Identification of Human Phosphoglycerate Mutase 1 (PGAM1) inhibitors using hybrid virtual screening approaches

Numan Yousaf[†1], Rima D. Alharthy[†2], Maryam[1], Iqra Kamal[1], Muhammad Saleem[3], Muhammad Muddassar*[1]

[1]Department of Biosciences, COMSATS University Islamabad, Park Road, Islamabad, Pakistan.

[2]Department of Chemistry, Science and Arts College, Rabigh Campus, King Abdulaziz University, Jeddah 21577, Saudi Arabia.

[3]School of Biological Sciences, University of the Punjab, Lahore, Pakistan.

[†] Equal Contribution

Corresponding Author:
Muhammad Muddassar[1]
Department of Biosciences, COMSATS University Islamabad, Park Road, Islamabad, Pakistan

22   Email address: mmuddassar@comsats.edu.pk

Abstract

26   PGAM1 plays a critical role in cancer cell metabolism through glycolysis and different

27   biosynthesis pathways to promote cancer. It is generally known as a crucial target for treating

28   pancreatic ductal adenocarcinoma, the deadliest known malignancy worldwide. In recent years

29   different studies have been reported that strived to find inhibitory agents to target PGAM1,

30   however, no validated inhibitor has been reported so far, and only a small number of different

31   inhibitors have been reported with limited potency at the molecular level. Our insilico studies

32   aimed to identify potential new PGAM1 inhibitors that could bind at the allosteric sites. At first,

33   shape and feature-based models were generated and optimized by performing receiver operating

34   characteristic (ROC) based enrichment studies. The best query model was then employed for

35   performing shape, color, and electrostatics complementarity-based virtual screening of the

36   ChemDiv database. The top two hundred and thirteen hits with greater than 1.2 TanimotoCombo

37   score were selected and then subjected to structure-based molecular docking studies. The hits

38   yielded better docking scores than reported compounds, were selected for subsequent structural

39   similarity-based clustering analysis to select the best hits from each cluster. Molecular dynamics

40   simulations and binding free energy calculations were performed to validate their plausible

41   binding modes and their binding affinities with the PGAM1 enzyme. The results showed that

42   these compounds were binding in the reported allosteric site of the enzyme and can serve as a

43   good starting point to design better active selective scaffolds against PGAM1enzyme.

Introduction

45   In metabolic pathways of cellular respiration, glycolysis is considered as the main step to

46   produce energy in the form of Adenosine triphosphate (ATP). The eighth step of glycolysis is

47   catalyzed by a glycolytic enzyme called Phosphoglycerate Mutase 1 (PGAM1). PGAM1

48   reversibly isomerizes 3-phosphoglycerate (3PG) to 2- Phosphoglycerate (2PG) through 2, 3-

49   bisphosphoglycerate (2BPG) intermediate in order to proceed aerobic glycolysis (Hitosugi et al.

50   2012; Liu et al. 2017). Glycolytic enzyme Phosphoglycerate Mutase 1 (PGAM1) has potential

51   stimulating roles in a number of human cancers. In cancer cells, PGAM1 is highly up regulated

52   that cause rapid production of cancer cells and tumor growth (Xu et al. 2016; Zhang et al. 2017).

53   Human genome contains two PGAM genes thereby two iso-types of dPGM genes exist, which

54   includes PGAM1 or PGAM-B and PGAM2 or PGAM-M. Their encoded products share high

55   degree of homology with 81 % identity, similar length of 254 and 253 amino acids respectively.

56   Moreover, they form homodimers and heterodimer that are functionally identical. The PGAM1 is

57   present in different body parts such as liver, brain, kidney, red blood cells (RBCs), and early fetal

58   skeletal muscle, whereas PGAM2 can be found in adult skeletal muscles and myocardium

59   (cardiac muscles). The MB heterodimer is found in heart and in the late fetal or neonatal

60   muscle(Xu et al. 2014).

61   In humans the active enzyme exists in the form of dimer which is formed by the C strands of two

62   monomers with antiparallel alignment. In the center of the dimer chloride ion is located which

63   strengthens the interactions of the two monomers. Low concentration of chloride ion accelerates

64   dPGM activity. Whereas its higher concentration inhibits dPGM activity, as chloride ion binds at

65   the enzyme's active site which is rich with positively charged residues and compete with the

66   substrate thereby inhibit enzyme's activity(Xu et al. 2014).

67   Currently PGAM1 has only few reported classes of inhibitors that include MJE3, PGMI-004A,

68   and a natural chemical epigallocatechin gallate (EGCG), the major polyphenol derived from

69   Camellia sinensis (green tea)(Evans et al. 2005) and Xanthone derivatives. MJE3 was initially

70   identified as PGAM1 inhibitor but it also inactivates PGAM1 by covalently modifying its K100

71   residue, because of its spiroepoxide substructure(Li et al. 2017). PGMI-004A is a verified

72   PGAM1 inhibitor; it exhibits anticancer properties in mice models carrying human xenograft

73   tumors. These two reported inhibitors have low molecular potency(Liu et al. 2017). Although

74   PGAM1 third inhibitor EGCG has 30 times higher molecular potency than PGMI-004A but due

75   to its polyphenol structure and the off-target effect, it cannot be considered as a good inhibitor(Li

76   et al. 2017). Recently xanthone derivatives (24 *N*-xanthone benzene sulfonamides) have been

77   designed and synthesized as PGAM1 inhibitors. Most of them exhibit high potency against

78   PGAM1 as compared to PGMI-004A and their anti-proliferation activity was moderate on

79   different cancer cell lines(Ma et al. 2014; Wang et al. 2018b). Due to few reported PGAM1

80   inhibitors, there is a need to design and identify better active compounds that can bind to

81   PGAM1 enzyme to block its metabolic activity.

Materials and Methods

Preparation of active, decoys and database compounds.

84   The reported inhibitors of PGAM1 were selected as active compounds dataset. Decoys dataset

85   was prepared using physicochemical properties of the PGMI-004A co-crystal ligand. The criteria

86   such as HBA<9, HBD<3, molecular weight and logP value ±463.38g/mol, ±4 respectively was

87   used to fetch the decoys [18] from ChEMBL database https://www.ebi.ac.uk/chembl/. Decoy

88   molecules exhibit the similar physicochemical properties like active molecules; however, they

89   are chemically distant and act as false positive results in validation studies. Chemdiv database

90   was prepared by implementing drug like filter criteria to remove unwanted chemical structures,

91   duplicates and salts. Stereoisomers of molecules were checked and tautomeric states were

92   generated at pH 7.0. OMEGA tool was used to generate one hundred conformers per molecule.

Query model generation and optimization

94   Shape and color based query models were built using vROCS tool (Poongavanam & Kongsted

95   2013) implemented in OpenEye scientific software https://www.eyesopen.com/ (Grant et al.

96   1996). PGMI-004A co-crystal ligand was used for generating query models which were

97   validated by plotting receiver operating characteristic (ROC) curve (Poongavanam & Kongsted

98   2013). The ROC curves usually discriminate true positive and false positive compounds in the

99   dataset. In our experiments, the graphs were represented by fraction of the active on Y-axis and

100 fraction of the decoy dataset on X-axis. Different features were iteratively added and removed to

101 obtain the best query model that can distinguish true actives from inactive compounds during

102 virtual screening.

ROCS shape and electrostatic potential similarity-based screening

104 ROCS shape and color-based screening method was used to align the best query model with

105 chemical database compounds on the basis of shape and color matching, the best match was

106 selected based on TanimotoCombo scoring function. Screened hits possessed higher

107 TanimotoCombo score were subjected for electrostatic potential similarity matches using the

108 EON software implemented in OpenEye software. It calculates the partial charges similarity

109   between database and query molecule, and then calculates the ET-Combo score which is the

110   combination of shape-Tanimoto and PB-electrostatic similarity coefficient. Hits having ET-

111   Combo >1 were selected for further screening (Imran et al. 2020).

112   **Protein and ligand structure preparation for molecular docking**

113   Human PGAM1 structure (PDB-ID: 5Y2I) was retrieved from protein data bank

114   (https://www.rcsb.org/)(Berman et al. 2003). The crystal structure of PGAM1 in complex with

115   ligand PGMI-004A was in dimeric form consisting of two chains (Chain A and B). Protein

116   preparation wizard embedded in Schrödinger's Maestro tool (Sastry et al. 2013) was used for

117   preprocessing of structure. All water molecules and other cocrystal ligands were removed and

118   structure was minimized using OPLS2005 forcefield to remove all steric clashes. Similarly, for

119   molecular docking studies compounds were prepared using Ligprep tool embedded in

120   Schrodinger software(Guasch et al. 2013). Different ionization states and tautomeric forms were

121   generated at pH 7.0 (Li et al. 2005).

122   **Grid Generation and Glide Docking**

123   Before docking studies, a Grid box was generated around the allosteric site of the PGAM1

124   structure. The size of the outer box was adjusted up to 17Å and size of inner box was fixed to

125   12Å, 12Å, and 15Å for X, Y and Z axis respectively. The Receptor Grid Generation panel was

126   used to generate the grid at the allosteric sites of the prepared protein. At the end by using Glide

127   SP docking (precision tool from Schrodinger software)(Friesner et al. 2004) prepared ligands

128   were docked within generated grid.

129   **Identification of Toxicological and Pharmacokinetics Properties**

130   Toxicological and Pharmacokinetics studies were conducted on virtually screened compounds

131   mined from Chemdiv database by using online tool PreADMETv.2.0

132   (https://preadmet.bmdrc.kr/adme/) (Lee et al. 2004; Lee et al. 2003). ADME data i.e., Toxicity,

133   Absorption, and Distribution were calculated by means of selected following parameters:

134   Toxicological in Ames Test, Carcinogenicity for Rats and Mouse, Absorption in Human

135   Intestinal Absorption, Cell Permeability, Cell permeability Maden Darby Canine Kidney, In-

136   Vitro P-glycoprotein inhibition, permeability in Caco2 cells and distribution in blood-brain

137   barrier, plasma protein binding(Araújo et al. 2020).

138 **Identification of Biological Activity of Screened Compounds**

139 Biological activities were predicted in order to test the possibility of the biological effects of the

140 selected compounds. Activity predictions were made using the online PASS server

141 (http://www.pharmaexpert.ru/passonline/)(Lagunin et al. 2000) which is the freely available

142 software to predict the biological effects of the compounds based on their chemical structures.

143 PASS server suggests the activity of the compounds based on their chemical formula using

144 multi-level atom neighbors (VMA) descriptors. Activities of compound with following effects

145 had been selected: Antiallergic, antiasthmatic, Alzheimer's disease treatment, and antidiabetic.

146 **Molecular Dynamic (MD) Simulation Method**

147 Molecular Dynamic simulation (MD) was carried out using the NAMD software

148 https://www.ks.uiuc.edu/Research/namd) [31] based on Charm run and molecular dynamic force

149 fields. The parameters of the inhibitors were described by General Amber Force Field (GAFF)

150 and atomic charges were calculated through Antechamber package of Amber tools 21

151 https://ambermd.org/AmberTools.php)(Macke et al. 2010). Subsequently, protein structures of

152 PGAM1 in complex with inhibitors were prepared for simulation using the tLeap module of

153 amber tools 21. The ff14SB force field and TIP3P water model were described for solvent-

154 solvent simulation. Protein ligand complexes were immersed in 10Å octahedral water box and

155 then neutralize by adding Na^+/Cl^- ions. After that, we initialize the system for simulation

156 includes: energy minimization of the complexes by removing undesired contacts between non-

157 bonded adjacent atoms then, the temperature of the system is increased gradually from 200k to

158 250k to 300k in order to stabilize the system(de Oliveira et al. 2019; Jorgensen et al. 1983).This

159 step is done in three further steps with each steps ran for about 5000 steps, at each integration

160 step velocities are reassigned and the temperature is increased. In the final step, the simulation

161 was run in "production" phase at 310K temperature and 1 atm pressure by using the NPT

162 ensemble, for 50 ns in order to study the behavior of proteins after binding with the ligands.

163 **Binding Free Energy Calculation**

164 Binding free energies of the protein-ligand complexes were calculated using molecular

165 mechanics Poison Boltzmann calculations MMGBSA free energy algorithm(Sun et al. 2014).

166 The ΔG_{total} was calculated using following equation:

167
$$\Delta G_{total} = \Delta G_{complex} - [\Delta G_{protein} + \Delta G_{ligand}]$$

168 ΔG_{total} is the sum of interaction energy of the gas phase in between protein-ligand complex.

169 However free energy calculation per residue was calculated by following equation:

170
$$\Delta G_{MMGBSA} = \Delta G_{vdW} + \Delta G_{ele} + \Delta G_{GB} + \Delta G_{Surf}$$

171 **Residues free energy decomposition**

172 In order to find the key residues involved in the binding of protein and ligand, MMGBSA free

173 energy decomposition program of Amber21 was also applied to the complexes. The interaction

174 of ligand and protein can be described by the following equation:

175
$$\Delta G_{inhibitor\_residue} = \Delta G_{vdW} + \Delta G_{ele} + \Delta G_{ele,sol} + \Delta G_{nonpol,\ sol}$$

176 **Computational Alanine Scanning**

177 The key residues involved in the protein-ligand binding free energy were mutated to alanine to

178 find their role in the total binding free energies of the complexes. The mutated complexes were

179 subjected to recalculate the binding free energy. The total binding free energy of the two systems

180 (wild type and mutant) can be calculated by the following equation:

181
$$\Delta\Delta G_{bind} = \Delta G_{bind}^{mutant} - \Delta G_{bind}^{wildtype}$$

182 The flowchart of overall study is shown in Fig 1.

183

184 **Results and Discussion**

185 **Molecular Docking of Reported Compound**

186 Different classes of PGAM1 inhibitors have been reported which include xanthone derivatives,

187 Anthraquinone derivative, MJE3 and Epigallocatechin (EGCG) (Evans et al. 2005; Wang et al.

188 2018a). Molecular docking studies were performed to find the sites in PGAM1 protein at which

189 reported compounds are binding. The analyses showed that among all reported classes xanthone

190 derivatives are binding at the allosteric site with reasonably good glide score values. The binding

191 interactions of protein residues such as hydrogen bonding, pi-pi stacking etc. with xanthone

192 derivatives were observed as reported in co-crystalized structure. The key residues involved in

193 the interactions are Phe22, Arg90, Tyr92 and Arg191(Wen et al. 2019b).

ROCS Query Model Generation and Virtual Screening

195 Since the co-crystal ligand 8kx (xanthone derivative) showed good binding score among other

196 reported compounds. Therefore, it was selected for the development of query model. Finest

197 ROCS query models were produced and optimized in four phases, in first phase, all

198 pharmacophore features were applied for model building. In second phase model was generated

199 by reducing the number of pharmacophoric features to 5A and 4R (A: Hydrogen bond Acceptor,

200 R: Ring), phase 3 have 4A and 3R pharmacophoric features and in last phase query model was

201 built on 3A, 3R features as shown in Fig 2A. The query models were sequentially validated to

202 observe their ability to discriminate the active compounds from the decoys. The dataset used for

203 the validation contained fifteen active compounds and six hundred decoys retrieved from

204 ChEMBL database. The query-1 showed the poor enrichment with Area under curve (AUC)

205 value of 0.78, whereas query 2 and 3 showed better screening performance and yielded AUC

206 0.84 and 0.83 respectively. The query-4 discriminated the active and decoys significantly better

207 than other three queries with an AUC value of 0.87. The ROC plots shown in Fig 2B are

208 representing the fraction of active ligands (the true positives) on the y-axis and the fraction of

209 decoys (the false positives) on the x-axis.

210 After generating optimized query model, the Chemdiv database was screened by matching the

211 query features with its compounds. Since the query contains both shape and color features,

212 therefore it was speculated that unknown compounds having similar characteristics with query

213 will show the biological activities (Kearnes & Pande 2016). Shape and color tanimoto scores

214 were used to rank the overlaying compounds. TanimotoCombo score ranges from 0 to 2 was

215 used to select the hits , which means hits with value closer to 2 will exactly overlay its shape and

216 features to query model, while the value closer to 0 indicates shape and chemical-features are not

217 matching (Kearnes & Pande 2016; Li et al. 2018). Using this scoring scheme, 540 compounds

218 were obtained with TanimotoCombo score ranges from 1.20-1.53. Then these hits were

219 subjected to compare their partial charges complementarity with cocrystal ligand using PB

220 Electrostatic Tanimoto and shape Tanimoto (ET_Combo score) implemented in EON(Rush et al.

221 2005).

Molecular Docking Based Virtual Screening and Clustering

223   To further narrow down the size of initially screened hits from ROCS and EON based

224   approaches, hits were docked at the allosteric site of PGAM-1 target protein. All the hits were

225   sorted based on the glide score and the compounds having glide score comparative to the binding

226   affinity of co-crystal ligand were selected for their structural clustering analysis. The purpose of

227   structure based clustering was to classify the compounds into groups based on structure

228   similarity(Mangiatordi et al. 2017). Based on the glide scores and structural diversity, five

229   compounds (**C1**=6144-0309, **C2**=E470-1348, **C3**=C301-8900, **C4**=F540-0157 and **C5**=L464-

230   0403) were selected from each cluster. Among the top five compounds, **C2** exhibit the good

231   binding pose having highest glide score of -6.902 Fig 3. On comparison, docking score of

232   virtually screened top hits compounds were relatively equal to the selected ligands of PGAM1.

233   Selected compounds also indicated the greater binding interaction within the allosteric site

234   residues. Compound **C1** formed one hydrogen bond and salt bridge with Lys100 and pi

235   interaction with Phe22. The oxygen of sulfonamide group of **C2** was involved in two hydrogen

236   bonds with Arg90, Arg191 while NH group formed one hydrogen bond with Tyr92.

237   Additionally, a pi interaction was observed with Phe22. Similarly, Arg191 and Tyr92 were

238   making hydrogen bonds with sulfonamide oxygen and NH group of compounds **C3**, while Phe22

239   was involved in pi-pi stacking. In complex of PGAM1-**C4**, Phe22 was involved in two types of

240   interactions. It made a hydrogen bond with carbonyl oxygen and pi-pi stacking with the benzene

241   ring. The **C5** compounds was also making the hydrogen bonds with Tyr92 and Arg191 and pi-pi

242   stacking with Phe22 as shown in Fig 4A. All these interactions of the screened compounds

243   demonstrated that these were bound to the reported key residues of the allosteric pocket of

244   PGMA1 enzyme (Wen et al. 2019b). The binding poses of these compounds were also analyzed

245   by comparing with the binding pose of the co-crystal ligand 8kx, which showed that these

246   compounds have similar plausible binding modes as the reference ligand as shown in Fig 4B.

247   **Analysis of Toxicological and Pharmacokinetic Properties**

248   Toxicological and pharmacokinetics properties of the selected hits were assessed using online

249   server PreADMET (https://preadmet.bmdrc.kr/adme/). First the absorption levels of the

250   compounds were predicted using different parameters like HIA (%), CaCo-2 (nm/sec), PMDCK

251   (nm/sec), P-gp Inhibition. All the compounds showed high value of HIA>94%, which is believed

252   the most important ADME properties. CaCo-2 permeability also showed significant results,

253 value greater than two showed easy absorption of the compound. Compound **C1** showed highest

254 22.2815nm/sec absorption rate in all reported inhibitors. Moreover, P_{MDCK} permeability value of

255 **C1** has also seen substantially increased 204.401nm/sec and **C4** have moderate and ideal value

256 of 4.13987nm/sec. Other important drug absorption barrier in the cell is P-glycoprotein (P-gp).

257 Data showed that it assists in the intestinal penetration of compounds except **C4** which dropped

258 the efflux method due to its passive permeability. The Distribution parameter of BBB% was

259 referred to characterize the distribution of the compounds in various tissues in vivo. Compounds

260 have BBB% >0.3 can cross the blood-brain membrane easily(Clark 1999). According to data, all

261 compounds were predicted to be good to cross the blood-brain barrier. It was observed the

262 compounds **C2, C3, and C4** have showed **BBB (%)** values less than one which means that these

263 compounds have no activity on central nervous system (CNS) shown in **Table 1**.

264 Data in **Table 2** illustrates that the values of possibility of biological activity of compounds are

265 higher as compare to possibility of inactive (Pa > Pi) generally in term of antineoplastic,

266 antiallergic and antidiabetic responses. The selected compounds showed Pa>Pi values, indicating

267 the relation with reported biological activities. **C1** possess best value for proteasome ATPase

268 inhibitor (Pa=0.784) and Insulysin inhibitor (Pa=0.637), **C2** (Pa=0.863) have good activity value

269 as Omptin inhibitor, **C3** (Pa=0.338) shows better inhibition against Antineoplastic, **C4**

270 (Pa=0.808) as an Anti-obesity and **C5** (Pa=0.734) for Nootropic disease. All the candidate

271 compounds have good biological activity values for different diseases and can be performed as

272 potential inhibitors against different targets.

273 **MD Simulation studies:**

274 Molecular Dynamic studies of best identified compounds were performed to evaluate the binding

275 modes and stability of selected compounds with PGAM1. MD-simulations studies help to

276 understand the dynamic movement of the protein ligand complex and to calculate their binding

277 free energy. The best fitted compounds (**C1**= 6144-0309, **C2**= E470-1348, **C3**= C301-8900,

278 **C4**= F540-0157, **C5**= L464-0403) in complex with PGAM1 were subjected to MD simulation

279 for 50 ns time scale. Trajectory analysis were performed by calculating RMSD of back bone

280 atoms, RMSF, and Radius of gyration of all the five compounds in complex with PGAM1. The

281 Fig 5, 6, and 7 showed that all compounds (except **C4**) in simulation system were stable,

282 compact and well fitted in the allosteric site. The protein backbone RMSD helps to analyze the

283 protein-ligand complex stability over the simulation time (Sargsyan et al. 2017). The RMSD

284 values of the complexes range from 0 Å 3 Å shown in Fig 5. It can be observed that all

285 complexes showed stability as the RMSD of all complexes remained in the range of ~2.5 to 3 Å

286 after being equilibrated at 5 ns. A small deviation in the RMSD value of PGAM1-**C1** complex

287 was observed during 35 to 40 ns but it gained stability towards the end of simulation. The overall

288 behavior of all complexes showed that protein was remained stable during the simulation.

289 Further, the different snapshots of protein-ligand complexes (Fig 6) were extracted from 0, 5, 10,

290 15, 20, 25, 30, 35, 40, 45, and 50 ns and aligned to detect the position of the docked compounds

291 and analyze the mobile regions of the protein. The aligned snapshots depicted that **C4** compound

292 did not remain at docked site (Fig 6D) while the other compounds were tightly bound to the

293 allosteric site. The elasticity and fluctuations in protein amino acid residues are calculated by

294 measuring the root mean square fluctuations. The higher RMSF value shows that the protein is

295 more flexible and vice versa(Martínez 2015). In RMSF analysis, the C and N terminal and the

296 loop regions gave higher fluctuation in all five complexes as shown in Fig 7. The protein

297 residues did not show major fluctuations during simulation however residues range from 100 to

298 150, showed some minor fluctuations. This region contained the binding sites of protein as well

299 as some loop regions. Besides these, no significant fluctuations were observed at the ligand

300 binding site of PGAM1. Overall PGAM1 in complex with the ligands found stable and no abrupt

301 deviations have been observed during whole simulation. Similarly, no uncertain behavior of the

302 complexes was seen during the 50 ns MD simulation which indicated that the virtual hit

303 compounds were well fitted at allosteric site of PGAM1 enzyme.

304 **Binding Free Energy Studies**

305 Molecular mechanics Generalized Born surface area (MMGBSA) method was used to calculate

306 the total binding free energy (ΔG_{total}) for all five complexes. ΔG_{total} value is usually used to

307 estimate the stability of protein-ligand complex. The lower values of ΔG_{total} indicates that the

308 complex is more stable (Wang et al. 2021) and vice versa. It was computed as a sum of protein-

309 ligand complex and the difference of PGAM1 protein and its ligands free energies. The total

310 binding free energy estimated using MMGBSA model is the outcome of the contribution of

311 various protein-ligand interactions such as van der Waals energy (VDWAALS), electrostatic

312 energy (EEL), EGB (electrostatic contribution to solvation free energy by Generalized Born).

313   The ΔG_{VDW} of **C2, C3 and C5** complexes were found to be -37.42 ± 0.26, -42.36 ± 0.19 and -

314   35.96 ± 0.32 kcal/mol respectively and contributed more in binding affinities. **C1** and **C4** had

315   limited contribution of -29.61 ± 0.17 and -17.79 ± 0.35 respectively. In the case of ΔG_{ELEC} the

316   energy component was -21.66 ± 0.57 in **C1**, -19.64 ± 0.29 in **C2**, -21.39 ± 0.34 in **C3**, -13.24 ±

317   0.72 in **C4** and -15.17 ± 0.50 in **C5**. ΔG_{ELEC} energy contribution of **C1** is highest among all other

318   complexes. After those surface energies were analyzed, **C3** complex exhibited highest surface

319   energy binding affinities of -5.18 ± 0.01 while **C4** showed lowest energy. The energy

320   contribution of each energy function of all complexes is shown in **Table 3**.

321   The ΔG_{total} of **C2** (-37.11 ± 0.36), **C3** (-41.53 ± 0.39) and **C5 (**-36.79 ± 0.62) were quite better

322   and similar compared to **C1** and **C4** hit. The differences in the binding energies were due to the

323   difference in the contribution of electrostatic, van der Waal, and surface energies in the protein-

324   ligand complexes. In order to observe the major energy contribution and favorable interactions of

325   residues with selected hits, the total binding free energy was decomposed for each residue of the

326   complexes (Chohan et al. 2016), shown in Fig 8. The decomposition analysis showed that hits

327   were interacting with various PGAM1 residues. The residues like Phe22, Tyr92, Arg90, Lys100,

328   Arg116, Asn188, and Arg191 showed favorable contribution in total binding free energy of all

329   complexes except for **C4** complex where Asn20 and Phe22 contribution was very minor.

330   However, Glu19 and Glu89 residue showed unfavorable interactions with **C2, C3**, and **C5** hits,

331   as these interactions had positive binding free energy values. The interactions of C**2**, **C3**, and **C5**

332   compounds with the PGAM1 allosteric site residues before and after MD simulations are shown

333   in Fig 9.

334   **Distance and Hydrogen Bond Analysis**

335   The distances of hit molecules with key residues (Phe22, Tyr92, Arg191) that majorly

336   contributed to the binding free energy, was measured in all trajectories. It was observed that the

337   nitrogen (N1) of Borazine group of C2 was away from the amino group hydrogen (H) of Phe22

338   in first 20 ns but after that both maintained an average distance of ~ 4 Å till the end of

339   simulation. Similarly, the distance between sulfur atom (S1) of thiophene ring of C3 and benzene

340   hydrogen (H4) of Phe22 was in the range of ~ 4-4.5 Å throughout the simulation while the

341   distance between Isoxazolidine ring of C5 and benzene ring of Phe22 was ~ 3.5-6 Å for whole

342   simulation time. The distance between amine group (N2) of C2 and hydroxyl group (OH) of

343 Tyr92 was less than 4 Å throughout the simulation. Similarly, the distance between amine group

344 (N3) of C3 and hydroxyl group (OH) of Tyr92 was ~ 2-2.5 Å all the time except for some frames

345 where the distance increased to ~ 5 Å. The distance between the amine group (N2) of C5 and

346 hydroxyl group (OH) of Tyr92 was less than 2 Å till 20 ns and then increased to 4 Å at 22 ns.

347 The distance remained in this range till 49 ns. The distance increased to ~ 6 Å towards the end of

348 simulation. Moreover, the distance between sulfonamide group oxygen (O3) of C2 and amine

349 group (NH2) of Arg191 was < 4 Å till 18 ns and exceed to 10 Å after 20 ns for some time and

350 then attained an average distance of ~7 Å till 50 ns except for small deviations at 43 to 45 ns.

351 The similar trend was also observed in the distance between sulfonamide group oxygen (O3) of

352 C3 and amine group (NH1) of Arg191. However, the distance between sulfonamide group

353 oxygen (O1) of C5 and amine group (NH1) of Arg191 was in the range of ~ 2-5 Å throughout

354 the simulation (Fig 10). Further, hydrogen bond analysis was performed to observe the hydrogen

355 bond patterns between the protein binding pocket and selected ligands. Hydrogen bond

356 formation or distortion is a principal factor during simulation so large number of hydrogen bonds

357 show stronger affinity between the ligand and protein (Surti et al. 2020). The hydrogen bonding

358 plots of the three complexes are shown in Fig 11. The average number of hydrogen bonds

359 between **C2** and protein in first 25 ns were 1.60 while in the second half, the average number of

360 hydrogen bonds were 1.18. Similarly, the average number of hydrogen bonds between **C3** and

361 protein were 1.45. Among the three hits, the average number of hydrogen bonds of **C5**-PGAM1

362 were higher than other two hits with a value of 2.21. The hydrogen bonding analysis showed that

363 the selected hits made strong and stable complexes with protein.

364 **Alanine scanning of selected complexes**

365 The three complexes (**C2, C3** and **C5**) showed maximum binding free energies were further

366 investigated to identify the key role of Phe22, Tyr92, and Arg191residues in the total binding

367 free energy of protein-ligand complexes using computational alanine scanning. This method

368 helps to calculate the energy contribution of individual residue in total binding free energy of

369 protein-ligand complex(Liu et al. 2018). One speculation of the computational alanine scanning

370 is the effect of mutation on the structural perturbation of system is so small that the binding free

371 energy of mutated system can be obtained from the wildtype trajectory. So, three complexes

372 were selected based on the better binding free energies and the contribution of Phe22, Tyr92, and

373 Arg191 in total binding free energy was estimated one by one. The change in the binding free

374 energy of three complexes and entropy changes can be observed in **Table 4**. In PGAM1-**C2**

375 complex, the mutation of Arg191 to alanine played a significant role in the loss of binding free

376 energy as total energy reduced to -30.70 ± 0.29 with an energy loss of -6.41 ± 2.52 as compared

377 to the other two mutations where energy loss of Tyr92 and Phe22 mutations was -3.74 ± 1.40

378 and -4.49 ± 1.08 respectively. Similarly, the contribution of Arg191 was higher in PGAM1-**C3**

379 complex as this mutation lead to an energy loss of -6.67 ± 4.26 that was higher than -5.56 ± 1.49,

380 and -4.49 ± 1.44 for mutations of Tyr92 and Phe22 respectively. In case of PGAM1-**C5**

381 complex, same trend was observed as Arg191 mutation caused a loss of -5.72 ± 2.90 energy

382 while the energy loss for Phe22 and Tyr92 mutation caused an energy loss of -5.19 ± 1.72 and -

383 4.41 ± 2.58 respectively. The major difference in binding free energies was due to the mutation

384 of Arg191 which indicates that Arg191 have critical role in binding with the ligands.

385 **Conclusions**

386 In this study the optimized shape and structure-based approaches yielded the five hit compounds

387 having the sulfonamide as linker group to attach different chemical fragments. Molecular

388 docking studies reveal that among finally selected hits, C3 binds at PGAM1 allosteric site better

389 than other screened virtual hits. Further molecular dynamics simulations and binding free

390 energies calculations suggest that C2, C3 and C5 hits are tightly binding at the allosteric binding

391 site of PGAM1 enzyme with -37.11 ± 0.3, -41.53 ± 0.39 and -36.79 ± 0.62 Kcal/mol binding free

392 energies respectively with no structural distortions throughout the simulations time.

393 Computational alanine scanning suggests that Phe22, Tyr92 and Arg191 residues are playing

394 important role in binding interactions of these hits with PGAM1 enzyme, which is in agreement

395 with the experimental findings (Wen et al. 2019b). The predicted ADMET properties also

396 suggest that these compounds can be used as lead compound after performing the biological

397 assays for further optimization.

398 **Conflicts of Interest**

399 We wish to confirm that there is no known conflict of interest associated with this publication.

400

401 **References:**

402 Araújo PH, Ramos RS, da Cruz JN, Silva SG, Ferreira EF, de Lima LR, Macêdo WJ, Espejo-Román JM, Campos
403   JM, and Santos CB. 2020. Identification of potential COX-2 inhibitors for the treatment of
404   inflammatory diseases using molecular modeling approaches. *Molecules* 25:4183.

405 Berman H, Henrick K, and Nakamura H. 2003. Announcing the worldwide protein data bank. *Nature*
406   *Structural & Molecular Biology* 10:980-980.

407 Chohan TA, Qian H-Y, Pan Y-L, and Chen J-ZJMB. 2016. Molecular simulation studies on the binding
408   selectivity of 2-anilino-4-(thiazol-5-yl)-pyrimidines in complexes with CDK2 and CDK7. 12:145-
409   161.

410 Clark DE. 1999. Rapid calculation of polar molecular surface area and its application to the prediction of
411   transport phenomena. 1. Prediction of intestinal absorption. *Journal of pharmaceutical sciences*
412   88:807-814.

413 de Oliveira MS, da Cruz JN, Silva SG, da Costa WA, de Sousa SHB, Bezerra FWF, Teixeira E, da Silva NJN, de
414   Aguiar Andrade EH, and Neto AMdJC. 2019. Phytochemical profile, antioxidant activity, inhibition
415   of acetylcholinesterase and interaction mechanism of the major components of the Piper
416   divaricatum essential oil obtained by supercritical CO_2. *The Journal of Supercritical Fluids* 145:74-
417   84.

418 Evans MJ, Saghatelian A, Sorensen EJ, and Cravatt BF. 2005. Target discovery in small-molecule cell-based
419   screens by in situ proteome reactivity profiling. *Nature biotechnology* 23:1303-1307.

420 Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, and
421   Perry JK. 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and
422   assessment of docking accuracy. *Journal of medicinal chemistry* 47:1739-1749.

423 Grant JA, Gallardo M, and Pickup BT. 1996. A fast method of molecular shape comparison: A simple
424   application of a Gaussian description of molecular shape. *Journal of computational chemistry*
425   17:1653-1666.

426 Guasch L, Sala E, Mulero M, Valls C, Salvadó MJ, Pujadas G, and Garcia-Vallvé S. 2013. Identification of
427   PPARgamma partial agonists of natural origin (II): in silico prediction in natural extracts with
428   known antidiabetic activity. *PloS one* 8:e55889.

429 Hitosugi T, Zhou L, Elf S, Fan J, Kang H-B, Seo JH, Shan C, Dai Q, Zhang L, and Xie J. 2012. Phosphoglycerate
430   mutase 1 coordinates glycolysis and biosynthesis to promote tumor growth. *Cancer cell* 22:585-
431   600.

432 Imran M, Waqar S, Ogata K, Ahmed M, Noreen Z, Javed S, Bibi N, Bokhari H, Amjad A, and Muddassar M.
433   2020. Identification of novel bacterial urease inhibitors through molecular shape and structure
434   based virtual screening approaches. *RSC Advances* 10:16061-16070.

435 Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, and Klein ML. 1983. Comparison of simple
436   potential functions for simulating liquid water. *The Journal of chemical physics* 79:926-935.

437 Kearnes S, and Pande VJJoc-amd. 2016. ROCS-derived features for virtual screening. 30:609-617.

438 Lagunin A, Stepanchikova A, Filimonov D, and Poroikov V. 2000. PASS: prediction of activity spectra for
439   biologically active substances. *Bioinformatics* 16:747-748.

440 Lee S, Chang G, Lee I, Chung J, Sung K, and No K. 2004. The PreADME: pc-based program for batch
441   prediction of adme properties. *EuroQSAR* 9:5-10.

442 Lee S, Lee I, Kim H, Chang G, Chung J, and No K. 2003. The PreADME Approach: Web-based program for
443   rapid prediction of physico-chemical, drug absorption and drug-like properties. *EuroQSAR 2002*
444   *Designing Drugs and Crop Protectants: processes, problems and solutions* 2003:418-420.

445 Li H, Robertson AD, and Jensen JH. 2005. Very fast empirical prediction and rationalization of protein pKa
446   values. *Proteins: Structure, Function, and Bioinformatics* 61:704-721.

447   Li Q, Yang H, Mo J, Chen Y, Wu Y, Kang C, Sun Y, and Sun HJP. 2018. Identification by shape-based virtual
448   screening and evaluation of new tyrosinase inhibitors. 6:e4206.
449   Li X, Tang S, Wang Q-Q, Leung EL-H, Jin H, Huang Y, Liu J, Geng M, Huang M, and Yuan S. 2017.
450   Identification of epigallocatechin-3-gallate as an inhibitor of phosphoglycerate mutase 1. *Frontiers*
451   *in pharmacology* 8:325.
452   Liu S-e, Hu J-c, Zhang H, Xu P, Wan W, Zheng M-y, Yu K-q, Ding H, Jiang H-l, and Zhou L. 2017. Conformation
453   and dynamics of the C-terminal region in human phosphoglycerate mutase 1. *Acta*
454   *Pharmacologica Sinica* 38:1673-1682.
455   Liu X, Peng L, Zhou Y, Zhang Y, and Zhang JZH. 2018. Computational Alanine Scanning with Interaction
456   Entropy for Protein–Ligand Binding Free Energies. *Journal of Chemical Theory and Computation*
457   14:1772-1780. 10.1021/acs.jctc.7b01295
458   Lobanov MY, Bogatyreva N, and Galzitskaya OJMB. 2008. Radius of gyration as an indicator of protein
459   structure compactness. 42:623-628.
460   Ma Y-C, Li C, Gao F, Xu Y, Jiang Z-B, Liu J-X, and Jin L-Y. 2014. Epigallocatechin gallate inhibits the growth
461   of human lung cancer by directly targeting the EGFR signaling pathway. *Oncology reports* 31:1343-
462   1349.
463   Macke TJ, Svrcek-Seiler W, Brown RA, Kolossváry I, Bomble YJ, Case DA, Zhang W, Hou T, Schafmeister C,
464   and Ross WS. 2010. AmberTools Users' Manual. Ver.
465   Mangiatordi GF, Trisciuzzi D, Alberga D, Denora N, Iacobazzi RM, Gadaleta D, Catto M, and Nicolotti O.
466   2017. Novel chemotypes targeting tubulin at the colchicine binding site and unbiasing P-
467   glycoprotein. *European journal of medicinal chemistry* 139:792-803.
468   Martínez LJPo. 2015. Automatic identification of mobile and rigid substructures in molecular dynamics
469   simulations and fractional structural fluctuation analysis. 10:e0119264.
470   Poongavanam V, and Kongsted J. 2013. Virtual screening models for prediction of HIV-1 RT associated
471   RNase H inhibition. *PloS one* 8:e73478.
472   Rush TS, Grant JA, Mosyak L, and Nicholls AJJomc. 2005. A shape-based 3-D scaffold hopping method and
473   its application to a bacterial protein– protein interaction. 48:1489-1495.
474   Sargsyan K, Grauffel C, Lim CJJoct, and computation. 2017. How molecular size impacts RMSD applications
475   in molecular dynamics simulations. 13:1518-1524.
476   Sastry GM, Adzhigirey M, Day T, Annabhimoju R, and Sherman W. 2013. Protein and ligand preparation:
477   parameters, protocols, and influence on virtual screening enrichments. *Journal of computer-aided*
478   *molecular design* 27:221-234.
479   Sun H, Li Y, Shen M, Tian S, Xu L, Pan P, Guan Y, and Hou T. 2014. Assessing the performance of MM/PBSA
480   and MM/GBSA methods. 5. Improved docking performance using high solute dielectric constant
481   MM/GBSA and MM/PBSA rescoring. *Physical Chemistry Chemical Physics* 16:22035-22045.
482   Surti M, Patel M, Adnan M, Moin A, Ashraf SA, Siddiqui AJ, Snoussi M, Deshpande S, and Reddy MNJRa.
483   2020. Ilimaquinone (marine sponge metabolite) as a novel inhibitor of SARS-CoV-2 key target
484   proteins in comparison with suggested COVID-19 drugs: designing, docking and molecular
485   dynamics simulation study. 10:37707-37720.
486   Wang P, Jiang L, Cao Y, Ye D, and Zhou LJM. 2018a. The design and synthesis of N-xanthone
487   benzenesulfonamides as novel phosphoglycerate mutase 1 (PGAM1) inhibitors. 23:1396.
488   Wang P, Jiang L, Cao Y, Zhang X, Chen B, Zhang S, Huang K, Ye D, and Zhou L. 2018b. Xanthone derivatives
489   as phosphoglycerate mutase 1 inhibitors: design, synthesis, and biological evaluation. *Bioorganic*
490   *& medicinal chemistry* 26:1961-1970.
491   Wang Y, Guo Y, Qiang S, Jin R, Li Z, Tang Y, Leung ELH, Guo H, and Yao XJFiP. 2021. 3D-QSAR, Molecular
492   Docking, and MD Simulations of Anthraquinone Derivatives as PGAM1 Inhibitors. 12.

493   Wen C-L, Huang K, Jiang L-L, Lu X-X, Dai Y-T, Shi M-M, Tang X-M, Wang Q-B, Zhang X-D, and Wang P-H.
494   2019a. An allosteric PGAM1 inhibitor effectively suppresses pancreatic ductal adenocarcinoma.
495   *Proceedings of the National Academy of Sciences* 116:23264-23273.
496   Wen C-L, Huang K, Jiang L-L, Lu X-X, Dai Y-T, Shi M-M, Tang X-M, Wang Q-B, Zhang X-D, and Wang P-
497   HJPotNAoS. 2019b. An allosteric PGAM1 inhibitor effectively suppresses pancreatic ductal
498   adenocarcinoma. 116:23264-23273.
499   Xu Y, Li F, Lv L, Li T, Zhou X, Deng C-X, Guan K-L, Lei Q-Y, and Xiong Y. 2014. Oxidative stress activates SIRT2
500   to deacetylate and stimulate phosphoglycerate mutase. *Cancer research* 74:3630-3642.
501   Xu Z, Gong J, Wang C, Wang Y, Song Y, Xu W, Liu Z, and Liu Y. 2016. The diagnostic value and functional
502   roles of phosphoglycerate mutase 1 in glioma. *Oncology reports* 36:2236-2244.
503   Zhang D, Jin N, Sun W, Li X, Liu B, Xie Z, Qu J, Xu J, Yang X, and Su Y. 2017. Phosphoglycerate mutase 1
504   promotes cancer cell migration independent of its metabolic activity. *Oncogene* 36:2900-2909.

505

ADME properties calculation of virtually screened compounds using PreADMET server (https://preadmet.bmdrc.kr/adme/).

Table 1: ADME properties calculation of virtually screened compounds using PreADMET
(https://preadmet.bmdrc.kr/adme/).

Compound	Human Intestinal Absorption	Cell Permeability	Cell Permeability Maden Darby Canine Kidney	In-Vitro P-glycoprotein inhibition	Permeability of Blood-Brain Barrier
	HIA (%)	CaCo-2 (nm/sec)	P_{MDCK} (nm/sec)	P-gp Inhibition	BBB (%)
C1	100.000	22.281	204.401	Non	1.499
C2	96.912	14.758	2.171	Non	0.154
C3	97.039	19.580	1.510	Non	0.181
C4	97.063	4.271	4.139	Inhibitor	0.736
C5	97.276	11.180	0.667	Non	1.564

Off target biological activities of top five virtually screened compounds (C1, C2, C3, C4, and C5), determined by web server PASS (http://www.pharmaexpert.ru/passonline/).

1 **Table 2:** Pharmacokinetics properties and biological activities of top five virtually screened compounds

2 (C1, C2, C3, C4, and C5), as determined by web server PASS (http://www.pharmaexpert.ru/passonline/).

COMPOUND	Possibility of Activity (p_a)	Possibility of Inactive(p_i)	Biological Activities
C1	0.150	0.141	Antipruritic, non-allergic
	0.162	0.153	Antibacterial
	0.784	0.006	Proteasome ATPase inhibitor
	0.637	0.018	Insulysin inhibitor
C2	0.411	0.049	Antiallergic
	0.863	0.004	Omptin inhibitor
	0.253	0.093	Alzheimer's disease treatment
	0.142	0.131	Antineoplastic (renal cancer)
C3	0.243	0.196	Insulysin inhibitor
	0.338	0.016	Antineoplastic
	0.090	0.073	Lipoxygenase inhibitor
	0.140	0.010	Sulfonylureas
C4	0.808	0.005	Ant obesity
	0.650	0.009	Antidiabetic
	0.290	0.070	Alzheimer's disease treatment
	0.258	0.138	Antineoplastic (Multi myeloma)
C5	0.231	0.021	Osteoarthritis treatment
	0.379	0.049	Antidiabetic
	0.734	0.031	Nootropic
	0.169	0.162	Diabetic nephropathy treatment

Binding free energy (kJ/mol) calculations of wild type complexes using MMGBSA

1   **Table 3:** Binding free energy (kJ/mol) calculation of complex C1, C2, C3, C4, and C5 determine by

2   MMGBSA.

Interactions	PGAM1-C1	PGAM1-C2	PGAM1-C3	PGAM1-C4	PGAM1-C5
ΔG_{vdW}	-29.61 ± 0.17	-37.42 ± 0.26	-42.36 ± 0.19	-17.79 ± 0.35	-35.96 ± 0.32
ΔG_{ele}	-21.66 ± 0.57	-19.64 ± 0.29	-21.39 ± 0.34	-13.24 ± 0.72	-15.17 ± 0.50
ΔG_{GB}	38.31 ± 0.49	25.04 ± 0.29	27.41 ± 0.25	25.98 ± 0.71	19.33 ± 0.28
ΔG_{surf}	-4.09 ± 0.02	-5.09 ± 0.02	-5.18 ± 0.01	-2.16 ± 0.04	-4.99 ± 0.04
ΔG_{gas}	-51.27 ± 0.55	-57.06 ± 0.32	-63.76 ± 0.33	-31.04 ± 0.55	-51.13 ± 0.74
ΔG_{solv}	34.21 ± 0.48	19.94 ± 0.29	22.22 ± 0.25	23.81 ± 0.73	14.34 ± 0.26
ΔG_{total}	-17.05 ± 0.23	-37.11 ± 0.36	-41.53 ± 0.39	-7.22 ± 0.34	-36.79 ± 0.62

MMGBSA based binding free energy (kJ/mol) calculations of mutant complexes.

1   **Table 4:** Binding free energy (kcal/mol) calculation of mutant complexes C2, C3 and C5.

Interactions	PGAM1-C2 R191A	PGAM1-C2 Y92A	PGAM1-C2 F22A	PGAM1-C3 R191A	PGAM1-C3 Y92A	PGAM1-C3 F22A	PGAM1-C5 R191A	PGAM1-C5 Y92A	PGAM1-C5 F22A
ΔG_{vdW}	-35.62±0.27	-33.90±0.23	-33.50±0.24	-41.48±0.18	-38.57±0.17	-38.53±0.17	-35.22±0.33	-33.16±0.30	-32.08±0.28
ΔG_{ele}	-11.10±0.29	-19.35±0.29	-19.11±0.28	-8.63±0.29	-19.58±0.34	-20.77±0.35	-4.13±0.29	-12.60±0.44	-14.51±0.51
ΔG_{GB}	20.83±0.28	25.11±0.30	24.63±0.28	20.22±0.20	27.04±0.23	26.99±0.24	13.04±0.20	17.98±0.29	19.41±0.31
ΔG_{surf}	-4.81±0.02	-4.82±0.02	-4.64±0.02	-4.96±0.01	-4.83±0.01	-4.72±0.01	-4.75±0.04	-4.60±0.03	-4.40±0.04
ΔG_{gas}	-46.72±0.28	-53.65±0.31	-52.61±0.31	-50.11±0.34	-58.15±0.35	-59.31±0.34	-39.35±0.53	-45.76±0.67	-46.60±0.71
ΔG_{solv}	16.01±0.28	20.28±0.30	19.98±0.28	15.25±0.20	22.21±0.23	22.27±0.24	8.28±0.21	13.38±0.27	15.00±0.29
$T\Delta S$	-22.62±0.84	-22.62±0.84	-22.62±0.84	-20.23±1.26	-20.23±1.26	-20.23±1.26	-24.62±1.15	-24.62±1.15	-24.62±1.15
$W\Delta G_{total}$	-37.11 ± 0.36	-37.11± 0.36	-37.11 ± 0.36	-41.53 ± 0.39	-41.53 ± 0.39	-41.53 ± 0.39	-36.79 ± 0.62	-36.79 ± 0.62	-36.79 ± 0.62
$M\Delta G_{total}$	-30.70±0.29	-33.37±0.34	-32.62±0.34	-34.86±0.25	-35.94±0.41	-37.03±0.40	-31.06±0.56	-32.37±0.55	-31.59±0.57
$\Delta\Delta G_{bind}$	-6.41±2.52	-3.74±1.40	-4.49±1.08	-6.67±4.26	-5.58±1.49	-4.49±1.44	-5.72±2.90	-4.41±2.58	-5.19±1.72

3    $T\Delta S$ = The entropy changes

4    $W\Delta G_{total}$ = Binding free energy of Wild type complex

5    $M\Delta G_{total}$ = Binding free energy of Mutant complex

6     $\Delta\Delta G_{bind}$ = Difference in mutant and wild type energy values

Figure 1

The flowchart of the study

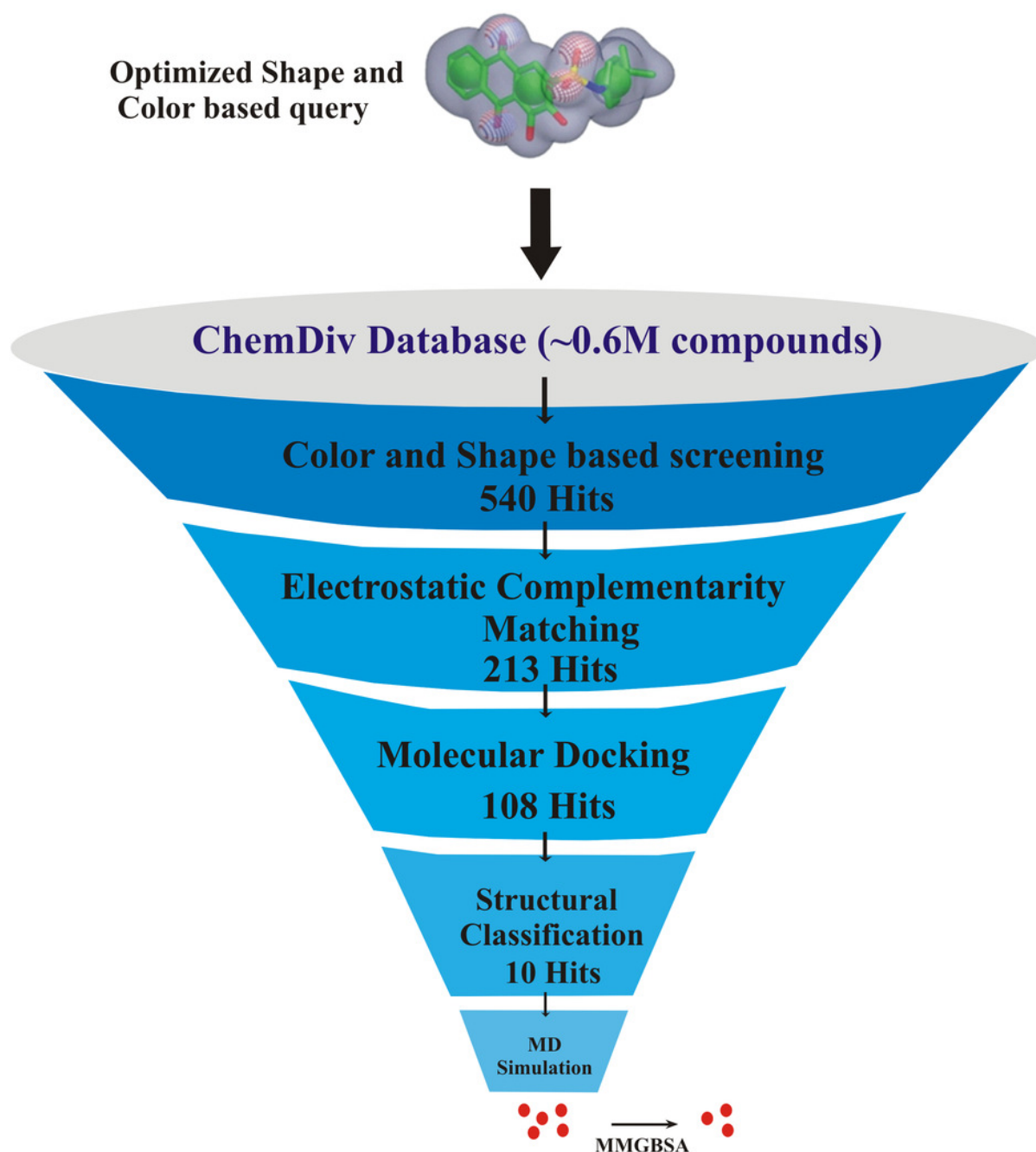



Figure 2

Query model optimization

(A) Multiple ROCS queries on cocrystal ligand. Green color is indicating hydrophobic regions i.e., four aromatic rings, four red regions are showing acceptors and a single blue region is representing donor. (B) Validation of multiple query model of co-crystal ligand 8kx represented by AUC curve.

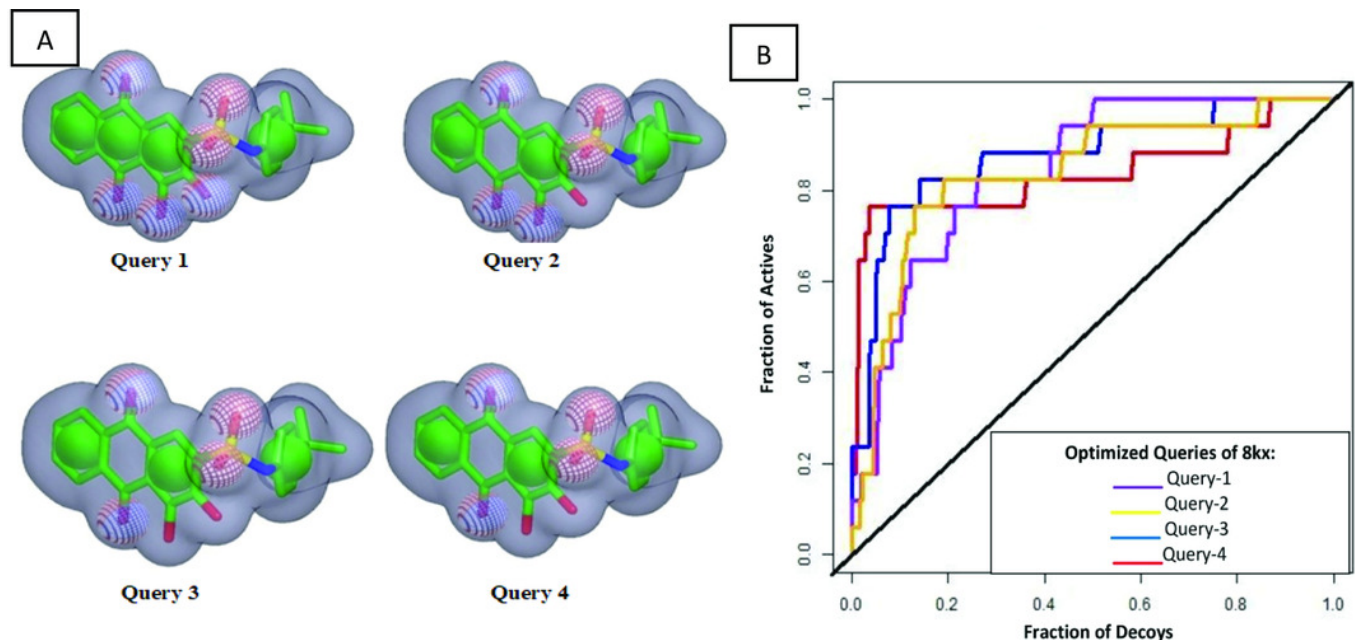
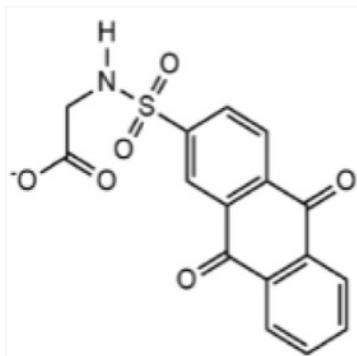
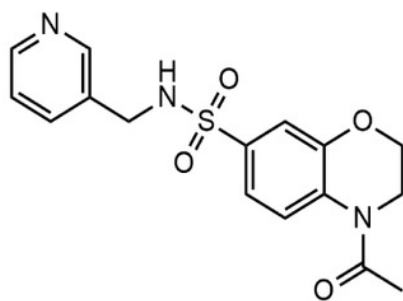



Figure 3

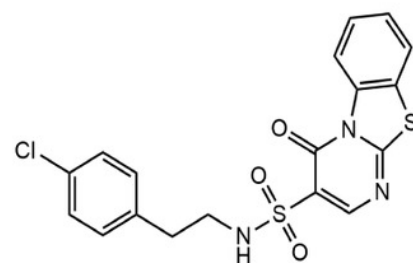
The best virtual screen hits

(C1= 6144-0309, C2= E470-1348, C3= C301-8900, C4= F540-0157, C5= L464-0403) from Chemdiv database with their chemical structures, ROCS TanimotoCombo scores (T.C) and glide docking scores (G.S)

C1

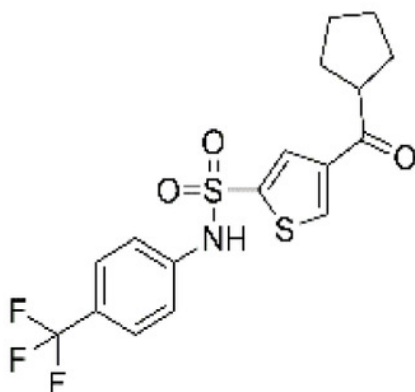
T.C= 1.4520

G. S= -5.570

C2

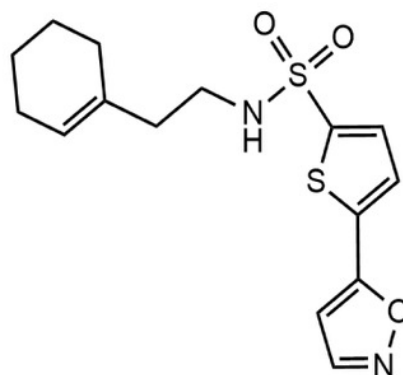
T.C= 1.5430

G. S= -6.902

C3

T.C= 1.3580

G. S= -6.054

C4

T.C= 1.2730

G. S= -6.446

C5

T.C= 1.4330

G. S= -6.128

Figure 4

Docked poses of hit compounds

Protein-ligand complexes (C1, C2, C3, C4, C5 and 8kx) with best screened compounds and co-crystal ligand in the binding site. A) 2D interactions of hits with the allosteric site residues. B) The binding poses of the best selected hits in the allosteric site with reference to cocrystal ligand 8kx.

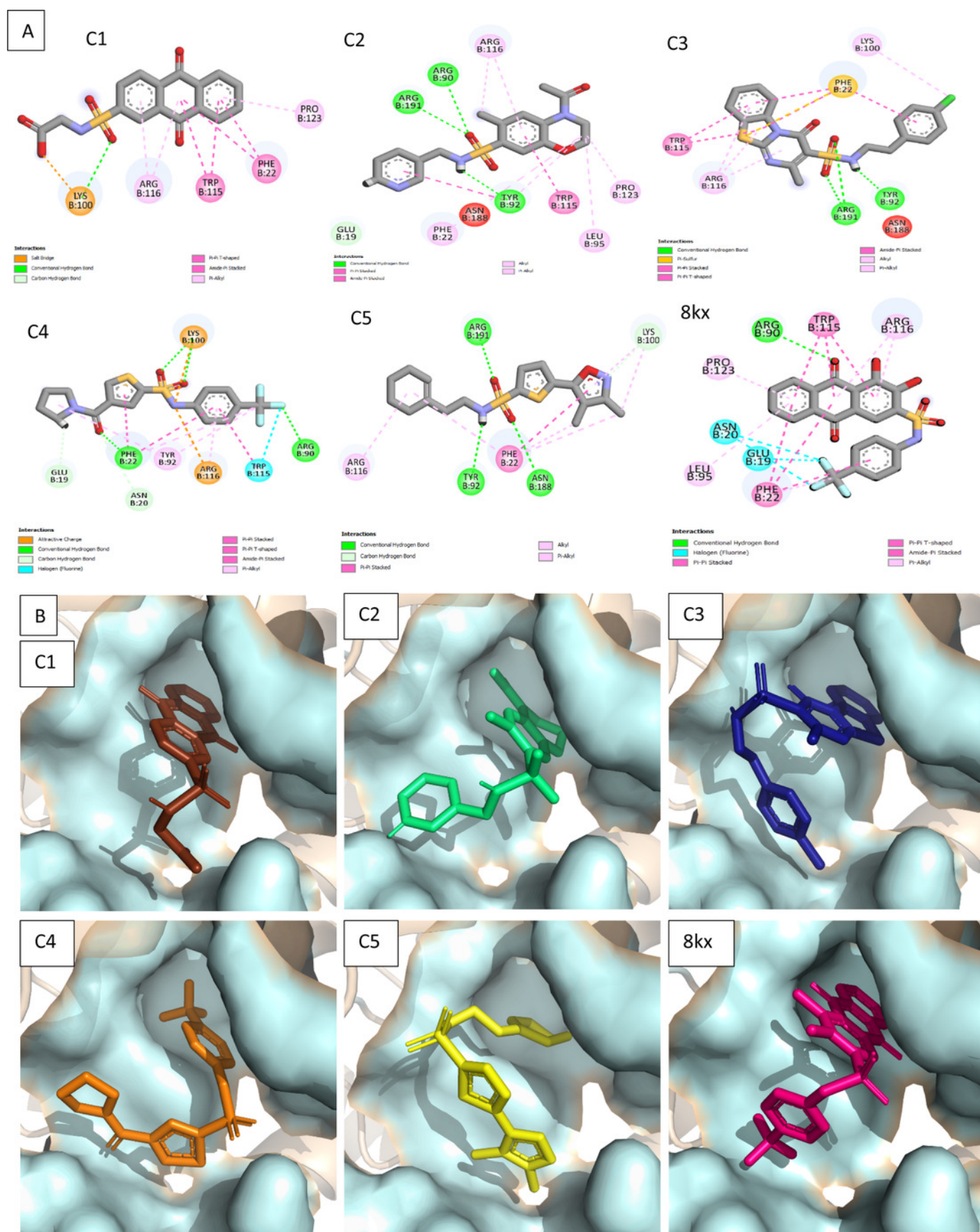



Figure 5

The RMSD(Å) vs time(ns) plots

PGAM-1 enzyme backbone atoms in complex with C1 (red), C2 (blue), C3 (green), C4 (orange), and C5 (Pink) hits.

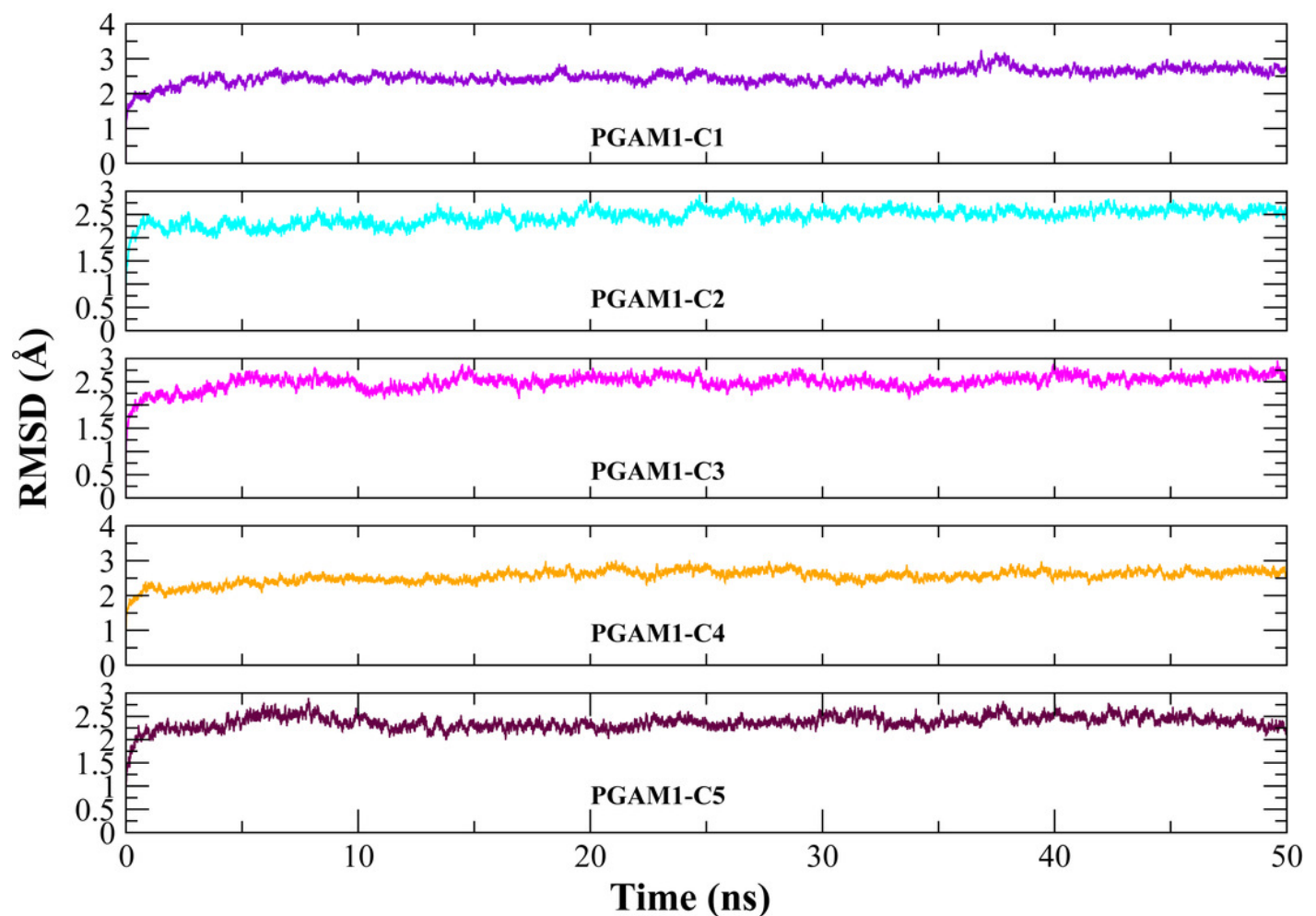



Figure 6

The superimposed snapshots of the MD trajectories of C1 to C5 complexes

(A) PGAM1- C1 (B) PGAM1-C2 (C) PGAM1-C3 (D) PGAM1-C4 (E) PGAM1-C5

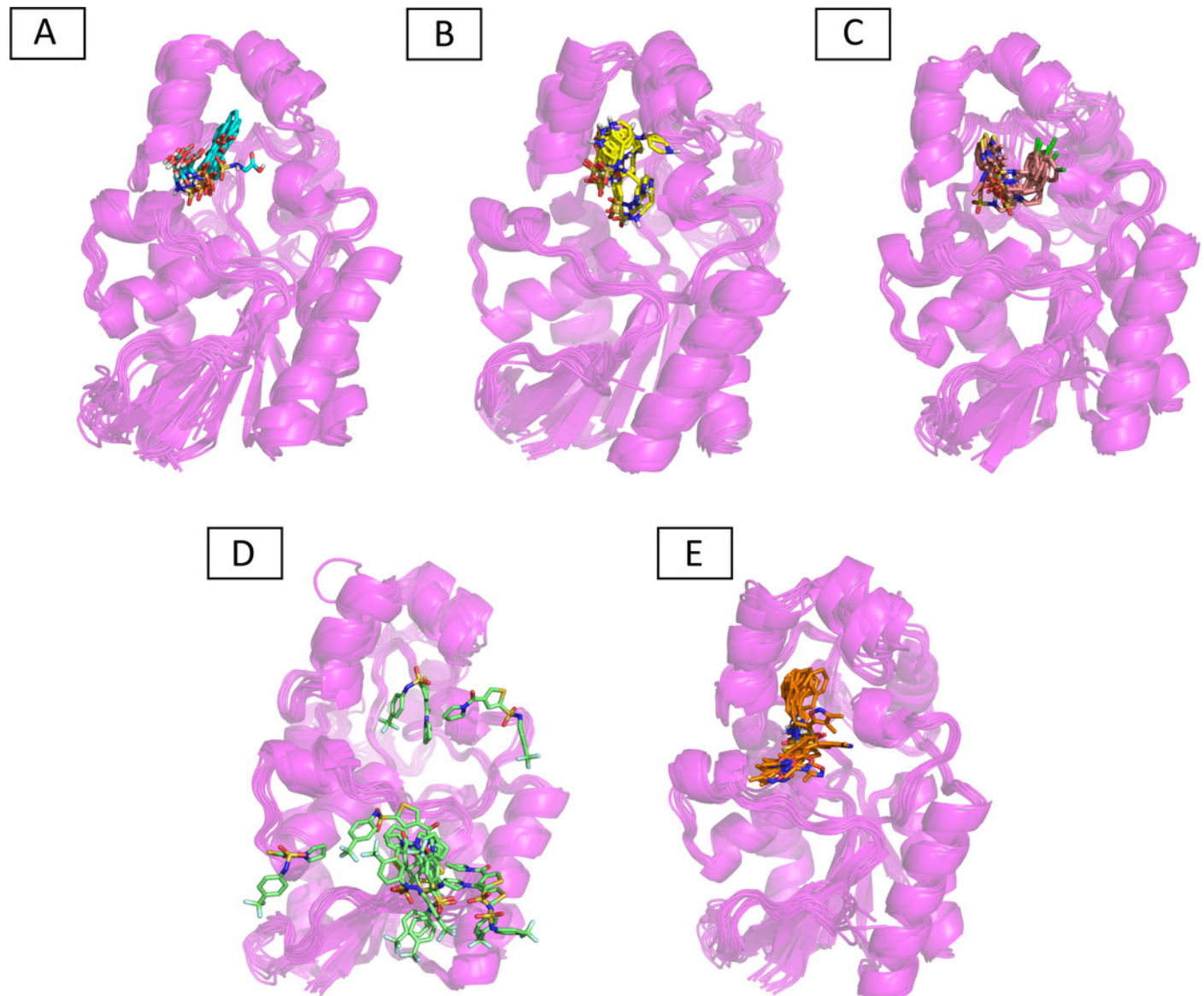



Figure 7

The RMSF (Å) plots of complexes

The RMSF (Å) plots of residues of PGAM-1 enzyme in complex with C1 (red), C2 (blue), C3 (green), C4 (orange), and C5 (Pink) hits

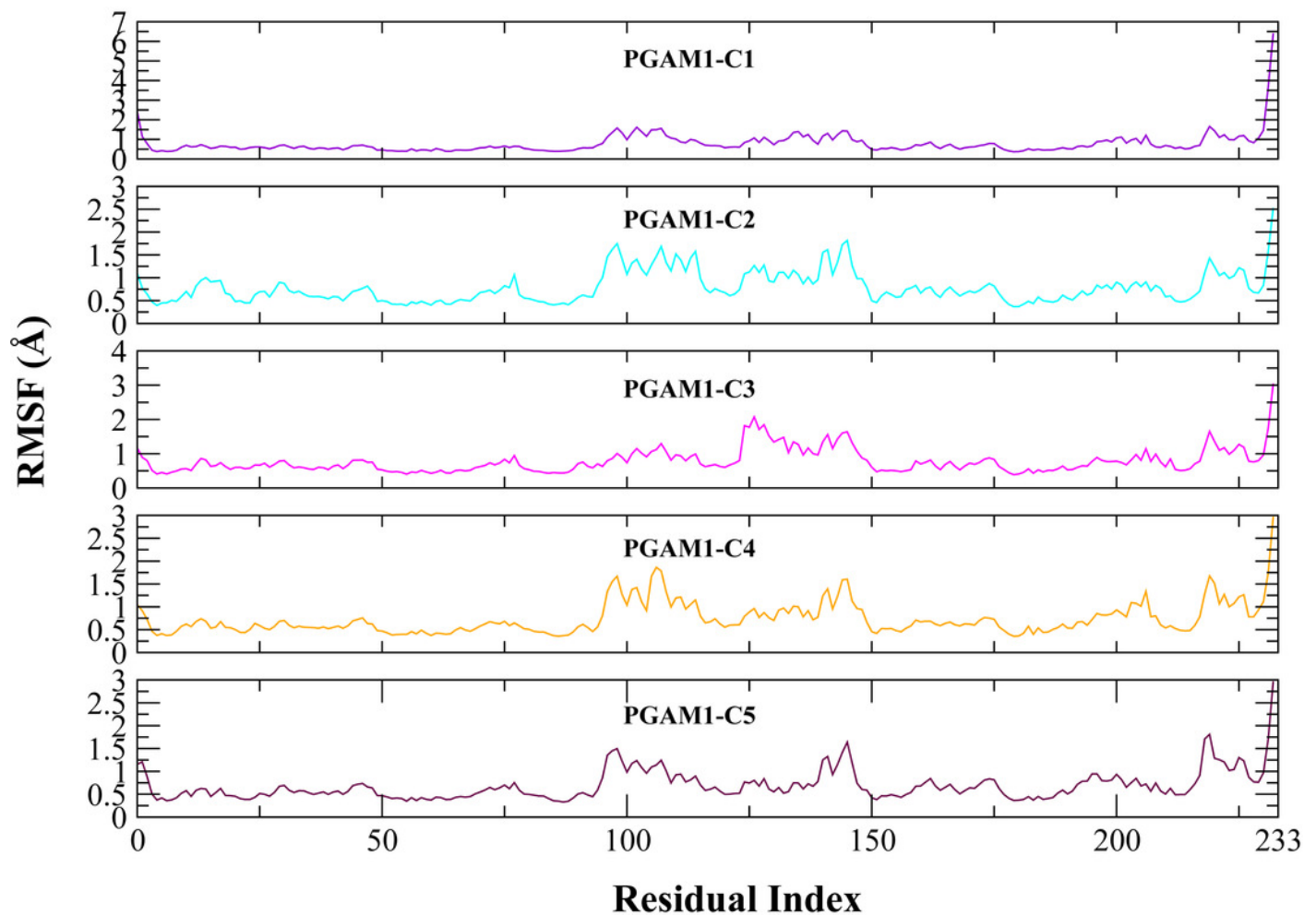



Figure 8

Total binding free energy decomposition of all five hits-PGAM-1 complexes

Key residues with major contribution (Phe22, Arg90, Tyr92, Arg116, and Arg191) are shown in peaks.

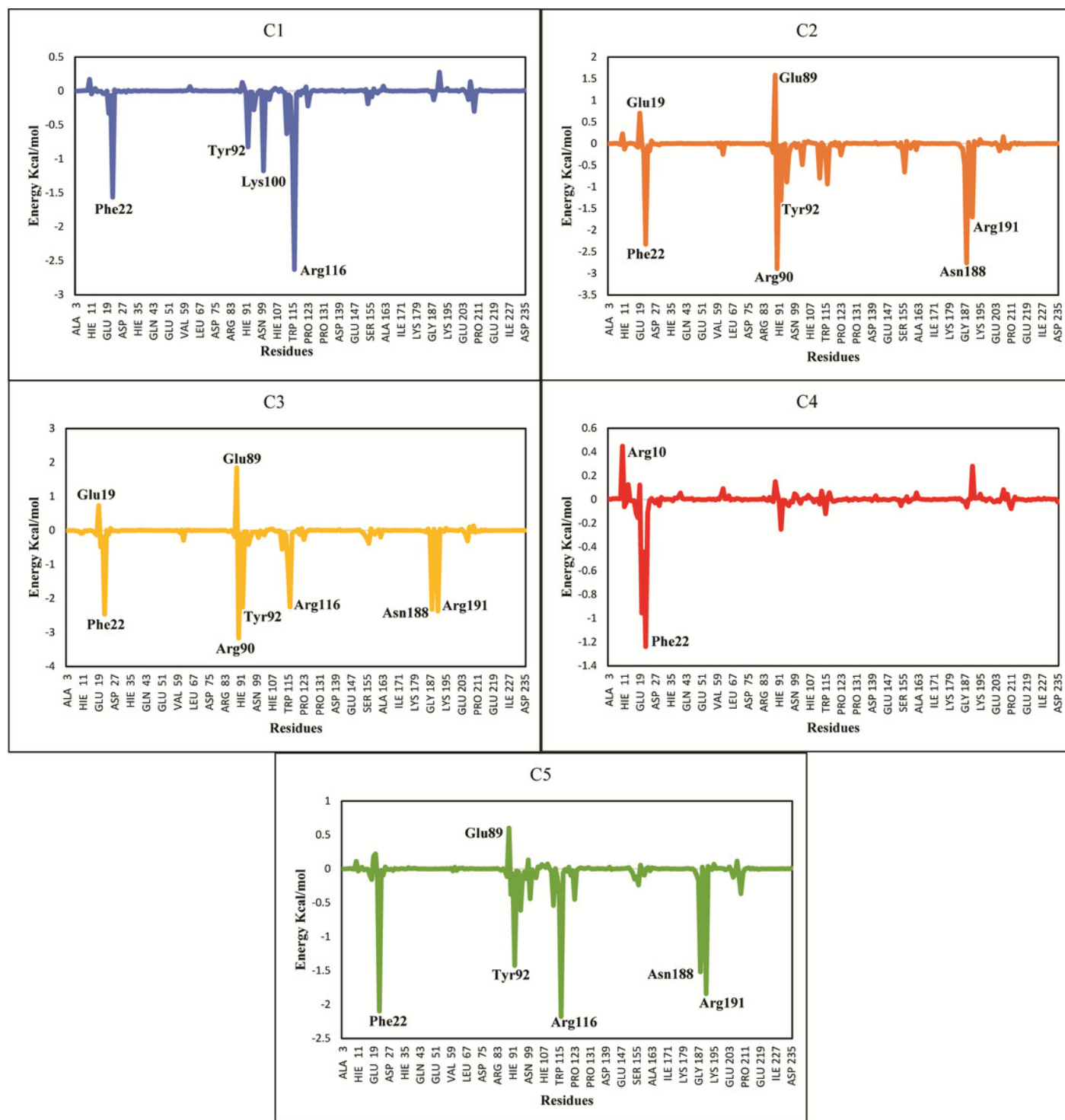



Figure 9

The interactions patterns of finally selected hits before and after MD simulations

(A) The interactions showed at 0 ns (first frame). (B) The interactions at 50 ns (last frame). (C) The binding modes of hits before and after MD simulations. The red sticks are showing poses (first frame) while the green sticks are indicating the poses in last frame.

C2   C3   C5

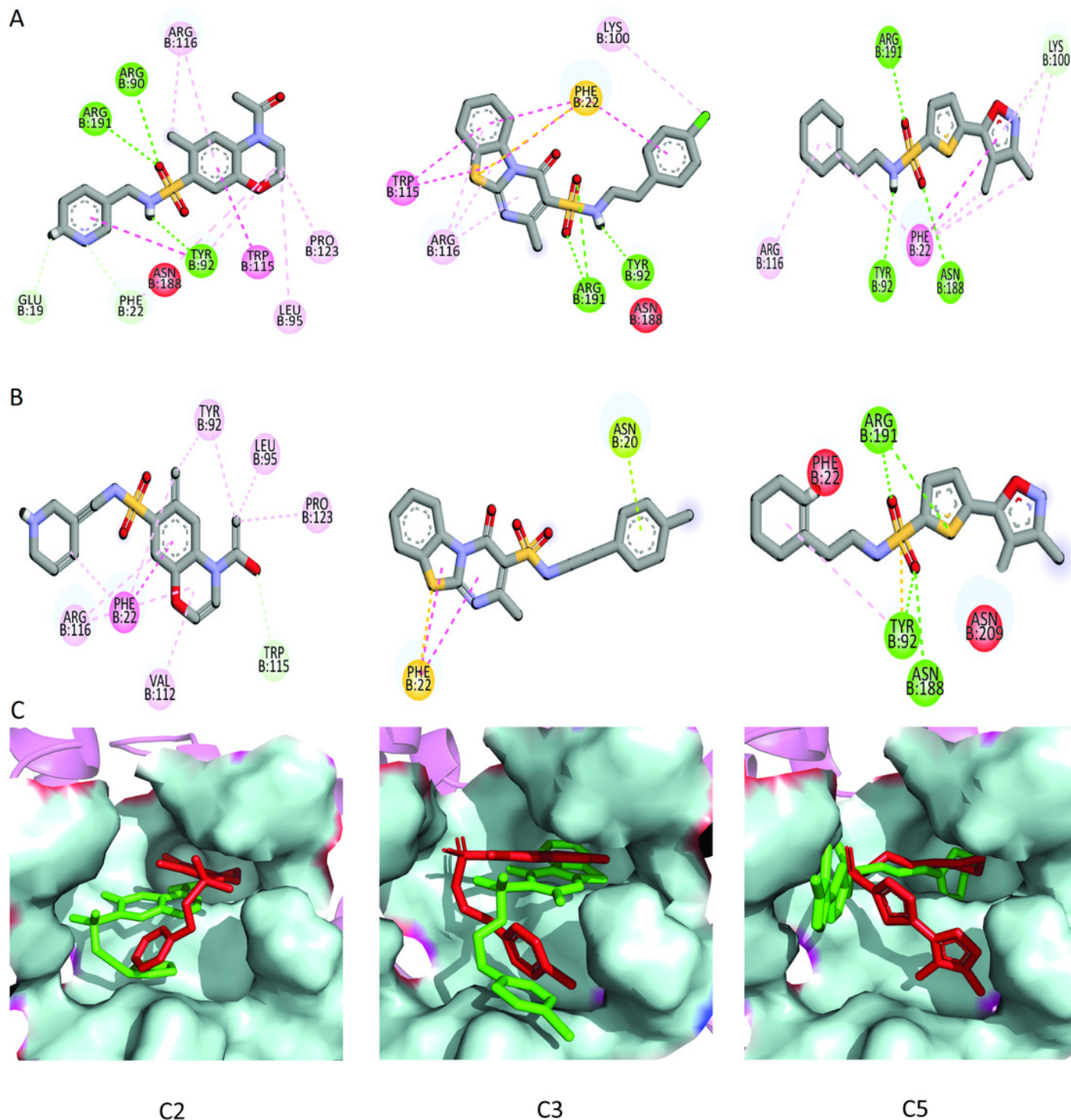



Figure 10

Distance plots among key residues and selected hits

Distance between the key residues (Phe22, Tyr92, Arg191) and selected hit compound throughout the MD simulation. Cyan plots indicate the distance between Phe22 and hit compounds, Magenta shows the Tyr92-Hits distances,while black plots show the distances between hits and Arg191.

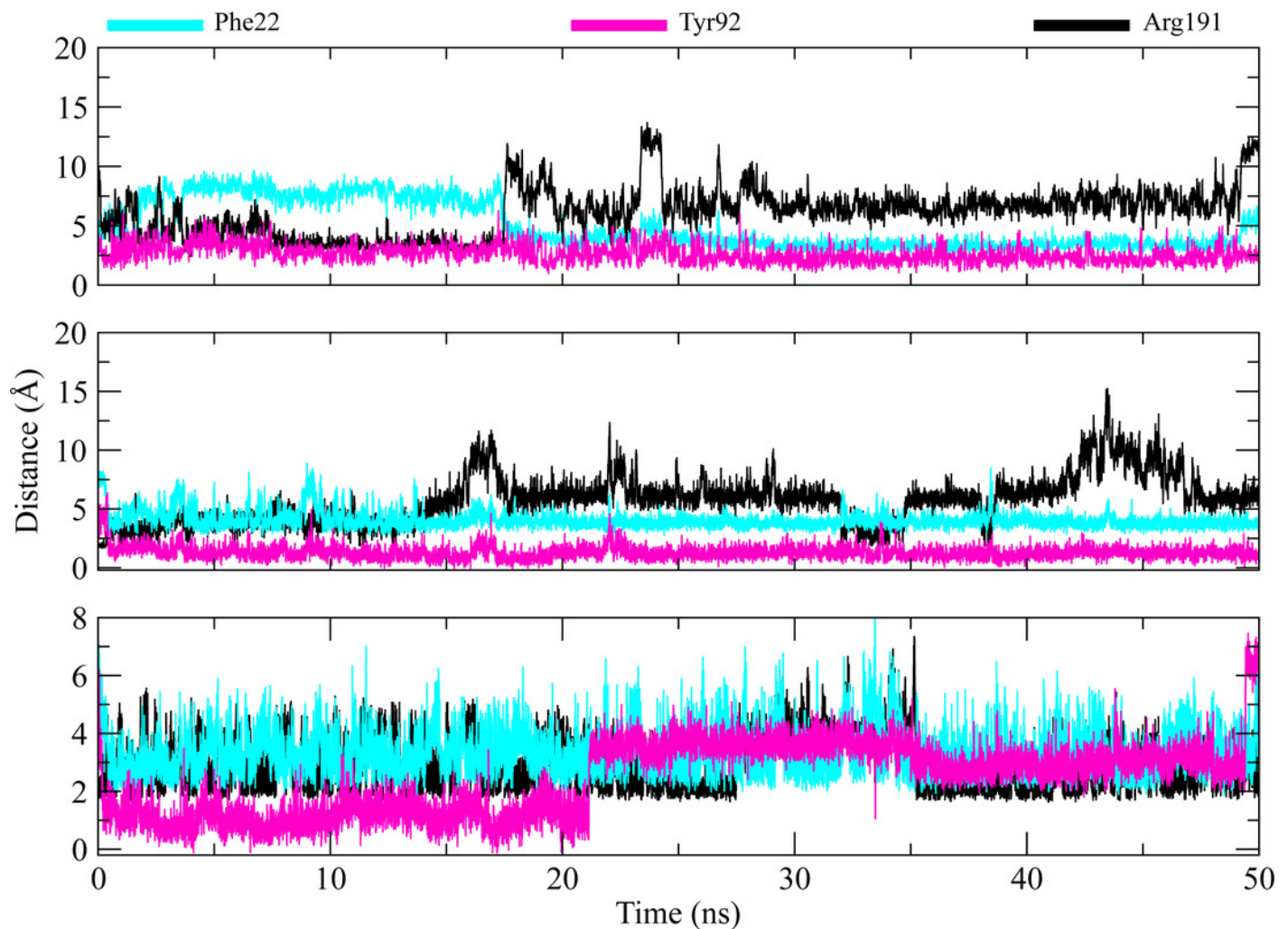



Figure 11

Hydrogen bond analysis of hits with key residues

The number of hydrogen bonds formed between the binding site of protein and selected hits (C2, C3, C5) during simulation. The plots show the consistent hydrogen bonding among all complexes.