Molecular docking studies of pancreatic cancer expressed proteins with Psidium guajava (guava) derived bioactive compounds

Tharrun Daniel Paul L, Roshan Kumar Sah, Thirunavukkarasu Palaniyandi*
Department of Biotechnology, Dr. M. G. R Educational and Research Institute, Chennai, Tamil Nadu, India

Article History:
Received on: 25 Sep 2020
Revised on: 20 Nov 2020
Accepted on: 25 Nov 2020

Keywords:
Psidium guajava, auto dock vina, discovery studio, quercetin, pancreatic cancer

The current study evaluates the binding affinities (kcal/mol) of different proteins expressed in pancreatic cancer with Psidium guajava derived bioactive compounds by performing molecular docking through auto dock vina. Auto dock vina was used to perform molecular docking between the proteins expressed in pancreatic cancer and P. guajava derived bioactive compounds. Nine proteins and nine ligands were chosen for molecular docking. Among the nine ligands, gemcitabine which is a commercial first-line drug used to treat pancreatic cancer, was selected. The docking output was visualized using the Biovia Discovery Studio visualizer. From the docking results, we found that, out of the nine ligands, quercetin had a better binding affinity than the other ligands and the commercial drug (gemcitabine). SNAI1 docked with quercetin had a binding affinity of -9.6 kcal/mol, which was found to be the highest. In conclusion, it can be said that the compound quercetin derived from the ethanolic extract of the P. guajava has the highest binding affinity, so it can be used for the treatment of pancreatic cancer after modification to its properties so that it has good efficacy and pharmacokinetic properties. Further studies will be based on the in-vitro testing of the extract and gene and protein expression analysis using RT-PCR and MALDI-TOF, respectively.

*Corresponding Author
Name: Thirunavukkarasu Palaniyandi
Phone: 9952172249
Email: ppthirunacas@gmail.com

INTRODUCTION
Cancer is not just the accumulation of tissue masses but also the heterotypic interaction of distinct cell types with one another (Hanahan and Weinberg, 2011). Cancer was the second leading cause of mortality in 2018, resulting in 9.6 million deaths (WHO, https://www.who.int/health-topics/cancer). Pancreatic cancer is the fourth leading cause of cancer deaths in the United States and this is because the annual death rate being equal to the annual incidence rates (Street, 2019). Pancreatic cancer has a survival rate of only 5 years that too the percent of survival rate in that is only 9% (Rahib et al., 2014). The reason behind this is the inability of the disease to be even diagnosed and cured (Warshaw and del Castillo, 1992). It is found that 80% of pancreatic cancers are unresectable in patients and also, the survival rate of the subjects after surgery is poor (Rahib et al., 2014). Pancreatic cancer is estimated to be the second leading cause of death by 2030 (Gilabert et al., 2017). Among all the other cancer types, pancreatic cancer exhibited high resistance to traditional chemotherapy and radiation therapy, including both de novo (intrinsic) and acquired (therapy-induced) chemo-resistance of the cancer cells (Long et al., 2011). The development of drug resistance is a critical barrier for the effec-
tive transportation of a drug to the target site in the tumor. There are two types of drug resistance, the one being de novo resistance and the other acquired resistance. De novo drug resistance subjects do not respond to the chemotherapy from the beginning of its administration and acquired resistance subjects are first sensitive to the drug but later, they become resistant to the treatment. The cancer cells also become resistant to other mechanically and structurally unrelated drugs, which is called multidrug resistance (Jäger, 2009). Thus, a treatment combining several drugs for different targets also fail (Wang et al., 2011; Gottesman et al., 2002).

Currently, there are many combination therapies for pancreatic cancer. Gemcitabine is used as the drug of choice for enfeebled patients and in combination along with other drugs such as irinotecan, oxaliplatin (also known as FOLFIRINOX), 5-fluorouracil, leucovorin in the case of healthy subjects (Kobayashi et al., 2017; Orlandi et al., 2016). But due to the short half-life range of 8 to 15 mins for the drug, the continuous administration of the drug parenterally causes renal and hematological toxicities (Dorjee and Long, 2018; Kasuya et al., 2011; Muranaka et al., 2017). Although gemcitabine is used as a first-line drug, it is not satisfactory due to the endogenous and exogenous resistance exhibited. Endogenous drug resistance occurs due to the drug metabolism, its transport, abnormal activation and inactivation of its signaling pathways, whereas exogenous drug resistance occurs due to impedance of drug delivery to the target (Hung et al., 2012; Nakano et al., 2007).

Psidium guajava (commonly called as guava) is a native fruit of Mexico, and it grows widely in South America, Europe, Africa and Asia. It has been reported that the main traditional use was as an anti-diarrheal. Gastroenteritis, dysentery, stomach disorders, antibacterial colic pathogenic germs of the intestine were the ailments for which this plant was used as a treatment (Gutiérrez et al., 2008). The leaves were used to prepare a decoction to treat cough (Heinrich et al., 1998). In Mexico it is used to treat gastrointestinal and respiratory disturbances and also as an anti-inflammatory medicine (Rehab et al., 2019). Psidium guajava leaves have shown to possess some ethnomedical uses such as to treat diarrhea and stomach in the form of infusion or decoction (Pontikis, 1996), Diabetes mellitus, hypertension (Oh et al., 2005; Ojewole, 2005), Febrifuge, antispasmodic, rhematism and several other treatments. The leaves were chewed to relieve toothache and applied on wounds, ulcers and rheumatic pain (Heinrich et al., 1998). It is used as an astrigent, drying agent and a diuretic in Latin America, Central and West Africa and in Southeast Asia. The decoction of the leaves and bark has been used in India for treating diarrhea, dysentery, vomiting, sore throat and menstrual cycle regulation. Amazon tribes have been using the decoction for the treatment of bleeding gums, rinse off vaginal discharge and to tone up the vaginal walls after labour (Kamath et al., 2008).

Ryu et al. (2012) has reported that compounds such as γ-sitosterol, Vitamin E and Squalene present in P. guajava contribute to the potent anti-cancer activity by exhibiting certain mechanisms such as suppression of signalling pathways, apoptosis induction and cell-cycle arrest. β-caryophyllene inhibits the lipopolysaccharide-stimulated proinflammatory cytokines (TNF-α and IL-1β) in peripheral blood, thereby exhibiting an anti-inflammatory efficacy (Gertsch et al., 2008). The study conducted by Ujiki et al. (2006) has shown the role of Api- genin in the inhibition of pancreatic cancer by suppressing the DNA synthesis and proliferation, G2/M stage, cyclin B associated cdc2 activity in these pancreatic cancer cell lines. Gemcitabine has been used as a drug of choice for pancreatic cancer either alone or with a combination of drugs (Orlandi et al., 2016). The compound quercetin is known to possess antioxidant capacity (Thaipong et al., 2005), hypoglycemic and anti-hypotensive effects (Ojewole, 2005; Wang et al., 2005).

Thus, pancreatic cancer has been seen to have an aggressive nature and needs more specific targeting of the drug to bring about an effective inhibition of cancer. This study focuses on the proteins expressed in pancreatic cancer as a target for the in silico molecular interaction with the active compounds present in P. guajava to find out the best binding compounds. The results of this study can be tested further through in vitro and in vivo approaches to bring out an efficient therapeutic molecule.

MATERIALS AND METHODS

Ligand preparation

The P. guajava derived bioactive compounds were selected for the docking studies. Apigenin (CID: 5280443), β-caryophyllene (CID: 5281515), γ-sitosterol (clianosterol) (CID: 94195), Glycolic acid (CID: 757), Ledol (CID: 92812), Quercetin (CID: 5280343), Squalene (CID: 638072), Vitamin E (CID: 14985), Gemcitabine (Commercial drug) (CID: 60750) were retrieved from the PubChem database.

The ligands were converted from sdf to pdb format using PyMol software. The ligand was prepared for docking using the Auto dock tools 4.2 software. Its
torsions were detected and modified and saved as a pdbqt file. This file was used for docking.

**Target protein preparation**
A literature survey was made for the proteins expressed in pancreatic cancer in conditions such as metastasis and drug resistance. The proteins which were expressed in the pancreatic cancer were chosen to be docked. The proteins β-catenin (PDB ID: 3sl9), Mesothelin (PDB ID: 4f3f), CD43 (PDB ID: 1kyj), SNAI1 (PDB ID: 3w5k), P-selectin (PDB ID: 1g1q), HSF1 (PDB ID: 5hdg), CD44 (PDB ID: 1p0z) were retrieved from RCSB PDB database, sequence for Claudin-4 (Uniprot ID: 014493) and PDGF-D (UniProt ID: Q9GZP0) were retrieved from Uniprot database and modelled with Swiss auto model and their 3D Structures were generated.

All proteins were prepared using the Auto dock tools 4.2 software and the pre-attached ligands and water molecules were removed and polar hydrogen and Kollman charges were added. And the file was saved as a new pdb file for docking. The grid box was set to the required size depending on the region of binding pockets present on that particular protein. The molecule was then used for docking.

**Active site preparation**
The active sites of the prepared target proteins were predicted using the Castp server ([Tian et al., 2018](#)). The pdb file was uploaded in the webserver, and the results were obtained in which the active site residues were mentioned. The regions with these active site residues were chosen for setting up the grid box dimension in the Auto dock tools software.

**Molecular docking**
Auto dock vina was used for molecular docking and calculating binding affinities ([Trott and Olson, 2009](#)). Auto dock Vina operates via the command line terminal. The target protein name and the ligand name were specified with pdbqt extension in the input parameter under receptor and ligand for auto dock vina configuration, and the sizes and centers x, y, z, were mentioned for the grid parameters. These configurations were different proteins. Two separate files were created as output, one as a text log file and the other a pdbqt file after running the docking.

**Visualization of docked structures**
Biovia Discovery studio visualizer was used to analyze the docked structures obtained from Auto dock Vina. The ligand interactions were made visible and labelled with the amino acid residues. The binding affinities of the ligands ([*P. guajava* derived bioactive compounds]) were compared with that of the control ligand (Gemcitabine).

### RESULTS AND DISCUSSION

**Molecular docking**
All docking results were retrieved in a log file in text format and output file in pdb format split by the command-line tool. Auto dock results output were obtained as Binding energy (kcal/mol) (negative value) in a log file and structure file (pdbqt). The binding scores of different ligands with different proteins were compared with the standard drug (gemcitabine). The binding energy less than standard drug was considered to be effective ligands and can be finalized for further studies.

![Figure 1: Molecular interaction between β-Catenin and *P. guajava* derived bioactive compounds](image1)

![Figure 2: Molecular interaction between CD43 and *P. guajava* derived bioactive compounds](image2)

**Analysis of docked structures**
The docking results were analyzed using the Dis-
Table 1: Binding scores (kcal/mol) for the molecular interaction of bioactive compounds from *P. guajava* with pancreatic cancer expressed proteins

<table>
<thead>
<tr>
<th>Protein</th>
<th>Beta-catenin</th>
<th>Claudin-4</th>
<th>Mesothelin</th>
<th>CD43-mucin</th>
<th>SNAI1</th>
<th>PDGF-D</th>
<th>P-selectin</th>
<th>HSF1</th>
<th>CD44</th>
</tr>
</thead>
<tbody>
<tr>
<td>(PDB ID: 3sl9)</td>
<td>(Uniprot ID: 014493)</td>
<td>(PDB ID: 4f3f)</td>
<td>(PDB ID: 1kyj)</td>
<td>(PDB ID: 3w5k)</td>
<td>(Uniprot ID: Q9GZP0)</td>
<td>(PDB ID: 1g1q)</td>
<td>(PDB ID: 5hdg)</td>
<td>(PDB ID: 1poz)</td>
<td></td>
</tr>
<tr>
<td>Apigenin</td>
<td>-7.2</td>
<td>-5.6</td>
<td>-7.3</td>
<td>-5.7</td>
<td>-8.1</td>
<td>-7.8</td>
<td>-6.9</td>
<td>-5.5</td>
<td>-6.1</td>
</tr>
<tr>
<td>Beta-caryophyllene</td>
<td>-5.9</td>
<td>-5.5</td>
<td>-6.1</td>
<td>-5.7</td>
<td>-6.3</td>
<td>-5.8</td>
<td>-5.9</td>
<td>-4.7</td>
<td>-5.3</td>
</tr>
<tr>
<td>Gamma sitosterol (Clianosteryl)</td>
<td>-5.9</td>
<td>-6.0</td>
<td>-7.0</td>
<td>-5.7</td>
<td>-6.5</td>
<td>-5.9</td>
<td>-6.0</td>
<td>-5.5</td>
<td>-4.7</td>
</tr>
<tr>
<td>Glycolic acid</td>
<td>-3.9</td>
<td>-3.2</td>
<td>-4.1</td>
<td>-3.2</td>
<td>-4.0</td>
<td>-3.3</td>
<td>-3.7</td>
<td>-3.2</td>
<td>-3.6</td>
</tr>
<tr>
<td>Ledol</td>
<td>-5.7</td>
<td>-4.8</td>
<td>-5.4</td>
<td>-6.0</td>
<td>-6.5</td>
<td>-6.0</td>
<td>-5.3</td>
<td>-4.7</td>
<td>-5.5</td>
</tr>
<tr>
<td>Quercetin</td>
<td>-7.7</td>
<td>-6.4</td>
<td>-7.9</td>
<td>-6.3</td>
<td>-9.6</td>
<td>-6.9</td>
<td>-7.2</td>
<td>-6.2</td>
<td>-6.4</td>
</tr>
<tr>
<td>Squalene</td>
<td>-5.3</td>
<td>-4.2</td>
<td>-4.4</td>
<td>-3.4</td>
<td>-5.6</td>
<td>-3.8</td>
<td>-4.3</td>
<td>-3.7</td>
<td>-4.1</td>
</tr>
<tr>
<td>Vitamin E</td>
<td>-5.6</td>
<td>-5.7</td>
<td>-5.4</td>
<td>-4.4</td>
<td>-6.5</td>
<td>-6.1</td>
<td>-6.2</td>
<td>-4.6</td>
<td>-4.5</td>
</tr>
<tr>
<td>Gemcitabine</td>
<td>-6.5</td>
<td>-5.6</td>
<td>-6.3</td>
<td>-4.9</td>
<td>-6.9</td>
<td>-5.1</td>
<td>-6.3</td>
<td>-4.9</td>
<td>-5.6</td>
</tr>
</tbody>
</table>

Figure 3: Molecular interaction between CD44 and *P. guajava* derived bioactive compounds

Figure 4: Molecular interaction between Claudin-4 and *P. guajava* derived bioactive compounds

From the results obtained from molecular docking, it was found that quercetin, a compound derived from the ethanolic extract of *P. guajava* leaves, has shown to have the highest binding affinity for 8 proteins, namely β-catenin, Claudin-4, Mesothelin, CD43-mucin, SNAI1, P-selectin, CD44, HSF1. Apigenin has shown to have the highest binding affinity for PDGF-D protein.

The results of the interaction and binding affinities have been given below in the order of Figure 1: β-catenin, Figure 2: CD43, Figure 3: CD44, Figure 4: Claudin-4, Figure 5: HSF1, Figure 6: Mesothelin, Figure 7: PDGF-D, Figure 8: P-selectin, Figure 9: SNAI1 along with the ligands in the order of, a. Quercetin, b. Apigenin, c. β-caryophyllene, d. γ-sitosterol, e. Ledol, f. Squalene, g. Vitamin E, h. Glycolic acid, i. Gemcitabine.
Figure 5: Molecular interaction between HSF-1 and *P. guajava* derived bioactive compounds

Figure 6: Molecular interaction between Mesothelin and *P. guajava* derived bioactive compounds

Figure 7: Molecular interaction between PDGF-D and *P. guajava* derived bioactive compounds

Figure 8: Molecular interaction between P-selectin and *P. guajava* derived bioactive compounds

The molecular docking was performed using auto dock vina and the results of the dockings were analyzed using the discovery studio visualizer. The necessity of protein-ligand docking is to find the best fit of ligand to the protein's three-dimensional structure. Nine receptors (proteins) and nine ligands (bioactive compounds from *P. guajava*) were chosen for docking. Among the nine ligands docked, quercetin was the ligand that showed the best docking score for eight proteins, namely β-catenin, Claudin-4, Mesothelin, CD43-mucin, SNAI1, P-selectin, CD44, HSF1 and apigenin was the ligand that showed the best docking score for the protein PDGF-D. The docking scores were represented in terms of binding affinity denoted by kcal/mol.

The docking between the protein SNAI1 and quercetin gave the best score of -9.6 kcal/mol. Quercetin bound to the amino acid residues CYS262, ARG264, ASP339, ASP426, ARG465, ASN469, SER527 of SNA1 1 with 4 H-bonds. The second highest binding affinity was found to be with the docked protein mesothelin with a score of -7.9 kcal/mol with the amino acid residues GLY41, GLY42, LYS43, VAL93 with 3 H-bonds. Apigenin docks with PDGF-D with a binding score of -7.8 at the amino acid residues VAL186, PRO203, LEU205, ALA215, GLU216 with 1 H-bond. Quercetin binds to β-catenin with a binding affinity of -7.7 kcal/mol at the amino acid residues ILE153, PRO154, ARG190, GLN193, ARG274, LEU279 with 3 H-bonds.
Quercetin binds with the protein P-selectin with a binding affinity of -7.2 kcal/mol at the amino acid residues GLN30, ASN31, LEU134, THR136, ILE137, CYS142, PRO151 with 5 H-bonds. The binding affinities of the proteins claudin-4 and CD44 with quercetin were found to be the same - 6.4 kcal/mol with amino acid residues LYS103, GLU105, ALA115 with 2 H-bonds and TYR42, TYR79, ALA98, SER112 with 2 H-bonds, respectively. Quercetin docks with the protein CD43 at the amino acid residues GLN7, GLY102, LYS103, ALA113, ALA115, MET118 at 4 H-bonds with a binding affinity of -6.3. Quercetin binds with the least docking score of -6.2 to the protein HSF1 at the amino acid residues VAL26, ASP32, TRP37, SER38, PRO39, ARG106 with 3 H-bonds. Following the least score with the HSF1 protein.

The commercial drug gemcitabine is commonly used for many types of cancer chemotherapy and especially for pancreatic cancer. This compound was found to have an efficient binding score of -6.9 kcal/mol for the protein SNAI1 with the amino acid residues ASP339, ARG465, SER468, ASN469 with 2 H-bonds. Gemcitabine showed the least binding affinities for two proteins, namely CD43 and HSF1 at the amino acid residues ASP76, ASN142, GLN145, ASP146, ARG158 with 2 H-bonds and MET75, TYR76, HIS101, LYS116, ARG117 at 1 H-bond with scores -4.9 for both the proteins (Table 1). The compound quercetin is a major flavonoid belonging to the class of flavonols. It is mostly found in foods like cauliflower, nuts, tea, apples, guavas and berries.

The anti-oxidant activity of the quercetin is responsible for its radical scavenging, lipid peroxidation inhibition and metal chelation mechanisms in vitro (Rice-Evans et al., 1996). The drug-likeness of quercetin was very good which can be understood by its molecular properties. It has 5 hydrogen bond donors (nOHNH) and 7 hydrogen bond acceptors (nON) with an octal water partial coefficient (log P) of 1.683, having a molecular weight of 302.238 g/mol, 22 non-hydrogen atoms, a single rotatable bond, the polar surface area of 131.351 A² and a molecular volume of 240.084.

The overall drug-likeness score for quercetin was reported 1.00 which is actually a good score (Islam et al., 2013). Compound quercetin was found to have a binding affinity of - 4.52 kcal/mol against cellular tumor antigen p53 in a human cervical cancer cell line (HELA) and bound with NF- kappa B with the least score of - 2.83 kcal/mol. (Muthukala et al., 2015).

Furthermore, lung cancer proteins viz., p53, caspase 3 and mucosal addressin cell adhesion molecule 1 were targeted for apigenin docking, which showed - 4.6 kcal/mol of binding affinity against the p53 protein, - 5.7 kcal/mol against caspase3, and - 5.3 kcal/mol against mucosal addressin cell adhesion molecule 1 (Kasilingam and Elengoe, 2018) and a drug-likeness score of apigenin was found to be 0.940 (Hashemi, 2012).

From the above information on the performed molecular docking studies and the available literature, it is understood that the compounds quercetin and apigenin have been found to be potent anti-cancer agent against pancreatic cancer and can be therapeutically effective.

CONCLUSION

The overall studies on the in silico molecular interaction between pancreatic cancer expressed proteins and P. guajava derived bioactive compounds using the Auto dock vina has been carried out efficiently and the binding affinities of the compounds have been identified. Quercetin and apigenin have been found to be the compounds with good binding affinities towards pancreatic cancer expressed proteins. Therefore, it has been well understood from the above approaches that specific compounds such as quercetin and apigenin present in the P. guajava have the potential to be developed as a very effective anti-cancer drug in future, provided its solubility, pharmacokinetics and toxicity level must be optimized through dry and wet lab approaches. Further studies will look forward to investigating this anti-cancer activity in depth by molecular dynamics simulation, in vitro and in vivo studies and gene expression analysis.
ACKNOWLEDGEMENT

The authors express their sincere thanks to the Head, Department of Biotechnology, President, and the Management of Dr. M. G. R Educational and Research Institute, Chennai, India, for their constant encouragement and support throughout the academic years. The authors also express their gratitude to the staffs and students of the Department of Biotechnology.

Conflict of Interest

The authors declare that they have no conflict of interest for this study.

Funding Support

The authors declare that they have no funding support for this study.

REFERENCES


