical physics*, 98(12), 10089-10092 (1993).
26. Hess, B., Bekker, H., Berendsen, H. J., & Fraaije, J. G. LINCS: a linear constraint solver for molecular simulations. *Journal of computational chemistry*, 18(12), 1463-1472 (1997).
27. Ryckaert, J. P., Ciccotti, G., & Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of computational physics*, 23(3), 327-341 (1977).
28. Kumari, R., Kumar, R., Open Source Drug Discovery Consortium, & Lynn, A. g\_mmpbsa: A GROMACS tool for high-throughput MM-PBSA calculations. *Journal of chemical information and modeling*, 54(7), 1951-1962 (2014).
29. Turner, P. J. "XMGRACE, Version 5.1. 19." Center for Coastal and Land-Margin Research, *Oregon Graduate Institute of Science and Technology, Beaverton, OR* (2005).

## Figures



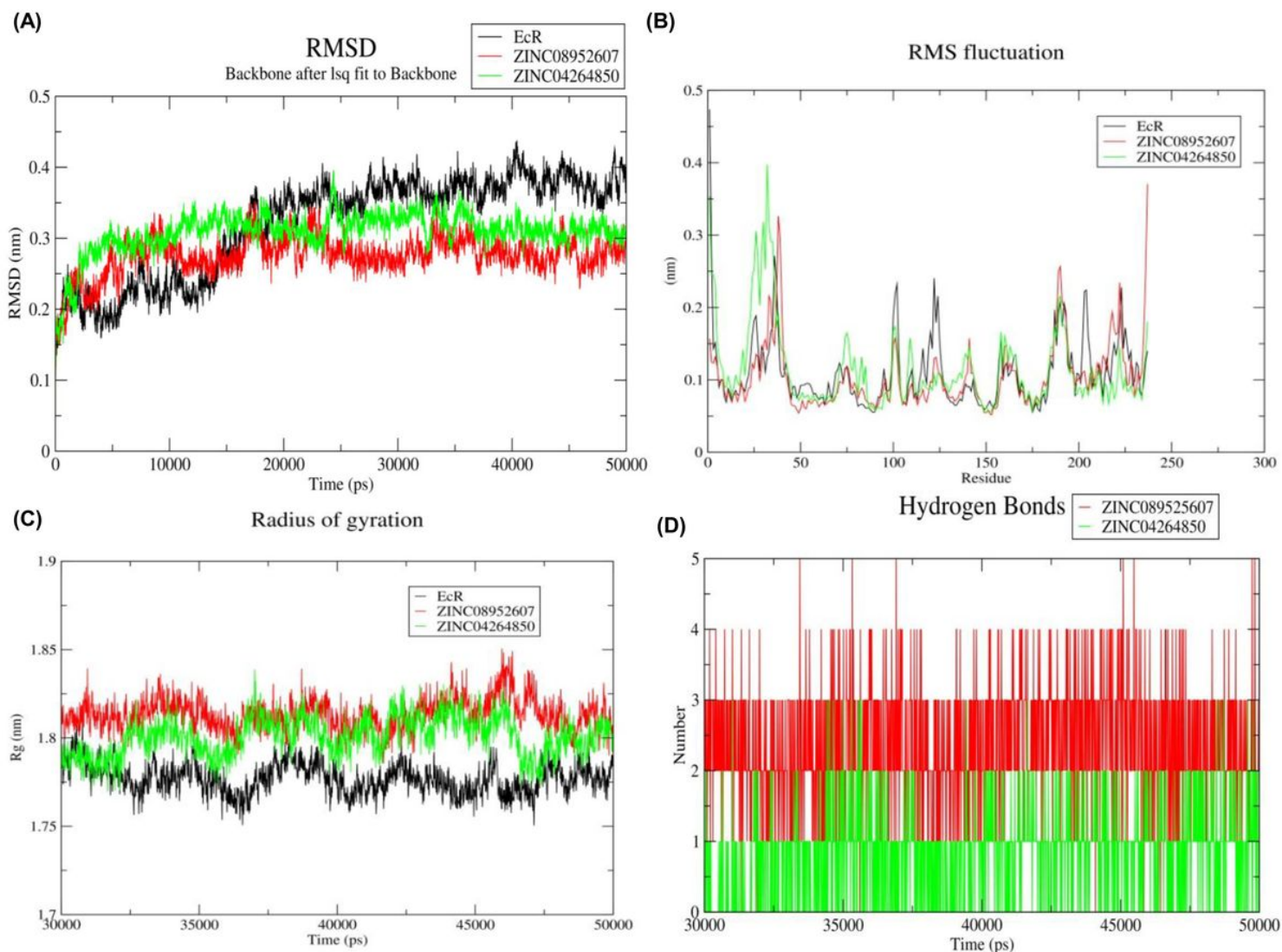
**Figure 1**

Protein-ligand interaction analysis: (A) BtEcR-ZINC08952607 docked complex structure generated through Ligplot showing hydrogen bonds with amino acid residues Thr231 and Ile227 as green lines and hydrophobic interactions as red arcs. (B) BtEcR-ZINC04264850 docked complex structure generated through Ligplot showing hydrogen bonds with amino acid residues Glu199 and Asn390 as green lines and hydrophobic interactions as red arcs.



**Figure 2**

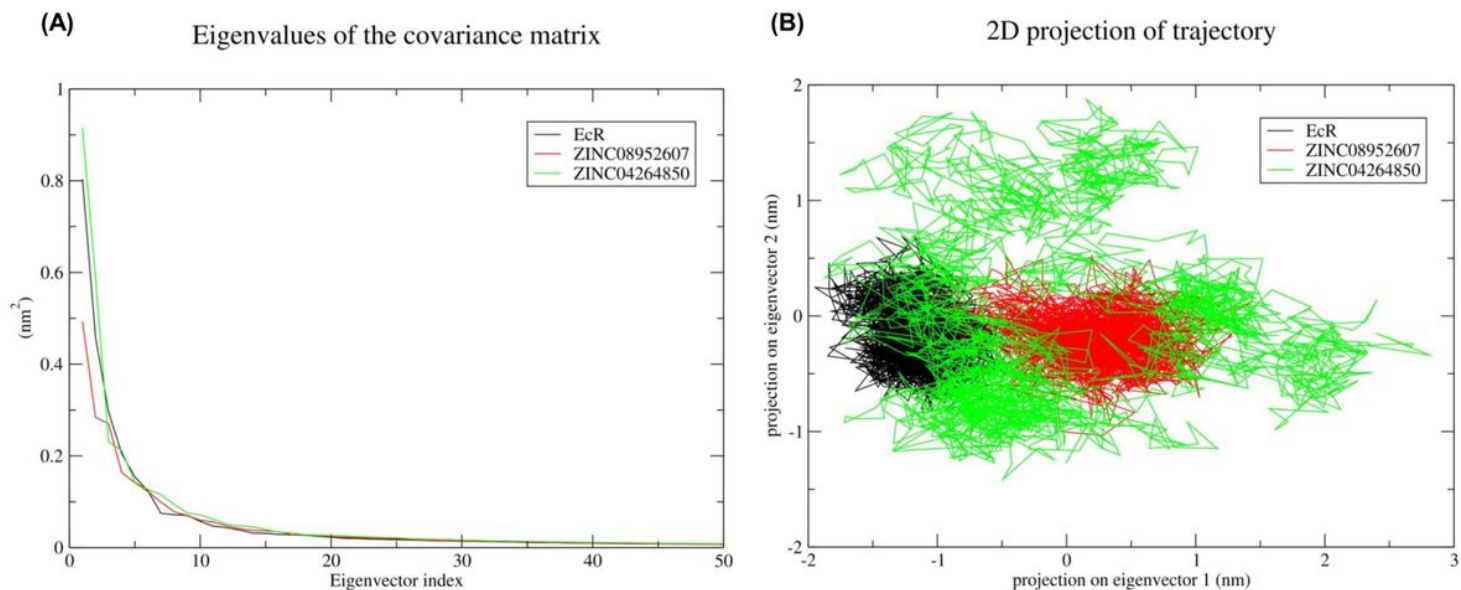
Values of principal descriptors for ZINC08952607 and ZINC 04264850



**Figure 3**

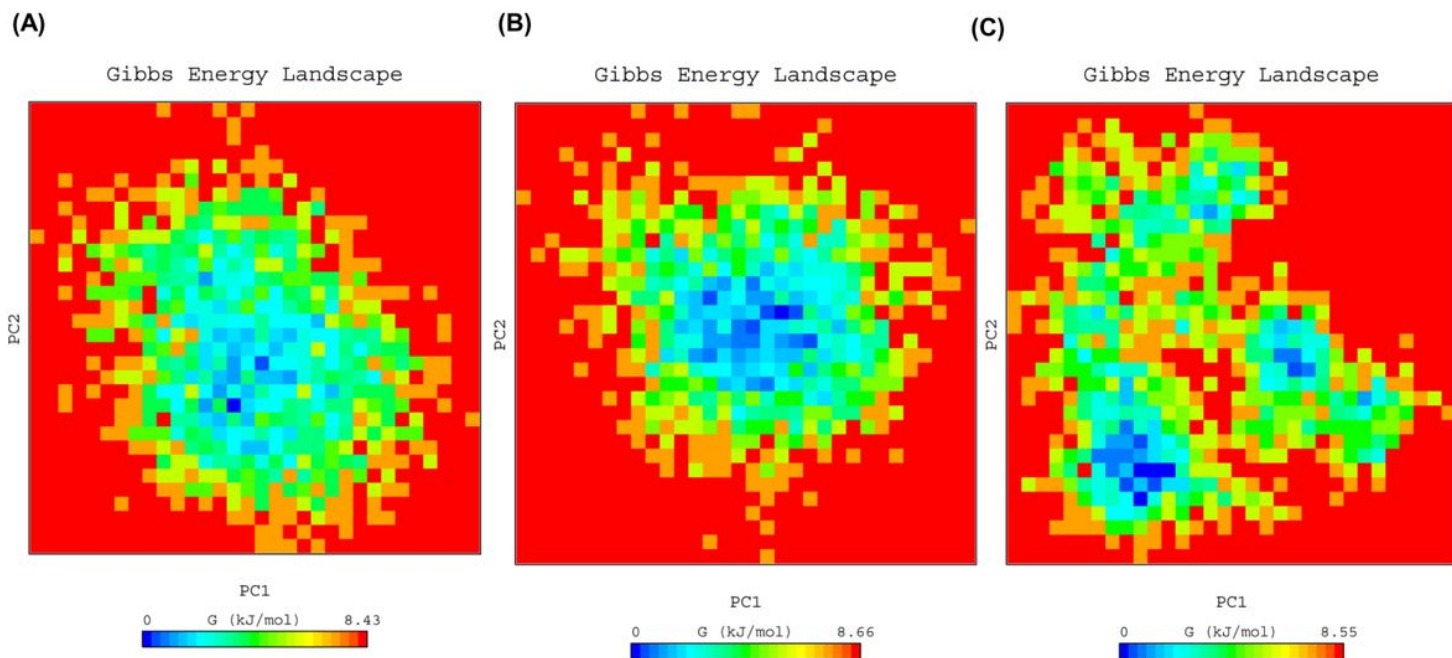
(A) Time dependent RMSD of c- $\alpha$  backbone of the BtEcR, BtEcR-ZINC08952607, and BtEcR-ZINC04264850. (B) The RMSF for c- $\alpha$  atoms of BtEcR, BtEcR-ZINC08952607, and BtEcR-ZINC04264850. (C) Plot of Rg vs time for BtEcR, BtEcR-ZINC08952607, and BtEcR-ZINC04264850. (D) Plot of number of hydrogen bonds vs time for the BtEcR-ZINC08952607, and BtEcR-ZINC04264850. Black, red and green color represents BtEcR, BtEcR-ZINC08952607, and BtEcR-ZINC04264850, respectively.





**Figure 4**

Principal component analysis. (A) Plot of eigen values vs eigenvector index was plotted taking into consideration the first 50 eigen vectors for BtEcR, BtEcR-ZINC08952607, and BtEcR-ZINC04264850. (B) Projection of protein motion in phase space along PC1 and PC2 for BtEcR, BtEcR-ZINC08952607, and BtEcR-ZINC04264850. Black, red and green color represent BtEcR, BtEcR-ZINC08952607 and BtEcR-ZINC04264850 complex, respectively.



**Figure 5**

PC1 and PC2 Gibbs free energy landscape for (A) BtEcR, (B) BtEcR-ZINC08952607, and (C) BtEcR-ZINC04264850