1

2

Title:

A new Virtual Screening approach for Protein Disulfide Isomerase inhibitors reveals potential candidates for antithrombotic agents

6
7

Authors

Noureddine Ben Khalaf [1*] and Moiz Bakhiet[1]

Authors affiliation

[1]*HH. Al-Jawhara Centre for Molecular Medicine and Inherited Diseases,The Arabian Gulf University, Manama, Bahrain*

*Corresponding author:

Dr. Noureddine Ben Khalaf, Al-Jawhara Centre for Molecular Medicine and Inherited Diseases, Building 61 King Abdulaziz Avenue, Block 328, Manama Bahrain. E-mail: noureddinek@agu.edu.bh

27

28

29

30

31

32

33

34

1

Abstract

Background: Arterial thrombosis causes heart attacks and strokes and constitutes one of the leading causes of morbidity and mortality in the world and few therapies are available for its treatment. Thus, new therapeutic approaches in the prevention and treatment of arterial thrombosis are needed. Protein disulfide isomerase (PDI) has been shown to be expressed on vascular cells following injury and to be involved in regulating thrombus formation in vivo. Since inhibition of PDI prevents platelet accumulation and fibrin generation, it makes it a valuable target for the development of new antithrombotics. Rutin, a flavonol glycoside derivative of Quercetin, was previously described for displaying decent potency against PDI, and it inhibited the agonist-induced platelets aggregation *in vivo*, however its utility is limited by its low solubility and its off-target activity. Rutin was recently reported to bind specifically to the b' domain of PDI affecting protein flexibility which results in the inhibition of its reductase activity. To investigate Rutin inhibitory mechanism we used docking and molecular dynamics simulation and we observed that Rutin binds to a specific hydrophobic pocket of the b' domain which reduces PDI flexibility. **Methods:** In an attempt to identify more potent, soluble and specific PDI inhibitors, we established an *in silico* approach based on similarity search in Zinc Drug-like library composed of more than 17 million compounds satisfying Lipiniski's rule of five. A KNIME workflow was established for selecting Rutin-similar compounds based on Tanimoto coefficients. Then, a virtual screening of selected compounds was performed using Autodock Vina on PDI target pocket. In order to select PDI specific probes, a counter-screen was run to eliminate hits binding Erp57 thioredoxin active site. Hits were then submitted to druglikeness prediction using Quantitative Estimate of Druglikeness (QED). A total of 5 compounds were selected and submitted to re-docking with Autodock Vina. Complexes were subject to Molecular Dynamics simulation using NAMD. **Results and Discussion:** a total of 4 compounds were shown to form stable complexes with PDI binding pocket and then could constitute promising candidates for lead optimization. In conclusion, our *in silico* approach lead to the identification of potential novel PDI inhibitors that may form suitable candidates for Arterial thrombosis drug discovery.

Keywords: Protein Disulfide Isomerase, Thrombosis, Virtual Screening, Molecular docking, Molecular Dynamics

2

Introduction

Arterial thrombosis causes heart attacks and strokes and is one of the leading causes of morbidity and mortality in the world. However, few adequate therapies are available for treating arterial thrombosis. Protein Dislufide Isomerase (PDI), the founding member of a large family of thiol oxydoreductases, was shown to be required for thrombosis, hemostasis and vascular inflammation (Cho 2013). PDI was discovered 50 years ago as the first folding catalyst. Since its discovery, it was demonstrated that PDI acts as a dithiol–disulfide oxidoreductase capable of reducing, oxidizing and isomerizing disulfide bonds. Independently of its redox activity, PDI can also act as a chaperone both in vitro (Cai et al. 1994) and in vivo (McLaughlin & Bulleid 1998).

Thioredoxin family comprises 20 members that vary in length and domain arrangement. Most PDI family members share in common catalytic and non-catalytic thioredoxin-like domains. PDI is organized in four thioredoxin-like domains, a, a', b and b', in addition to a linker domain; x. The catalytic domains a and a' contain catalytic CGHC motifs reacting with thiol groups in substrate proteins. Non catalytic domains b and b' were shown to be involved in substrate recognition and recruitment (Kozlov et al. 2010).

Although PDI is assisting newly synthesized proteins to fold in the endoplasmic reticulum (ER) of majority of cells (Vaux et al. 1990), extracellular roles for PDI have been reported by previous studies, namely in the initiation of thrombus formation (Cho et al. 2008). Upon vascular injury, endothelial cells and platelets are activated and secrete PDI and other thiol isomerases. PDI, ERp5 and ERp57 have been shown to be involved in the initiation of thrombus formation in vivo (Furie & Flaumenhaft 2014). Following laser-induced vascular injury, PDI was demonstrated to be secreted from activated endothelial cells, then secreted from the bound platelets, to be finally

3

87 associated with thrombus growth. Inhibition of PDI with a blocking antibody completely

88 inhibited both platelet thrombus formation and fibrin generation (Cho et al. 2008; Jasuja et al.

89 2010). All these data point towards PDI as valuable and emerging drug target for thrombosis.

90 Quercetin-3-Rutinoside, known also as Rutin, a flavonol glycoside derivative of Quercetin, was

91 previously described for displaying decent potency against PDI (IC_{50}=6 μM), and it inhibited the

92 agonist-induced platelets aggregation at 30 μM (Jasuja et al. 2012), however its utility is limited

93 by its low solubility and its off-target activity. Interestingly, a recent study demonstrated that

94 Rutin inhibits PDI by binding directly to its b'x domain and not to the catalytic domains as

95 expected (Lin et al. 2015), which revealed new insights into PDI inhibition mechanisms.

96 In an attempt to identify more potent, soluble and specific PDI inhibitors, we established an *in*

97 *silico* approach based on similarity search for Quercetin-similar compounds. A virtual screening

98 followed by Molecular Dynamics Simulation allowed the identification of 4 compounds that

99 were shown to form specific and stable complexes with PDI b'x domain and then could

100 constitute promising candidates for lead optimization in preventing arterial thrombosis.

101 **Material and Methods**

102 **Ligand library preparation**

103 A subset of 17,900,742 Drug-like compounds was obtained in SDF format from ZINC database

104 (http://zinc.docking.org/). The compounds were filtered according to Lipinsky's rule of five (

105 MW <= 500, MW >= 150, XlogP <= 5, Rb <=7 and PSA < 150, Hbond_donors <= 5 ,

106 Hbond_acceptors <= 10) (Lipinski 2000). Rutin structures were sketched and obtained from

107 Pubchem database (PubChem CID:5280805). Quercetin was obtained from the ZINC database

108 (http://zinc.docking.org/; ZINC03869685). A KNIME (Nicola et al. 2015) workflow was set to

4

109   filter compounds with similarity to Quercetin according to Figure 1. In brief, compounds in SDF

110   format were read and Morgan fingerprints were generated using RDKit KNIME node. A

111   Tanimoto index (Bajusz et al. 2015) for similarity was calculated and a total number of 9533

112   compounds showing a similarity index more than 0.3 to Quercetin were selected for virtual

113   screening.

Target Preparation

115   Reduced human Protein Disulfide Isomerase (PDI) and ERp57 crystal structure (PDB ID: 4EKZ

116   (Wang et al. 2013) and PDB ID: 3F8U (Dong et al. 2009)) were obtained from RCSB database

117   (http://www.rcsb.org/) . Residues numbers were edited in the pdb file and input structure was

118   prepared by removing all water molecules, adding and merging non polar Hydrogen, and

119   computing Gasteiger charges, using Autodock tools (Morris et al. 2009). Structures were

120   converted to PDBQT format for docking with Autodock Vina (Morris et al. 2009).

Pocket analysis

122   Binding sites on PDI and Erp57 were analyzed using Pockets plugin for Vegazz software.

123   Pockets is the graphic interface for fpocket developed by Vincent Le Guilloux and Peter

124   Schmidtke to detect protein cavities. This program uses an optimized algorithm based on

125   Voronoi tessellation that is very fast and allows to analyze large molecules waiting for a

126   reasonable time. Parameters were set as follows: radius of a-sphere ranging from 0.3 to 0.6,

127   minimum number of a-sphere = 30, minimum apolar neighboring = 3, maximum distance for

128   clustering = 1.73, maximum distance for single linkage = 2.5, maximum distance between

129   spheres = 4.5, and 2500 iterations for volume calculation. A score for each pocket was calculated

130   using a scoring function of different pocket descriptors including Number of alpha spheres,

131 Density of the cavity, Polarity Score, Mean local hydrophobic density, Proportion of apolar alpha

132 spheres, Composition of amino acids, Maximum distance between two alpha sphere,

133 Hydrophobicity , Charge, Volume and B-factor scores. Output pdb result were visualized and

134 analyzed by Pymol.

135 **Rutin Docking and Pose prediction**

136 Rutin structure was converted to pdbqt format using OpenBabel (O'Boyle et al. 2011) and

137 docked to PDI using Autodock Vina (Trott & Olson 2010). Grid boxes were centered on PDI

138 binding pockets at the b' domain and box size was set to 13824 $\mathring{A}^3$. Docking was performed with

139 exhaustiveness=48 and 9 poses/run. Best Rutin conformation (E=-8 Kcal/mol) was saved to

140 generate protein-ligand complexes for molecular dynamics study and protein-ligand interactions.

141 A docking on ERp57 equivalent site was also performed for Rutin using the same approach as

142 for PDI.

143 **Virtual Screening**

144 Filtered compounds were subject to a first round docking to PDI b' domain using Autodock Vina

145 (Trott & Olson 2010) with an exhaustiveness of 12 and an affinity cutoff of -8.8 Kcal/mol

146 (Average affinity + 2SD). Selected compounds (69 conformations were selected) were subject to

147 a counter screening on ERp57 b' domain for target selectivity. All docking were performed using

148 Autodock Vina (Trott & Olson 2010) with an exhaustiveness of 12 and a grid box of 13824 $\mathring{A}^3$

149 centered on b' domain. A scoring function was established for compound ranking as follows:

150 $$Score = E_{PDI}*(1 + ((E_{Rutin\_Erp57}-E_{Erp57})/E_{Rutin\_Erp57}))$$

6

151   Where E_{PDI} are compound affinities for PDI b' domain, E_{Erp57} is compound affinity for Erp57 is

152   compound affinity, and E_{Rutin_Erp57} is Rutin affinity for Erp57.

153   Compounds were ranked according to the scoring function, a total of 23 were selected with

154   respective score higher than Rutin's score. Selected Hits were subject to Drug-likeness filtering

155   using DruLito(Bickerton et al. 2012) by applying Unweighted Quantitative estimate of drug-

156   likeness filter with a cut-off of 0.5 (QED)). A total of five molecules were shown to satisfy those

157   filters and were subject to a second round docking to PDI using Autodock Vina. Grid boxes were

158   centered on b' domain and box size was set to 13824 $Å^3$. Docking was performed with

159   exhaustiveness=48 and 9 poses/run. Best conformations for were saved to generate protein-

160   ligand complexes used for protein-ligand interaction analysis by lig-Inetraction plugin of

161   Maestro (Schrödinger Release 2015-4: Maestro, version 10.4, Schrödinger, LLC, New York,

162   NY, 2015).

163   **Molecular Dynamics Simulation**

164   Based on the docking results, molecular dynamics simulations were performed for PDI alone or

165   in complex with a total of 6 ligands; five of them were issued from the virtual screening

166   workflow, in addition to the previously described PDI inhibitor; Rutin (Jasuja et al. 2012).

167   Swissparam server (Zoete et al. 2011) was used to generate ligand topology and parameters files

168   based on the Merck molecular force field (MMFF) and compatible with the CHARMM force

169   field. PDI alone or in complex with ligand was solvated in a cubic Waterbox with periodic

170   boundary condition and a 2.4 Å layer of water for each direction of the coordinate structure using

171   the VMD solvation plugin (Humphrey et al. 1996). All MD simulation were performed using

172   NAMD (NAnoscale Molecular Dynamics program; v 2.9) (Phillips et al. 2005). Structures were

7

173   relaxed through 50000 steps of steepest descent energy minimization followed by 1 ns NVT. For

174   the production run, we used 1 fs time step, switching distance of 9 Å, cutoff of 10 Å, electrostatic

175   contribution evaluated every 2 steps, Particle Mesh Ewald with 1 Å spacing, NVT ensemble and

176   Langevin thermostat with 310 K as target temperature and pressure of 1.013 bar. Trajectories

177   were analyzed by VMD and VegaZZ software.

178   **Results**

179   **Docking Analysis**

180   Human Protein Disulfide Isomerase (PDI) and ERp57 crystal structure were prepared for

181   docking with Autodock Vina (Morris et al. 2009). Pocket analysis showed the presence of 31

182   potential binding sites on PDI, among them, a pocket lying in the b' domain showed the highest

183   overall score of 34.5714 according to fpocket scoring function with a total volume of 1526.7041

184    $A^{\circ 3}$. Figure 2 shows the pocket surface representation with a hydrophobic cavity surrounded by

185   charged (E239 and D297) and polar (N298 and T428) residues. This pocket was selected for

186   docking analysis of Rutin binding mode. Ligands were prepared and docked using

187   AutodockVina using a Grid of 13824 $A^{\circ 3}$ centered on the considered pocket. Docking poses

188   fitting to the considered pocket were selected and were analyzed manually, the one with the

189   lowest binding energy (-8.0 kcal/mol) were selected for building protein-ligand complex for

190   molecular dynamics simulations. The same approach was applied to Erp57 (lowest energy = -7.3

191   Kcal/mol).

192   **Similarity Search and Virtual Screening**

193   A subset of 17,900,742 Drug-like compounds filtered according to Lipinsky's rule of five from

194   the ZINC database was analyzed by a KNIME workflow (see Figure 1) to filter compounds

8

195   showing a similarity Tanimoto index above 0.3 comparing to Quercetin structure. A total number

196   of 9533 compounds were selected and were subject to a first round docking to PDI b' domain

197   then to a counter screening on ERp57 b' domain for target selectivity.

198   Compounds were ranked according to the scoring function, a total of 23 were selected with

199   respective score higher than Rutin's score (see Table 1). Drug-likeness filtering was performed

200   using DruLito(Bickerton et al. 2012) and identified five molecules, listed in Table 3. Selected

201   compounds were subject to a second round docking to PDI b'domain and best conformations for

202   were used to generate protein-ligand complexes for protein-ligand interactions and molecular

203   dynamics study.

204   **Ligand-protein Interaction analysis**

205   Ligand-protein interaction analysis Rutin fitting into the hydrophobic cavity in the b' domain

206   pocket through its Quercetin backbone as shown by Figure 3. Rutin forms H-bonds between

207   hydroxyl groups and PDI charged residues E239 and D297 and displays a significant exposure to

208   solvent mainly through the Rutinoside' hydroxyl groups.

209   Similar poses were predicted to the selected ligands. Figure 4 shows the superposition of all

210   compounds inside the pocket. Figure 5 showed all the compounds having important surface

211   interaction with the hydrophobic cavity with a partial exposure to solvent and a proximity to

212   charged residues. No polar bonds were predicted. Except for ZINC19928318, all the compounds

213   established Pi-Pi stacking interaction with F304.

214

215

9

Molecular Dynamics Simulation

For Molecular Dynamics Simulations, PDI alone or in complex with ligand was solvated in a cubic Waterbox with periodic boundary condition Structures were relaxed through 50000 steps of steepest descent energy minimization followed by 500 *ps* NVT with 310 K as target temperature and pressure of 1.013 bar. Trajectories were analyzed by VMD and MD simulation results are summaraized in Table 2. Trajectory analysis for PDI alone showed an average RMSD of 4.91 Å ± 0.37 for the total protein. Rutin was shown to form a more stable complex with PDI as shown by Figure 6 with an average RMSD of 4.76 Å ±0.35. Among selected ligands, four were shown to stabilize PDI structure; ZINC19928318, ZINC24834252, ZINC24601822 and ZINC24601767 with RMSDs values of 4.78±0.36, 4.77±0.37, 4.76±0.365 and 4.75±0.36, respectively in a similar manner to Rutin. Only one compound ZINC19632922 showed an increase in RMSD average comparing to PDI alone (4.97±0.38).

Discussion

PDI is secreted by platelets and endothelium cells following vascular injury and is shown to bind to β3 integrins on activated cells causing thrombus initiation (Cho et al. 2008). Blocking of PDI by a specific antibody reduced significantly both platelet thrombus formation and fibrin generation in mouse thrombosis model (Flaumenhaft et al. 2015). Quercetin-3-Rutinoside; Rutin, was reported as PDI reductase activity inhibitor and a potent anti-thrombotic agent in vitro and in vivo (Jasuja et al. 2012). In a recent study published by Lin *et al.* (Wang et al. 2013), Rutin was reported to directly bind to the b' domain of PDI with a 1:1 stoichiochemistry restricting conformational flexibility of the protein allowing a more compact conformation. PDI b' fragment was shown to contain the major binding site of Rutin and the infusion of the b'x

238   fragment in mouse thrombus model reversed Rutin inhibition of platelet thrombus formation

239   (Wang et al. 2013).

240   In an attempt to investigate the structural insights of PDI inhibition by Rutin, we

241   analyzed pocket distribution in the full length protein structure of reduced human Protein

242   Disulfide Isomerase (PDI), PDB ID: 4EKZ (4), and we identified a pocket with significant score

243   that can be used for Rutin binding prediction in the b' domain of the protein (see Figure 2). The

244   pocket contains a cavity with a hydrophobic environment surrounded by charged and polar

245   residues, as shown by figure 2. We used docking simulation to predict the best conformation for

246   Rutin binding (see Figure 3) in the considered pocket. Interaction analysis showed Rutin partial

247   interaction with the hydrophobic cavity of the binding site and significant exposure to solvent,

248   mainly through the Rutinoside hydroxyl groups. H-bonds were predicted between Rutin

249   hydroxyl groups and PDI charged residues E239 and D297. In the study of Jasuja *et al.* (Jasuja et

250   al. 2012). the evaluation of structure activity relationships demonstrated that a sugar at 3'

251   position in the C ring of quercetin-3-Rutinoside is critical for its ability to inhibit PDI (see Table

252   3). All analogs tested with a sugar in this position inhibited PDI, while analogs lacking this sugar

253   failed to demonstrate inhibition.

254   In order to investigate the binding effect of Rutin on PDI stability, we carried Molecular

255   Dynamics simulation. MD simulation showed PDI to have a significant flexibility with an

256   average RMSD of 4.91 Å $\pm$ 0.37. The Interdomain flexibility was previously reported in several

257   studies, namely on Yeast PDI (Tian et al. 2008; Tian et al. 2006) where interdomain flexibility

258   was reported to be essential for protein catalytic activity and this may be the main reason why

259   PDI was resistant to crystallization efforts for a long time. Among all PDIs with x-linker,

260   mobility between the b' and the a' domain was reported to be more pronounced than between a

11

261   and b domains (Kozlov et al. 2010). Although, a and a' domains are essential for the catalytic

262   activity of the enzyme, b' domain constitutes the major binding site for unfolded substrate

263   proteins and plays a central role in substrate recognition. Indeed, homology analysis showed

264   significant divergence of b' domain among PDI superfamily allowing substrate specificity

265   (Kozlov et al. 2010). Structural investigation showed b' hydrophobic cavity located between

266   helices α1 and α3 and formed by several residues; Phe223, ala228, Phe232, Ile284, Phe287,

267   Phe288, Leu303 and Met307 side chains, is involved in substrate recognition and binding (Byrne

268   et al. 2009). In our study, Rutin displayed predicted hydrophobic exposure to many of these

269   residues, in addition, Rutin was able to establish polar contact with charged residues surrounding

270   the cavity and to be exposed to the solvent. We hypothesize that the binding mode of Rutin to the

271   b' domain pocket allows stability of the protein-ligand complex and hence limits protein

272   interdomain flexibility between b' and a' domains. This hypothesis was confirmed *in silico* by

273   Molecular Dynamics simulation. In fact, Rutin was shown to form a more stable complex with

274   an average RMSD of 4.76 Å ±0.35 comparing to PDI alone. In addition to the hydrophobic

275   interaction with the protein b' cavity, solvent exposure of the ligand and polar contacts with

276   surrounding residues appear to be essential to stabilize the complex and reduce protein flexibility

277   which results in catalytic activity inhibition.

278   Despite its potency, the low solubility and off-target activity of Rutin limited its further

279   use as antithrombotic agent. In order to identify potential novel PDI Inhibitors with interesting

280   pharmacological properties, we used Quercetin structure to search for analog compounds that

281   could satisfy binding mode constraints for PDI catalytic activity inhibition, in a similar way to

282   Rutin. A Morgan Fingerprint filtering workflow allowed the identification of 9533 compounds

283   with a similarity Tanimoto coefficient above 0.3. Compounds were docked to the binding pocket

284 of PDI b' domain using a virtual screening approach. A counter screen was run using Erp57 as

285 target since equivalent b' pocket is divergent with human PDI' one (20% homology) and

286 contains negatively charged residues (Kozlov et al. 2010). Following screening workflow,

287 compounds were ranked according to the scoring function and a total of 23 were selected for

288 Drug-likeness filtering. Five compounds were selected and re-docked to PDI b' domain. The

289 compounds are listed in Table 3. Interaction analysis showed all the compounds displaying

290 surface exposure to the hydrophobic cavity through the benzo-furan group of ZINC19928318,

291 ZINC24601822; ZINC24601767 as well as the naphthalene group of ZINC24834252. The

292 benzyl residue attached to the piperazin allows partial exposure to solvent and a proximity to

293 charged residues despite that no polar bonds was predicted. Except for ZINC19928318, all the

294 compounds established Pi-Pi stacking interaction with F304, a residue part of the hydrophobic

295 cavity of b' domain. These structural features are coherent with the Rutin inhibitory binding

296 mode. Selected compounds were finally used for MD simulation in complex with PDI. As

297 expected, four compounds were shown to stabilize PDI structure and reduce protein flexibility;

298 ZINC19928318, ZINC24834252, ZINC24601822 and ZINC24601767 with respective RMSDs

299 lower than PDI alone, in a similar way to Rutin (see Table 2). We believe that adding charged

300 groups on the benzyl group attached to piperazin may increase protein-ligand stability by

301 establishing polar bonds with charged residues harboring the binding pocket, and hence limits

302 protein flexibility and catalytic activity. Only one compound ZINC19632922 showed an increase

303 in RMSD (see Table 2) which can be due to benzo-furan group solvent exposure (see Figure 5).

304 This compound is a drug known as Befuraline, it was previously reported as inhibitor of

305 Proteinase-activated receptor 1, and Anandamide amidohydrolase (Planty et al. 2010; Vincent et

306 al. 2009).

13

Conclusion

In conclusion, we reported in this study new insights into PDI inhibition mechanism through structure-activity relationship and molecular dynamics simulation of a reported PDI inhibitor; Rutin. We have successfully established a new virtual screening workflow to identify novel and pharmacologically relevant compounds that can constitute potential leads in targeting PDI-catalyzed thrombus formation and a new class of anti-thrombotic agents.

References

Bajusz D, Racz A, and Heberger K. 2015. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *J Cheminform* 7:20. 10.1186/s13321-015-0069-3

Bickerton GR, Paolini GV, Besnard J, Muresan S, and Hopkins AL. 2012. Quantifying the chemical beauty of drugs. *Nat Chem* 4:90-98. 10.1038/nchem.1243

Byrne LJ, Sidhu A, Wallis AK, Ruddock LW, Freedman RB, Howard MJ, and Williamson RA. 2009. Mapping of the ligand-binding site on the b' domain of human PDI: interaction with peptide ligands and the x-linker region. *Biochem J* 423:209-217. 10.1042/BJ20090565

Cai H, Wang CC, and Tsou CL. 1994. Chaperone-like activity of protein disulfide isomerase in the refolding of a protein with no disulfide bonds. *J Biol Chem* 269:24550-24552.

Cho J. 2013. Protein disulfide isomerase in thrombosis and vascular inflammation. *J Thromb Haemost* 11:2084-2091. 10.1111/jth.12413

Cho J, Furie BC, Coughlin SR, and Furie B. 2008. A critical role for extracellular protein disulfide isomerase during thrombus formation in mice. *J Clin Invest* 118:1123-1131. 10.1172/JCI34134

Dong G, Wearsch PA, Peaper DR, Cresswell P, and Reinisch KM. 2009. Insights into MHC class I peptide loading from the structure of the tapasin-ERp57 thiol oxidoreductase heterodimer. *Immunity* 30:21-32. 10.1016/j.immuni.2008.10.018

Flaumenhaft R, Furie B, and Zwicker JI. 2015. Therapeutic implications of protein disulfide isomerase inhibition in thrombotic disease. *Arterioscler Thromb Vasc Biol* 35:16-23. 10.1161/ATVBAHA.114.303410

Furie B, and Flaumenhaft R. 2014. Thiol isomerases in thrombus formation. *Circ Res* 114:1162-1173. 10.1161/CIRCRESAHA.114.301808

Humphrey W, Dalke A, and Schulten K. 1996. VMD: visual molecular dynamics. *J Mol Graph* 14:33-38, 27-38.

Jasuja R, Furie B, and Furie BC. 2010. Endothelium-derived but not platelet-derived protein disulfide isomerase is required for thrombus formation in vivo. *Blood* 116:4665-4674. 10.1182/blood-2010-04-278184

14

342 Jasuja R, Passam FH, Kennedy DR, Kim SH, van Hessem L, Lin L, Bowley SR, Joshi SS, Dilks JR,
343   Furie B, Furie BC, and Flaumenhaft R. 2012. Protein disulfide isomerase inhibitors
344   constitute a new class of antithrombotic agents. *J Clin Invest* 122:2104-2113.
345   10.1172/JCI61228
346 Kozlov G, Maattanen P, Thomas DY, and Gehring K. 2010. A structural overview of the PDI
347   family of proteins. *FEBS J* 277:3924-3936. 10.1111/j.1742-4658.2010.07793.x
348 Lin L, Gopal S, Sharda A, Passam F, Bowley SR, Stopa J, Xue G, Yuan C, Furie BC, Flaumenhaft R,
349   Huang M, and Furie B. 2015. Quercetin-3-rutinoside Inhibits Protein Disulfide Isomerase
350   by Binding to Its b'x Domain. *J Biol Chem* 290:23543-23552. 10.1074/jbc.M115.666180
351 Lipinski CA. 2000. Drug-like properties and the causes of poor solubility and poor permeability. *J
352   Pharmacol Toxicol Methods* 44:235-249.
353 McLaughlin SH, and Bulleid NJ. 1998. Thiol-independent interaction of protein disulphide
354   isomerase with type X collagen during intra-cellular folding and assembly. *Biochem J* 331
355   (Pt 3):793-800.
356 Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, and Olson AJ. 2009.
357   AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J
358   Comput Chem* 30:2785-2791. 10.1002/jcc.21256
359 Nicola G, Berthold MR, Hedrick MP, and Gilson MK. 2015. Connecting proteins with drug-like
360   compounds: Open source drug discovery workflows with BindingDB and KNIME.
361   *Database (Oxford)* 2015. 10.1093/database/bav087
362 O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, and Hutchison GR. 2011. Open
363   Babel: An open chemical toolbox. *J Cheminform* 3:33. 10.1186/1758-2946-3-33
364 Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, and
365   Schulten K. 2005. Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781-
366   1802. 10.1002/jcc.20289
367 Planty B, Pujol C, Lamothe M, Maraval C, Horn C, Le Grand B, and Perez M. 2010. Exploration of
368   a new series of PAR1 antagonists. *Bioorg Med Chem Lett* 20:1735-1739.
369   10.1016/j.bmcl.2010.01.050
370 Tian G, Kober FX, Lewandrowski U, Sickmann A, Lennarz WJ, and Schindelin H. 2008. The
371   catalytic activity of protein-disulfide isomerase requires a conformationally flexible
372   molecule. *J Biol Chem* 283:33630-33640. 10.1074/jbc.M806026200
373 Tian G, Xiang S, Noiva R, Lennarz WJ, and Schindelin H. 2006. The crystal structure of yeast
374   protein disulfide isomerase suggests cooperativity between its active sites. *Cell* 124:61-
375   73. 10.1016/j.cell.2005.10.044
376 Trott O, and Olson AJ. 2010. AutoDock Vina: improving the speed and accuracy of docking with
377   a new scoring function, efficient optimization, and multithreading. *J Comput Chem*
378   31:455-461. 10.1002/jcc.21334
379 Vaux D, Tooze J, and Fuller S. 1990. Identification by anti-idiotype antibodies of an intracellular
380   membrane protein that recognizes a mammalian endoplasmic reticulum retention
381   signal. *Nature* 345:495-502. 10.1038/345495a0
382 Vincent F, Nguyen MT, Emerling DE, Kelly MG, and Duncton MA. 2009. Mining biologically-
383   active molecules for inhibitors of fatty acid amide hydrolase (FAAH): identification of
384   phenmedipham and amperozide as FAAH inhibitors. *Bioorg Med Chem Lett* 19:6793-
385   6796. 10.1016/j.bmcl.2009.09.086

386   Wang C, Li W, Ren J, Fang J, Ke H, Gong W, Feng W, and Wang CC. 2013. Structural insights into
387   the redox-regulated dynamic conformations of human protein disulfide isomerase.
388   *Antioxid Redox Signal* 19:36-45. 10.1089/ars.2012.4630
389   Zoete V, Cuendet MA, Grosdidier A, and Michielin O. 2011. SwissParam: a fast force field
390   generation tool for small organic molecules. *J Comput Chem* 32:2359-2368.
391   10.1002/jcc.21816

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407 **Figures Legends**

408 Figure 1. KNIME workflow for database filtering. ZINC Drug-like database is read in SDF
409 format and molecule descriptors are calculated. Morgan fingerprint is generated for each
410 molecule and compared to Rutin fingerprint using RDkit Fingerprint similarity node.
411 Compounds showing a Tanimoto coefficient above 0.3 are selected and written in SDF format.

412

413 Figure 2. Surface representation of the target pocket in b' domain. The pocket is organized in a
414 hydrophobic cavity (Green) surrounded by charged residues (Red) and polar residues (Light
415 Blue).

416

417 Figure 3. Ligand-protein interaction analysis for PDI-Rutin complex. A. Rutin is fitting into the
418 hydrophobic cavity in the b' domain pocket through its Quercetin backbone. B. Rutin displays a
419 significant exposure to solvent mainly through the Rutinoside' hydroxyl groups which form H-
420 bonds with PDI charged residues E239 and D297 harboring the binding pocket.

421

422 Figure 4. Superposition of docking poses of selected compounds inside the b'domain binding
423 pocket. The hydrophobic cavity is represented in green, by charged residues in red and polar
424 residues in light blue.

425

426 Figure 5. Protein Ligand Interaction Plot.

427

428 Figure 6. RMSD variation for PDI alone and in complex with Rutin. Rutin forms a stable
429 complex with PDI reducing the protein flexibility.
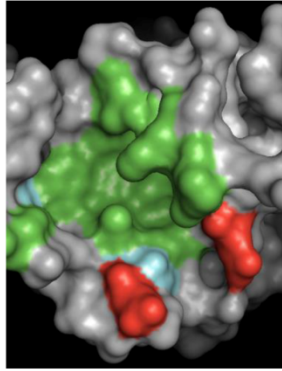
430

431

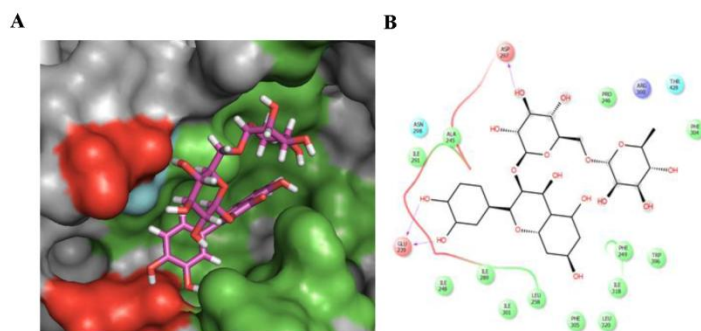432

433
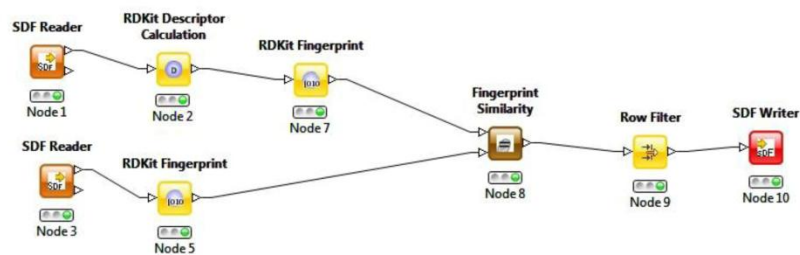
17

434   Figure 1

435

436   Figure 2

437
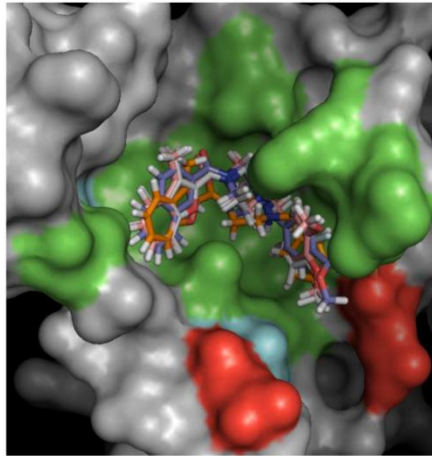
438   Figure 3

439

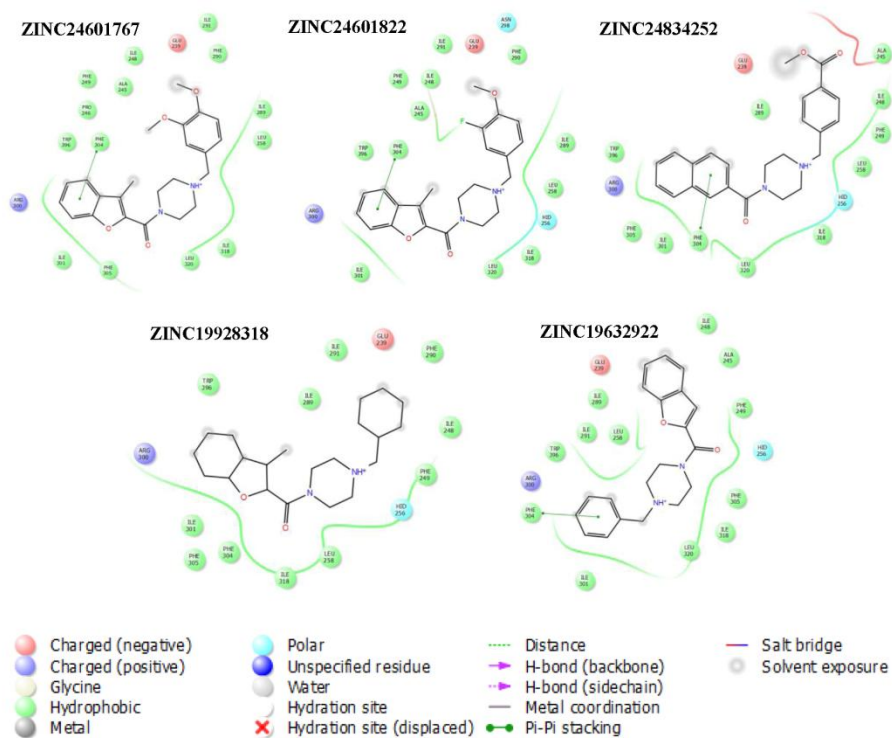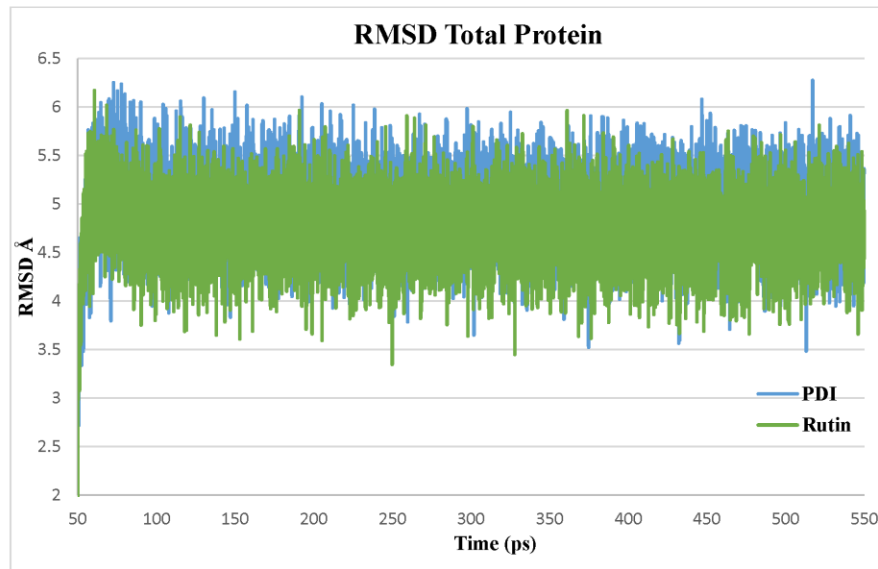440   Figure 4

441

442   Figure 5

443

444   Figure 6

445

446 **Table 1. Ranking of docking energies of selected compounds against PDI and Erp57.**

Number	Name	Energy PDI (kcal/mol)	Energy Erp57 (kcal/mol)	Score
1	ZINC20464669	-9.4	-7.3	-9.4
2	ZINC23460729	-10	-7.8	-9.31507
3	ZINC24834252	-9.2	-7.4	-9.07397
4	ZINC19928318	-9.4	-7.6	-9.0137
5	ZINC13687653	-9.1	-7.4	-8.97534
6	ZINC19632922	-9.2	-7.5	-8.94795
7	ZINC24601822	-9.2	-7.5	-8.94795
8	ZINC23909252	-9.5	-7.9	-8.71918
9	ZINC24601767	-8.8	-7.4	-8.67945
10	ZINC23957133	-8.9	-7.6	-8.53425
11	ZINC21868339	-9.2	-7.9	-8.44384
12	ZINC11771953	-8.8	-7.6	-8.43836
13	ZINC14977925	-9	-7.8	-8.38356
14	ZINC22830797	-9.6	-8.3	-8.28493
15	ZINC12385770	-9.2	-8.1	-8.19178
16	ZINC13678813	-8.9	-7.9	-8.16849
17	ZINC16474806	-8.9	-7.9	-8.16849
18	ZINC19889230	-8.9	-7.9	-8.16849
19	ZINC20731042	-9.3	-8.2	-8.15342
20	ZINC20045482	-9.1	-8.1	-8.10274
21	ZINC77200565	-9.1	-8.1	-8.10274
22	ZINC19700487	-8.8	-7.9	-8.07671
23	ZINC12439199	-8.9	-8	-8.04658
24	Rutin	-8	-7.3	-8

447

448

449

450

451

452

24

453

Table 2. Average RMSD for Protein-Ligand complexes

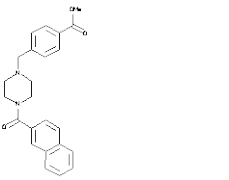| Protein	Ligand	RMSD (Å)	SD
PDI	-	4.91	0.37
PDI	Rutin	4.76	0.35
PDI	ZINC19928318	4.78	0.36
PDI	ZINC24834252	4.77	0.37
PDI	ZINC24601822	4.76	0.36
PDI	ZINC24601767	4.74	0.36
PDI	ZINC19632922	4.97	0.37

454

455

456

457

458

459

460

461

462

463

464

465

466

25
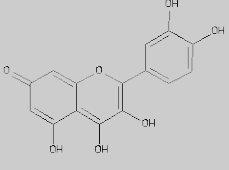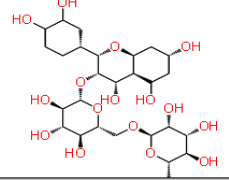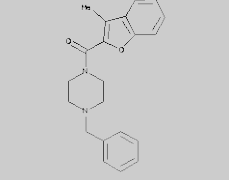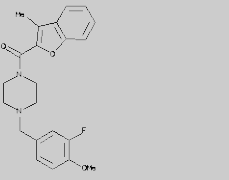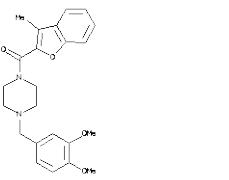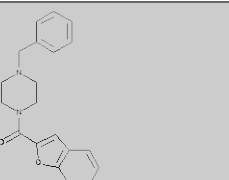
467

Table 3. Compounds Chemical names and Structure

Compound Id	Chemical name	Structure
ZINC03869685	Quercetin	
Pubchem CID 5280805	Rutin	
ZINC19928318	(4-benzylpiperazin-1-yl)-(3-methylbenzofuran-2-yl)methanone	
ZINC24834252	methyl 4-[[4-(naphthalene-2-carbonyl)piperazin-1-yl]methyl]benzoate	
ZINC24601822	[4-[(3-fluoro-4-methoxy-phenyl)methyl]piperazin-1-yl]-(3-methylbenzofuran-2-yl)methanone	
ZINC24601767	[4-[(3,4-dimethoxyphenyl)methyl]piperazin-1-yl]-(3-methylbenzofuran-2-yl)methanone	
ZINC19632922	BEFURALINE	

468

469