Prediction of enzyme inhibition and mode of inhibitory action based on calculation of distances between hydrogen bond donor/acceptor groups of the molecule and docking analysis: An application on the discovery of novel effective PTP1B inhibitors

P. Eleftheriou, A. Petrou, A. Geronikaki, K. Liaras, S. Dirnali & M. Anna

a Department of Medical Laboratory Studies, School of Health and Medical Care, Alexander Technological Educational Institute of Thessaloniki, Greece
b Department of Medicinal Chemistry, School of Medicine, Aristotle University of Thessaloniki, Greece

Published online: 21 Aug 2015.

To cite this article: P. Eleftheriou, A. Petrou, A. Geronikaki, K. Liaras, S. Dirnali & M. Anna (2015): Prediction of enzyme inhibition and mode of inhibitory action based on calculation of distances between hydrogen bond donor/acceptor groups of the molecule and docking analysis: An application on the discovery of novel effective PTP1B inhibitors, SAR and QSAR in Environmental Research, DOI: 10.1080/1062936X.2015.1074939

To link to this article: http://dx.doi.org/10.1080/1062936X.2015.1074939
Prediction of enzyme inhibition and mode of inhibitory action based on calculation of distances between hydrogen bond donor/acceptor groups of the molecule and docking analysis: An application on the discovery of novel effective PTP1B inhibitors

P. Eleftherioua*, A. Petroub, A. Geronikakib, K. Liarasb, S. Dirnali and M. Annaa

aDepartment of Medical Laboratory Studies, School of Health and Medical Care, Alexander Technological Educational Institute of Thessaloniki, Greece; bDepartment of Medicinal Chemistry, School of Medicine, Aristotle University of Thessaloniki, Greece

(Received 28 June 2015; in final form 16 July 2015)

PTP1B is a protein tyrosine phosphatase involved in insulin receptor desensitization. PTP1B inhibition prolongs the activated state of the receptor, practically enhancing the effect of insulin. Thus PTP1B has become a drug target for the treatment of type II diabetes. PTP1b is an enzyme with multiple binding sites for competitive and allosteric inhibitors. Prediction of inhibitory action using docking analysis has limited success in case of enzymes with multiple binding sites, since the selection of the right crystal structure depends on the kind of inhibitor. In the present study, a two-step strategy for the prediction of PTP1b inhibitory action was applied to 12 compounds. Based on the study of known inhibitors, we isolated the structural characteristics required for binding to each binding site. As a first step, 3D-structures of the molecules were produced and their structural parameters were measured and used for prediction of the binding site of the compound. These results were used for the selection of the appropriate crystal structure for docking analysis of each compound, and the final prediction was based on the estimated binding energies. This strategy effectively predicted the activity of all compounds. A linear correlation was found between estimated binding energy and inhibition measured in vitro ($r = -0.894$).

Keywords: PTP1b inhibitors; docking analysis; structural characteristics; active centre; allosteric centre; competitive inhibitors; uncompetitive inhibitors; thiazols; diabetes

1. Introduction

Enzymes represent one of the main classes of drug targets with 124 members, which places them in second position after receptors with 193 members. As a result, 234 drugs which modify enzyme action had already been approved up to 2011 [1]. These drugs belong to several drug categories including not steroidal anti-inflammatory drugs, antihypertensive drugs, drugs with anti-diabetic activity, antiviral drugs, etc. So, the effort to find novel enzyme inhibitors is an established aim of many research groups all over the world.

Resemblance to the substrate is the primary strategy for designing competitive inhibitors which is followed by all researchers when dealing with a totally new drug target. However,
as the number of known inhibitors of an enzyme increases, several other tools, based on the
structure of known inhibitors, can be used to facilitate successful design. Among these is the
incorporation of groups (fragments of molecules) already present in other inhibitors as parts
of the newly designed molecule. The systematic application of fragment-based design
includes the preparation of a library of molecule fragments of known inhibitors, selection of
the fragments with the highest positive influence on inhibition and use of them for the design
of novel inhibitors [2]. This solves the problem of the chemical groups that can be used, but
not the problem of orientation of these groups within the molecule.

In addition, several methods based on the similarity with structures from a database of
compounds with known activity are used. However, in these methods relatively large sub-
structures are used as predefined moieties for the comparison, which reduces the chance of
finding compounds with novel scaffold characteristics [3,4]. Moreover, the value of results is
limited in the case of relatively new drug targets because of the small number of molecules
with the desired activity in the training set. Although with restrictions, the value of such
programs cannot be overlooked. The prediction program PASS (Prediction of Activity Spectra
for Substances), which also predicts pharmacological effects and biochemical mechanisms on
the basis of the structural formula of the substances, can be used for the estimation of the
activity of designed compounds. PASS has been successfully used to identify agents with
several pharmaceutical activities, including multi-target agents [5,6].

Docking analysis may be used as a prediction tool if crystal structures of the enzyme
target are available [7,8], but it may not give reliable results when more than one probable
binding site exists at the enzyme molecule. It is well known that enzymes are flexible
molecules and transformation of their structure may take place when interacting with their
substrate or inhibitor. So, selection of the right crystal structure may be crucial. The use of
different crystal structures for docking analysis of the same compound usually produces
different estimated binding energy. Comparable results for a series of compounds can
normally be obtained by using the same crystal structure for all the compounds. However, the
use of multiple available structures and calculation of an average estimated binding energy
has also been proposed [8–10]. A more complicated problem is faced when multiple sites of
probable binding of a molecule at the same protein exist, as in enzymes with competitive and
allosteric inhibition.

PTP1b is an enzyme with multiple binding sites. Competitive inhibitors of the
enzyme may occupy only the main area around the active centre or may simultaneously be
connected to other areas around the active centre, while an allosteric binding site also exists
[11–14].

In the present study, we attempted to use available tools such as 3D structure designing
programs, crystal structures and docking analysis for the prediction of the ability of certain
compounds to act as enzyme inhibitors. The strategy emphasizes the selection of appropriate
crystal structures for the docking analyses. The study of known inhibitors which bind to each
binding site and the surrounding amino acids was used to elucidate the structural characteris-
tics required for binding to each specific binding area. Based on these criteria, the potential
binding site of each compound was predicted. According to the predicted binding site, an
appropriate crystal structure of the enzyme, derived from complexes with inhibitors which
bind to the same binding site, was selected and used for docking analysis. Prediction of
activity was finally done, based on the estimated binding energies of the compounds. The
prediction strategy was applied to 12 compounds, and the prediction results were evaluated
by the in vitro measurement of PTP1b inhibitory action.
PTP1B is a protein tyrosine phosphatase involved in insulin receptor desensitization. PTP1B inhibition, resulting in prolonged maintenance of the activated state of the receptor, practically enhances the effect of insulin. Thus, PTP1B has become a drug target for the treatment of type II diabetes [15,16]. Although a great number of inhibitors have been developed, among which are inhibitors binding to the catalytic site and allosteric inhibitors, the development of novel, highly effective and selective compounds remains a challenge.

2. Methods

2.1 Study, comparison and selection of the 3D structure of PTP1b

The 3D structure of the PTP1b enzyme complexes were obtained from the National Centre of Biotechnology Information, US National Library of Medicine, Protein Data Bank (NCBI PDB), available through the PubMed/structure option (NCBI/Computational Biology/Branch/Structure Group/Macromolecular Structures/3D Macromolecular structures). Yellow/light grey colour designation of characteristic amino acid residues was performed through the options of this database (Figures 1(a), 2) [17]. The RCSB (Research Collaboratory for Structural Bioinformatics) PDB was used for protein structure alignment (Figure 1(b)), measurement of distances between atoms, download of information concerning ligand–protein interactions (existence of hydrogen bonds, aromatic and hydrophobic interactions) and Ki/IC50 values of the inhibitors [18,19].

2.2 Design of 3D structure of the studied compounds

For the design of the 3D structure of the compounds, CORINA 3.6 software was used [20]. CORINA software uses the 2D structure of small molecules to produce the lowest energy 3D structure of the free molecule, and enables the user to measure the distances between selected atoms or the angles formed in the 3D structure. However, this enables only a draft estimation of probable distances and angles, since the molecule may obtain a different favourable structure when bound to the enzyme molecule. In spite of these restrictions, we considered this draft estimation, derived from the 3D structure of the compounds, as an easy tool for a first-step evaluation and decided to use it in a two-step prediction process which involves docking analysis as the second and more crucial step of a prediction process.

2.3 Docking analysis

Docking analysis was performed using the AutoDock 4.2 program. Reference protein coordinates for the docking studies were taken from PDB. The protein–ligand complexes of PTP1b which were used were the complex coded as 2VEW in which inhibitor was bound at the A1 and A2 areas of the binding site of PTP1b, the complex with code number 1T4J in which the inhibitor was bound at the allosteric site of the enzyme, and the complex with the code number 2F71 in which the inhibitor was bound at the areas A1 and B.

Docking calculations were carried out using software AutoDock 4.2. This is one of the most suitable methods for performing molecular docking of ligands to their macromolecular receptors, as well as for discriminating potential inhibitors. The ligand molecule is in an arbitrary conformation, orientation and position, and this molecular docking program [21] finds favourable poses in a protein-binding site using Lamarckian genetic algorithms implemented therein to search for the best conformers. The free energy of binding (ΔG) of the selected PTP1b structure in complex with the inhibitors was generated using this molecular docking program.
For PTP1b preparation, polar hydrogens were added. For ligand preparation, Gasteiger partial charges were added, non-polar hydrogen atoms merged, and rotatable bonds defined. The grid maps of docking studies were computed using the AutoGrid program. The grid must surround the region of interest in the macromolecule. The centre of the grid docking box was determined by coordinates X, Y and Z of sulphur of the Cys215 residue for docking at the active centre; for 1T4J these were 44.89, 16.841 and 7.923, for 2VEW these were 15.7872, 18.916 and 78.1173, and for 2f71 these were 40.7353, 10.7451 and 15.3277. The centre of the grid docking box was determined by coordinates X, Y and Z of carbon 2 of the Phe280 residue for docking at the allosteric centre; for 1T4J these were 56.7801, 31.1552 and 23.1272, for 2VEW these were 44.921, 1.84 and 84.563, and for 2f71 these were 29.634, 26.77 and –10.964. The three-dimensional structures of the aforementioned compounds were constructed using Chem 3D ultra 14.0 software (Chemical Structure Drawing Standard; Perkin Elmer Informatics, Waltham, MA, USA). The compounds were subjected to flexible docking using the pre-computed grid files. All parameters used in docking were default. A preliminary blind docking study was performed in order to discriminate the preferential binding sites of the ligand to the receptor. The translation, quaternion and torsions steps were taken from default values in AutoDock. The Lamarckian genetic algorithm and the pseudo-Solis and Wets methods were applied for minimization using default parameters. The number of docking runs was 100. After docking, the 100 solutions were clustered into groups with RMS lower than 1.0 E. The clusters were ranked by the lowest energy representative of each cluster. In order to describe the ligand–binding pocket interactions, the top-ranked binding mode was found by AutoDock in complex with the binding pocket of enzyme.

The estimated free binding energy (∆G), the molecular mechanical force fields 94 (MMFF94) energy, the binding affinity score, hydrogen bond acceptor, hydrogen bond donor interactions, hydrophobic interactions and aromatic interactions were among the parameters calculated using the AutoDock program. In most cases, the structure with the lowest binding energy (∆G) exhibited the best values of all parameters, including the highest number of donor acceptor interactions, etc. However, when a different ranking of probable structures was obtained according to different parameters, the lowest binding energy structure was selected among the structures of the higher number of hydrogen donor/acceptor interactions or overall interactions.

Apart from AutoDock 4.2, AutoDock Vina was also used for estimation of ligand–enzyme binding capacity. The two programs produced similar results concerning the binding mode and binding energy ranking of the compounds.

2.4 Biological evaluation of the PTP1b inhibitory action

PTP1-B inhibition activity was tested using human recombinant PTP1-B (Calbiochem). Incubation was carried out at 25°C for 60 min in a 100 μl reaction mixture containing 40 mM Tris-HCl pH 7.5, 100 mM NaCl, 1mM DTT, 1mM EDTA, 0.5 mg/ml BSA, 2.5 mM of p-nitro-phenyl phosphate and 4.5 U of the enzyme. Volumes of 10 μl of each compound, dissolved in DMSO were added. Compounds were incubated with the enzyme mixture for 15 min at room temperature before addition of the substrate. The reaction was stopped by the addition of equal volume of 2.5 N NaOH. Enzyme activity was estimated by measuring the absorbance of p-nitrophenol produced after dephosphorylation of the substrate at 405 nm with appropriate corrections for absorbance of the compounds [2,22]. Different substrate concentrations were used in order to elucidate the mode of inhibitory action, and different compound concentrations were used in order to calculate the IC50 value of the active compounds.
However, whenever percentage inhibition is noted, substrate concentration was 2.5 mM and compound concentration was 50 μM. The experiments were performed in triplicate.

2.5 Estimation of linear correlation

For the estimation of correlation between parameters, a linear regression fit using the method of reduction to major axis was applied.

3. Results and discussion

3.1 Structural characteristics of PTP1b inhibitors according to their binding site

Several kinds of competitive inhibitors of PTP1b have been found. All competitive inhibitors are bound at the active site, area A1 (Figure 1), interacting with the amino acids surrounding Cys215 which is the crucial amino acid for catalytic activity [15,23]. However, most potent inhibitors also bind to a second site near the active centre, such as area A2 around amino acids Tyr46, Arg47 and Asp48. Arg47 and Asp48 are the amino acids mostly involved in hydrogen bond interactions with this site. A third kind of inhibitor can simultaneously bind to site A1 (or sites A1/A2) and to area B near Arg24 [24], while a fourth kind of inhibitor can be simultaneously connected to sites A1, A2 and area C, interacting with Lys41 [13]. Occupation of sites B and C is considered essential for the development of specific inhibitors since TCPTP, the protein tyrosine phosphatase with the greatest resemblance with PTP1b, differentiates in amino acids of these areas, such as Lys41(→Arg) of area C and Ala27(→Ser) of area B [12,15].

Most importantly, apart from the inhibitors which interact with the catalytic site, there are inhibitors which interact with a second allosteric site of the molecule around the amino acids K197, N193, A189 and F280 [14].

Study of the binding mode of several inhibitors and the amino acids surrounding each binding site reveals that a hydrophobic (Hph1) group surrounded by two hydrogen donor/acceptor groups (Hd/a1 and Hd/a2) at a distance around 9.3 Å may favour binding to the A1 and A2 sites [11]. A second hydrophilic group (Hph2) may be present. If this second hydrophilic group exists, connected to the Hd/a2 group, it has to be able to be placed at the right position within enzyme molecule, which is accompanied by specific requirements concerning the angle between Hd/a1-Hd/a2 and Hd/a2-Hph2 bonds and specific distances between Hd/a1-Hph2 groups or Hph1-Hph2 groups (Figure 2(a)).

Molecules capable of interaction also with the area of site B (B1 or B2) need a third Hd/a group at a distance of 15–21 Å from Hd/a1 and 10–14Å from Hd/a2 (Figure 2(b)) [11]. A third Hd/a group at a distance of about 19 Å from Hd/a1 and 10–12Å from Hd/a2 is needed for molecules which will occupy site C as well (Figure 2(c)) [11]. The existence of at least two Hd/a groups is common even in molecules which interact only with amino acids of area A1.

In the case of molecules which interact with the allosteric site, Hd/a groups also seem to participate in the interaction and distances around 7.6 Å seem to be favourable, although a small number of such inhibitors have been found (Figure 2(d)) [11].

Among the PTP1b crystal structures which are available in the PDB, the structure 1A5Y concerns the cysteinyl-phosphate intermediate of the enzyme and does not involve co-crystalization with a substrate or an inhibitor [25]. The remaining structures are derived from PTP1b complexes with inhibitors and can be separated into categories according to the binding mode of the inhibitor (Table 1, Figure 3).
Table 1. Crystal structures of PTP1b complexes available in PDB. Binding area and \( K_i/IC_{50} \) values of the inhibitors [14,24,26–35].

<table>
<thead>
<tr>
<th>PTP1b structure</th>
<th>Inhibitor</th>
<th>Inhibitor binding sites</th>
<th>( K_i ) (µM)</th>
<th>( IC_{50} ) (µM)</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>2AZR</td>
<td>Thieno[2,3-b]pyridine-2-carboxylic acid</td>
<td>A1</td>
<td>230</td>
<td></td>
<td>26</td>
</tr>
<tr>
<td>2F6W</td>
<td>(2-methyl-5-phenylpyrazol-3-yl)sulfamic acid</td>
<td>A1</td>
<td>&gt;500</td>
<td></td>
<td>27</td>
</tr>
<tr>
<td>2F6Y</td>
<td>3(R)-methyl carbamoyl-7-sulfoamino-3,4-dihydro-1H-isooquinoline-2-carboxylic acid tetr-butyl ester</td>
<td>A1, Y46 (Hd/a – Ar)</td>
<td>24</td>
<td>134.8</td>
<td>27</td>
</tr>
<tr>
<td>1ONZ</td>
<td>2-[(7-hydroxy-naphthalen-1-yl)-oxalyl-amino]-benzoic acid</td>
<td>A1, Y46 (Hd/a – Ar)</td>
<td>21</td>
<td></td>
<td>28</td>
</tr>
<tr>
<td>1KAK</td>
<td>2,7-di(di-fluoro-phosphono-methyl)-naphthalene</td>
<td>A1, A2 (Hph – Hd/a)</td>
<td></td>
<td></td>
<td>29</td>
</tr>
<tr>
<td>1KAV</td>
<td>and 1,4-di[4(di-fluoro-phosphono-methyl]-phenyl]-butan</td>
<td>A1, Y46 (Hd/a – Ar)</td>
<td>1.5</td>
<td>4.4</td>
<td>26</td>
</tr>
<tr>
<td>2F6T</td>
<td><em>Bicyclic and tricyclic thiophenes</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2F6V</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2B07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1JF7</td>
<td>2-(carboxymethoxy)-5-[(2S)-2-[(2-methylpropan-2-yloxy)carbonylamino]-3-phenylpropanoyl]amino]-3-oxo-3-(pentylamino) propylbenzoic acid</td>
<td>A1, A2 (Hd/a – Hd/a)</td>
<td>0.37</td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>1Q6J</td>
<td>(Benzotriazol-1-yl)-di-phenyl-[4(di-fluoro-phosphonomethyl]-phenyl-methane derivatives (1Q6J, 1Q6M, 1Q6P) and (benzotriazol-1-yl)-(1,3-benzothiazol-2-yl)-</td>
<td>A1, A2 (Hd/a – Hd/a)</td>
<td>0.016</td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>1Q6M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1Q6P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1Q6N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2FJN</td>
<td>[[4-[(E,2R)-2-(benzotriazol-1-yl)-2-(4-methoxy carbonyl) phenyl]-5-phenylpent-4-enyl]phenyl]-difluoromethyl]phosphonic acid</td>
<td>A1, A2 (Hd/a – Hd/a)</td>
<td>0.039</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>2H4G</td>
<td>4-bromo-3-(carboxymethoxy)-5-(4-hydroxyphenyl) thiophene-2-carboxylic acid</td>
<td>A1, A2 (Hd/a – Hd/a)</td>
<td>0.3</td>
<td></td>
<td>32</td>
</tr>
<tr>
<td>2VEW</td>
<td>3-Fluoro-N-[((1s)-1-[(4-[[2-Fluorophenyl]methyl]imidazol-2-yl]-2-[[5s]-1,1,3-Trioxo-1,2-Thiazolidin-5-yl]phenyl[ethyl]benzenesulfonamide</td>
<td>A1, A2 (Hd/a – Hd/a)</td>
<td>0.064</td>
<td></td>
<td>33</td>
</tr>
<tr>
<td>1G7G</td>
<td>Sulfo tyrosyl tripeptide analog</td>
<td>A1, A2 (Hd/a – Hd/a)</td>
<td>0.25</td>
<td></td>
<td>34</td>
</tr>
<tr>
<td>2F6Z</td>
<td>{3(S)-methylcarbamoyl-2-[3-(3-sulfoamino-phenyl)propionyl]-1,2,3,4-tetrahydro isoquinolin-7-yl]-sulfamic acid</td>
<td>A1, B (Hd/a – Hd/a)</td>
<td>4.8</td>
<td></td>
<td>27</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>PTP1b structure</th>
<th>Inhibitor</th>
<th>Inhibitor binding sites</th>
<th>$K_i$ (µM)</th>
<th>$IC_{50}$ (µM)</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>2F71</td>
<td>3-[3-(3(S)-methylcarbamoyl-7-sulfoamino-3,4,5,6-tetrahydro-1H-pyridin-2-yl)-3-oxo-propyl]-benzoic acid</td>
<td>A1, A2, B (Hd/a–Ar–Hd/a)</td>
<td>2.5</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>2F70</td>
<td>UN6; 3-[3-oxo-3-[3-(sulfoamino)phenyl]methylamino]propyl]-sulfamic acid</td>
<td>A1, A2, B (Hd/a–Ar–Hd/a)</td>
<td>33.5</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>1Q6S</td>
<td>[4-(8-phosphonoquinolin-6-yl)-phenyl]-(Benzotriazol-1-yl)-(4[di-fluoro-phosphono-methyl]-phenyl)-phenyl-methane derivatives</td>
<td>A1, A2, B (Hd/a–Hd/a–Hd/a)</td>
<td>0.012</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>1Q6T</td>
<td>1Q6T</td>
<td>A1, A2, B (Hd/a–Hd/a–Hd/a)</td>
<td>0.005</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>1PXH</td>
<td>3-[2-[4-[difluoro(phosphono)methyl]phenyl] acetyl]amino]-4-[[2S]-3-[4-[di-fluoro(phosphono)methyl]phenyl]-1-oxo-1-{(pentylamino)propan-2-yl]amino]-4-oxobutanoic acid</td>
<td>A1, A2, C (Hd/a–Hd/a–Hd/a)</td>
<td>0.0018</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>1T4J</td>
<td>3-(3,5-dibromo-4-hydroxybenzoyl)-2-ethyl-N-[4-(1,3-thiazol-2-ylsulfamoyl)phenyl]-1-benzofuran-6-sulfonamide</td>
<td>Allosteric 2Hd/a, 3Ar</td>
<td>8</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>1T49</td>
<td>3-(3,5-dibromo-4-hydroxybenzoyl)-2-ethyl-N-(4-sulfamoylphenyl)-1-benzofuran-6-sulfonamide</td>
<td>Allosteric 2Hd/a, 2Ar</td>
<td>22</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>1T48</td>
<td>3-(3,5-dibromo-4-hydroxybenzoyl)-2-ethyl-N,N-dimethyl-1-benzofuran-6-sulfonamide</td>
<td>Allosteric 1Hd/a, 2Ar</td>
<td>350</td>
<td>14</td>
<td></td>
</tr>
</tbody>
</table>
The structures 2AZR and 2F6W come from complexes with small inhibitors which occupy only the A1 binding site. These inhibitors have very low inhibitory action, with $K_i$ values over 200 µM.

Hydrogen bond formation with residues of the A1 site and aromatic interaction with Tyr46, neighbouring Arg47 and Asp48 of the A2 site are observed in structures 2F6Y, 1KAV, 1ONZ, 2F6T, 2F6V. These compounds practically bind at the A1 site, but exhibit a slight interaction with an amino acid neighbouring the strong core of the A2 site. The opposite phenomenon with hydrophobic interactions with amino acids of the A1 site and hydrogen bond formation with amino acids of the A2 site is observed in the case of the 1KAK complex. In the cases mentioned above, the inhibitors exhibit a moderate inhibitory action, with $K_i$ or IC50 values at the micromolar range (Table 1).

In the structures 2B07, 1JF7, 1Q6J, 1Q6M, 1Q6N, 1Q6P, 2FJN, 2H4G, 2VEW and 1G7G the inhibitor occupies areas of both A1 and A2 sites, strongly interacting with these sites via hydrogen bond formation. Aromatic and hydrophobic interactions with amino acids of both sites enforce complex stabilization, resulting in more potent inhibitors with $K_i$ or IC50 values at the sub-micromolar range, with the exception of the 1JF7 inhibitor with $K_i = 2$ µM.

The structure 2F6Z represents a complex with an inhibitor interacting with sites A1 and B. The inhibitor occupies simultaneously the three sites A1, A2 and B in structures 2F71, 2F70, 1Q6S and 1Q6T. However, aromatic interactions are involved between inhibitor and A2 site in the first two structures, while hydrogen bond formation with all three sites is present only in case of the two last structures (1Q6S and 1Q6T) which correspond to the most potent inhibitors (IC50 values at the nanomolar range).

In the structure 1PXH the inhibitor occupies simultaneously the sites A1, A2 and C. Interaction of Hd/a groups of the molecule with residues of all three sites are observed.

In the case of structures 1T4J, 1T49 and 1T48, the inhibitors bind at the allosteric site. The number of hydrogen bonds and the overall number of interactions are in accordance with the IC50 values of the inhibitors. Two Hd/a interactions are observed in case of the most active inhibitors, 1T4J and 1T49 (IC50: 8 µM and 22 µM, respectively), while one Hd/a interaction and two aromatic interactions are observed in case of the practically inactive inhibitor of complex 1T48 (IC50 = 350 µM).

### 3.2 Structural characteristics of the studied compounds; prediction of the binding site of the compounds based on structural characteristics

The studied compounds were thiazolyl derivatives belonging to two subgroups, Scheme 1. The 3D structure of the compounds was designed using CORINA 3.6 software and structural parameters were measured. As a first step, we measured the number of Hd/a clusters in the molecules and the distances between these groups in order to estimate the probability to bind at sites A1, A2, B and C around the active centre. Long distances between Hd/a groups, which would enable simultaneous binding to sites A1, A2 and B or C, do not exist in these molecules. However, the measured distances were close to that enabling binding at the A1 and A2 sites in some compounds, or close to that favouring binding at the allosteric site of the enzyme. The rank of the compounds according to the presence of two Hd/a groups and the distance between these groups is presented in Table 2.

As shown at Table 2, four of the 12 compounds, A3, A7, Et-m and Et-p, had Hd/a1–Hd/a2 distances higher than 8.3 Å and closer to the favourable distance of 9.3 Å which seems to be a requirement for Hd/a interaction with both A1 and A2 sites. So, these compounds have
increased probability to bind at the active site and are predicted to act as A1, A2-binding competitive inhibitors. Six of the compounds, A1, A2, A5, A6, A9 and A8, had Hd/a1–Hd/a2 distances lower than 8.3 Å and closer to the distance of 7.6 Å, which may favour binding to the allosteric site, and were predicted to preferentially act as allosteric inhibitors.

The other two compounds, A4 and A10, bearing only one Hd/a cluster, are not expected to have a high probability for stable binding to the active site of the enzyme. Apart from the

![Scheme 1. Structure of the studied compounds.](image)

Table 2. Distances between Hd/a1 and Hd/a2 groups of the studied compounds.

<table>
<thead>
<tr>
<th></th>
<th>Hd/a1-Hd/a2 (Å)</th>
<th>x-8.3 (Å) x: Hd/a1-Hd/a2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A7</td>
<td>9.7</td>
<td>1.4</td>
</tr>
<tr>
<td>Et-p</td>
<td>9.2</td>
<td>0.9</td>
</tr>
<tr>
<td>Et-m</td>
<td>8.7</td>
<td>0.1</td>
</tr>
<tr>
<td>A3</td>
<td>8.4</td>
<td>−0.8</td>
</tr>
<tr>
<td>A5</td>
<td>7.5</td>
<td>−0.8</td>
</tr>
<tr>
<td>A6</td>
<td>7.5</td>
<td>−0.8</td>
</tr>
<tr>
<td>A8</td>
<td>7.5</td>
<td>−0.8</td>
</tr>
<tr>
<td>A9</td>
<td>7.5</td>
<td>−0.8</td>
</tr>
<tr>
<td>A2</td>
<td>7.8</td>
<td>−0.5</td>
</tr>
<tr>
<td>A1</td>
<td>6.7</td>
<td>−1.6</td>
</tr>
<tr>
<td>A4</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>A10</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>
absence of a second Hd/a group at the right position, the hydrophobic groups surrounding the
Hd/a cluster are at a short distance from the Hd/a cluster, as in case of the practically inactive
molecules of the structures 2AZR and 2F6W ($K_i = 230$ µM and >500 µM, respectively).

More precisely, in compound A3, bearing a NO$_2$ group at the 4’ position of the phenyl
group, a distance of 8.4 Å was measured between O of the NO$_2$ group and N of the
benzothiazolyl moiety.

In compound A7, bearing two F atoms, one at each one of the aromatic rings of the mole-
cule, a favourable distance of 9.7 Å was measured between 6-F of the benzothiazolyl moiety
and O atom of the –SO$_2$- group.

In compounds A6, A8 and A9, bearing only one F atom at position 4’ of the phenyl ring,
the distance of 7.5 Å between F and N of the thiazolyl moiety was the longer distance mea-
sured between Hd/a groups of the molecule. The same distance was measured between O of
the OCH$_3$ substituent of the phenyl ring and N atom of the benzothiazolyl moiety, in case of
compound A5.

Compounds A1 and A2 bear a Cl atom as a substituent at one or both the aromatic rings
of the molecule. Cl is not among the most electronegative atoms, such as F, O or N which
participate in hydrogen bond formation. So, these compounds could be considered as having
only one Hd/a cluster. However, Cl may also participate in weaker hydrogen bonds with H
atoms of classic hydrogen donors. The distance between 4’-Cl of the phenyl ring and N of
the benzothiazolyl moiety was 7.8 Å in the case of compound A2 and 6.5 Å in the case of
A1. The distance of the 4-Cl atom of the benzothiazolyl moiety and O of the –SO$_2$- group
was the longer distance measured in compound A1 (6.7 Å).

In compound Et-p, a distance of 9.2 Å was measured between O of the NO$_2$ group and O
of the keto group and was the closer measurement to the favourable distance of 9.3 Å. The
same distance was 8.7 Å in case of compound Et-m.

Figure 1. (a) Location of the A1, A2, B and C binding sites and of the allosteric site of PTP1b
enzyme, indicated on the enzyme complex with an A1 site binding inhibitor (PDB code: 1ONZ) [28].
(b) 3D structure alignment of PTP1b enzyme complex with A1 site binding inhibitor 1ONZ (dark grey)
with the PTP1b enzyme complex with A1, A2 sites binding inhibitor 2VEW (light grey) [17,18,36].
The position of Phe182 in both structures is indicated in coloured cycles.
3.3 Prediction of inhibitory action based on docking results

The first goal in docking analysis is the selection of the right crystal structure. This is essential and may strongly affect the results, especially in case of enzymes with multiple binding sites such as PTP1b. This mainly comes from the property of the enzymes to slightly change their conformation following interaction with other molecules such as substrates or inhibitors. An example of the difference between crystal structures is shown in Figure 1(b). The figure shows the 3D structure alignment of PTP1b enzyme complex with an A1-site binding inhibitor, 1ONZ (dark grey), with the PTP1b enzyme complex with an A1, A2-sites binding inhibitor, 2VEW (light grey). The image clearly indicates the difference of the location of the residue Phe182, in the area of A1 binding site, which belongs to the amino acids surrounding the active centre and is often involved in hydrogen bond formation and enzyme complex stabilization (Figure 1(b)).

According to the results derived from the structural characteristics (Table 2), four of the compounds, A3, A7, Et-m and Et-p fulfil the requirements for binding at the active centre of the enzyme, interacting with areas A1 and A2 via Hd/a interactions. Aromatic and hydrophobic interactions may also enforce complex stabilization. Among the available structures in the PDB database, one corresponds to the enzyme without inhibitor and 11 are derived from crystal structures of the enzyme with inhibitors bound to the allosteric site or inhibitors bound to

Figure 2. Distances between the main hydrogen acceptor/donor groups of inhibitors binding at different sites of PTP1b. (a) Binding at the sites A1 and A2, structure 2VEW. (b) Binding at the sites A1, A2 and B, structure 1Q6T [24]. (c) Binding at the sites A1, A2 and C, structure 1PXH [35]. (d) Binding at the allosteric site, structure 1T4J [14]. The numbers in white colour represent distances in Å.
solely A1 site or simultaneously occupying the A1, A2 and B or A1, A2 and C sites. These structures were excluded. Of the remaining structures, six are derived from complexes exhibiting hydrogen bond interactions with the A1 site and aromatic interaction with Tyr46, neighbouring Arg47 and Asp48 of the inner, main core amino acids of the A2 site, while one

Figure 3. 2D structures of known PTP1b inhibitors.
involves Hd/a interactions only with the A2 site. The other seven structures are derived from complexes with Hd/a interactions with both A1 and A2 sites, along with aromatic/hydrophobic interactions with both sites (2B07, 1Q6M, 1Q6N, 1JF7, 2FJN, 2H4G, 2VEW).

This last group of structures have the conformation enabling full occupation of the A1 and A2 sites by the inhibitors and were considered as the most suitable for the group of compounds, A3, A7, Et-m and Et-p, which fulfill the requirements for Hd/a interaction with the A1, A2 sites. The structure 2VEW was randomly selected among the seven suitable structures for docking analysis of these compounds.

For the compounds with Hd/a1–Hd/a2 distances lower than 8.3 Å, which have a higher probability to bind at the allosteric site (A1, A2, A5, A6, A8 and A9), the crystal structure 1T4J was randomly selected from among the three available structures (1T4J, 1T48 and 1T49) which are derived from a complex of the enzyme with an allosteric inhibitor. The same structure was used for docking analysis of the compounds A4 and A10 with no second Hd/a group, which are expected to have very low inhibitory action. The binding energies of the compounds are shown in Table 3.

The binding energies at the active site of 2VEW varied between −6.5 and −7.1 kcal/mol, for compounds A3, A7, Et-m and Et-p, for which a stable binding at the A1, A2 areas of the active site had been predicted.

Calculation of the binding energies of the compounds A1, A2, A5, A6, A8, A9, A4 and A10 at the allosteric site of 1T4J crystal structure revealed that estimated binding energy was higher than −5.7 kcal/mol for nearly all compounds, with the exception of compounds A1, A2 and A9, with binding energies −7.6 kcal/mol, −8.4 kcal/mol, and −6.1 kcal/mol, respectively. According to the binding energies, compounds A1 and A2 are expected to be active inhibitors, while compound A10 with a binding energy close to −6.0 is expected to have lower, if not no inhibitory action. In general, compounds with a binding energy higher than −6.0 kcal/mol are not expected to form stable complexes, and consequently exhibit minor or no inhibitory action.

<table>
<thead>
<tr>
<th>Compounds</th>
<th>2VEW (structure complexed with A1, A2 binding inhibitor)</th>
<th>1T4J (structure complexed with allosteric inhibitor)</th>
<th>2F71 (complexed at A1, B sites, Hd/a–Hd/a distance 17–21 Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimated binding energy (kcal/mol)</td>
<td>Estimated binding energy (kcal/mol)</td>
<td>Estimated binding energy (kcal/mol)</td>
</tr>
<tr>
<td>A7</td>
<td>−6.5</td>
<td>−6.9</td>
<td>−6.7</td>
</tr>
<tr>
<td>Et-p</td>
<td>−6.6</td>
<td>−6.6</td>
<td>−7.2</td>
</tr>
<tr>
<td>Et-m</td>
<td>−7.0</td>
<td>−6.4</td>
<td>−7.5</td>
</tr>
<tr>
<td>A3</td>
<td>−7.1</td>
<td>−8.0</td>
<td>−6.7</td>
</tr>
<tr>
<td>A5</td>
<td>−7.2</td>
<td>−5.1</td>
<td>−7.5</td>
</tr>
<tr>
<td>A2</td>
<td>−7.2</td>
<td>−6.8</td>
<td>−6.1</td>
</tr>
<tr>
<td>A9</td>
<td>−7.5</td>
<td>−5.5</td>
<td>−5.3</td>
</tr>
<tr>
<td>A8</td>
<td>−8.0</td>
<td>−4.2</td>
<td>−5.7</td>
</tr>
<tr>
<td>A6</td>
<td>−8.0</td>
<td>−4.9</td>
<td>−5.3</td>
</tr>
<tr>
<td>A1</td>
<td>−6.1</td>
<td>−7.0</td>
<td>−5.8</td>
</tr>
<tr>
<td>A4</td>
<td>−6.2</td>
<td>−4.4</td>
<td>−5.6</td>
</tr>
<tr>
<td>A10</td>
<td>−6.1</td>
<td>−5.0</td>
<td>−5.6</td>
</tr>
</tbody>
</table>
Binding energies of the same compound when docked either at the active or at the allosteric site of the same crystal structure were calculated for many of the compounds. Interestingly, docking analysis revealed lower energy for binding of the compounds at the site predicted as most probable according to structural characteristics, with the exception of the results derived from docking at the 2VEW structure of the compound A1, which was expected to bind at the allosteric site and for which 1T4J was selected as the appropriate crystal structure for docking (2VEW: active centre: −6.1, allosteric: −5.8). Docking of the same compound at the 1T4J structure, considered as the appropriate for an allosteric inhibitor, revealed lower energy for binding at the allosteric site (A1-1T4J: active centre: −7.0 kcal/mol, allosteric site: −7.6 kcal/mol), placing it among the allosteric inhibitors predicted as active.

### 3.4 In vitro evaluation of the biological activity of the compounds and estimation of prediction accuracy

The PTP1b inhibitory action of the compounds was evaluated using human recombinant PTP1b. p-nitrophenylphosphate at concentrations 2.5 μM and 5.0 μM was used as a substrate and the incubation was performed in the presence or absence of 50 μM of the compounds. Inhibition, expressed as a percentage of the total activity, is shown in Table 4. IC₅₀ values of the active compounds were also calculated and their values are presented in the table. The mode of inhibitory action was estimated by measuring the activity of the enzyme at two different substrate concentrations in the presence and absence of the inhibitor (Table 4). The Kᵢ values of the compounds were calculated using the equations proposed by Cerl et al. [37] for each specific mode of inhibitory action. Since the substrate concentration used for the evaluation of inhibition % and for IC₅₀ calculation was equal to the Kₘ of the enzyme, Kᵢ values for both competitive and uncompetitive inhibitors are equal to half the IC₅₀ value of the compound.

All compounds predicted to be active showed more than 45% inhibition, while all other compounds exhibited zero (compounds A4, A5, A6, A8 and A10) or near zero inhibition (inhibition 6.5% for compound A9).

<table>
<thead>
<tr>
<th>Prediction results</th>
<th>Biological results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated binding energy (kcal/mol)</td>
<td>Probable connection</td>
</tr>
<tr>
<td>2VEW (active centre)</td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>−7.1</td>
</tr>
<tr>
<td>Et-m</td>
<td>−7.0</td>
</tr>
<tr>
<td>Et-p</td>
<td>−6.6</td>
</tr>
<tr>
<td>A7</td>
<td>−6.5</td>
</tr>
<tr>
<td>1T4J (allosteric)</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>−8.4</td>
</tr>
<tr>
<td>A1</td>
<td>−7.6</td>
</tr>
<tr>
<td>A4</td>
<td>−5.6</td>
</tr>
<tr>
<td>A5</td>
<td>−5.3</td>
</tr>
<tr>
<td>A6</td>
<td>−5.3</td>
</tr>
<tr>
<td>A8</td>
<td>−5.7</td>
</tr>
<tr>
<td>A9</td>
<td>−6.1</td>
</tr>
<tr>
<td>A10</td>
<td>−5.6</td>
</tr>
</tbody>
</table>

Table 4. Estimated binding energy and prediction of binding site, inhibition activity and mode of inhibitory action of the active compounds.
Interestingly, the compounds which were predicted to bind at the active centre were found to act as competitive inhibitors, while those predicted to bind at the allosteric site were found to act as uncompetitive inhibitors. This kind of inhibitor binds to the enzyme after the substrate is bound at the active site, and they most probably bind to other sites than the active site of the enzyme (allosteric sites).

As shown in Figure 4, the inhibition percentage of the active compounds seems to correlate well with estimated binding energy. The compounds with lower estimated binding energy exhibited higher inhibition, and all compounds with zero to near zero inhibition had estimated binding energy higher than \(-6.0\) kcal/mol (with the exception of A9 which exhibited \(-6.1\) kcal/mol).

Investigation of the existence of a linear correlation between percentage of inhibition and estimated binding energy by the application of a simple linear regression program revealed that the values can follow the equation \(y = -\left(75+19.8x\right)\), where \(y\) is inhibition\% and \(x\) is estimated binding energy (kcal/mol), with a correlation coefficient \(r = -0.894\) (Figure 5(a)).

According to the results presented in Table 4 and Figures 4 and 5(a), the whole prediction strategy used seems to effectively predict the inhibitory action of the compounds.

In a further step, we investigated the effectiveness of prediction using solely docking results derived from docking of all compounds only at one binding site of the same crystal structure. The structure was randomly selected from the group of crystal structures considered as inappropriate. The crystal structure selected was 2F71, in which PTP1b forms a complex
with an inhibitor binding to sites A1 and B, and docking was performed at the active site of the enzyme which is more usually selected as the ligand target. None of the compounds has the ability to simultaneously interact with these two sites (requirement of a distance of 17–21 Å between two Hd/a groups). The estimated binding energies are shown in Table 3. As shown in Figure 6, a correlation between inhibition percentage and estimated binding energy does not seem to exist in this case, and a binding energy higher than −6.0 kcal/mol was calculated for the active compound A1 (inhibition: 68.3%, est. binding energy: −5.8 kcal/mol), placing it in the group of the compounds predicted to be inactive. No equation was produced by linear regression analysis of these results. As shown in Figure 5(b), no linear correlation exists (correlation coefficient $r = -0.188$).

Conclusively, while docking analysis following structural characteristics-guided selection of an appropriate crystal structure was proved to be effective for prediction of PTP1b inhibitory action, docking using a randomly selected structure failed to effectively predict inhibition activity of the compounds.

Comparison of the biological activity of the compounds with known inhibitors places them among the moderate inhibitors. Although known allosteric inhibitors have IC$_{50}$ values at the micromolar rage (8–350 μM, Table 1) comparable with the studied compounds A2 and A1 (5.2 and 22.9 μM, respectively), the IC$_{50}$ values of inhibitors interacting with areas A1 and A2 of the active site vary from the micromolar to the nanomolar range, with the most potent compounds exhibiting values around 2 nM or even less. However, more potent inhibitors have been mentioned among compounds occupying simultaneously the areas A1, A2 and C of the enzyme (Table 1).

The comparison of IC$_{50}$ values, although always attempted by researchers, may only give approximate results. Since inhibition percentage is affected by the concentration of the

Figure 5. Correlation of the percentage of inhibition of the studied compounds and the estimated binding energy which was derived by docking analysis (a) using the crystal structures selected according to predicted binding sites or (b) using the crystal structure 2F71 in which the inhibitor occupies sites different than the predicted for the studied compounds.
substrate in a way which differs according to the type of binding mode, the IC₅₀ value may vary if different substrate concentrations are used. Lower substrate concentrations may lead to lower IC₅₀ values in the case of competitive inhibitors, and to higher IC₅₀ values in case of uncompetitive inhibitors. Moreover, the kinetic properties of the enzyme, such as $K_m$, as well as the IC₅₀ and $K_i$ values of the inhibitors, may differentiate when different substrates are used. At least two different small molecules, p-nitro-phenyl-phosphate (pNPP) and fluorescein diphosphate (FDP) are commonly used by investigators for the measurement of PTP1b activity, and oligopeptides have also been used. A more reliable comparison could be done if reference compounds were used as inhibitors under the same conditions as the tested compounds. Since PTP1b is a relatively new drug target, no approved inhibitor, which could be used as a common reference compound by all researchers, exists in the market. This absolute comparison is not needed when the inhibition potential of the compounds is used for the evaluation of the prediction strategy used in the study, but may be important for the evaluation of the most active compounds as potential anti-diabetic drugs.

Conclusively, the studied compounds represent rather moderate PTP1b inhibitors. However, the active molecules may be used as initiative scaffolds for the design and synthesis of more effective inhibitors in the future.

Figure 6. Estimated binding energy of the studied compounds docked at the structure 2F71 and evaluated inhibition % of the studied compounds.
4. Conclusions

The two-step prediction strategy, which uses distances of crucial groups and structural characteristics of the compounds as a first step in order to predict the binding site of the compounds, and uses this prediction for the selection of the optimum crystal structure and the appropriate binding target for docking analysis, was effectively applied for the prediction of PTP1b activity of our compounds. Conclusively, this strategy may be a promising successful prediction approach in the case of enzymes with multiple binding sites.

Disclosure statement

The authors declare that they have no conflicts of interest.

References


