

## Central Lancashire Online Knowledge (CLOK)

|          |                                                                                                                                                                                                                                                                                                                                                                                                                        |
|----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Title    | Identification of C-β-d-Glucopyranosyl Azole-Type Inhibitors of Glycogen Phosphorylase That Reduce Glycogenolysis in Hepatocytes: In Silico Design, Synthesis, in Vitro Kinetics, and ex Vivo Studies                                                                                                                                                                                                                  |
| Type     | Article                                                                                                                                                                                                                                                                                                                                                                                                                |
| URL      | <a href="https://clock.uclan.ac.uk/id/eprint/29014/">https://clock.uclan.ac.uk/id/eprint/29014/</a>                                                                                                                                                                                                                                                                                                                    |
| DOI      | <a href="https://doi.org/10.1021/acscchembio.9b00172">https://doi.org/10.1021/acscchembio.9b00172</a>                                                                                                                                                                                                                                                                                                                  |
| Date     | 2019                                                                                                                                                                                                                                                                                                                                                                                                                   |
| Citation | Barr, Daniel, Szennyés, Eszter, Bokor, Éva, Al-Oanzi, Ziad H., Moffatt, Colin, Kun, Sándor, Docsa, Tibor, Sipos, Ádám, Davies, Matthew et al (2019) Identification of C-β-d-Glucopyranosyl Azole-Type Inhibitors of Glycogen Phosphorylase That Reduce Glycogenolysis in Hepatocytes: In Silico Design, Synthesis, in Vitro Kinetics, and ex Vivo Studies. ACS Chemical Biology, 14 (7). pp. 1460-1470. ISSN 1554-8929 |
| Creators | Barr, Daniel, Szennyés, Eszter, Bokor, Éva, Al-Oanzi, Ziad H., Moffatt, Colin, Kun, Sándor, Docsa, Tibor, Sipos, Ádám, Davies, Matthew, Mathomes, Rachel Thelma, Snape, Timothy James, Agius, Lorraine, Somsák, László and Hayes, Joseph                                                                                                                                                                               |

It is advisable to refer to the publisher's version if you intend to cite from the work.  
<https://doi.org/10.1021/acscchembio.9b00172>

For information about Research at UCLan please go to <http://www.uclan.ac.uk/research/>

All outputs in CLOK are protected by Intellectual Property Rights law, including Copyright law. Copyright, IPR and Moral Rights for the works on this site are retained by the individual authors and/or other copyright owners. Terms and conditions for use of this material are defined in the <http://clock.uclan.ac.uk/policies/>

**Identification of C- $\beta$ -D-glucopyranosyl azole type inhibitors of glycogen phosphorylase that reduce glycogenolysis in hepatocytes: *in silico* design, synthesis, *in vitro* kinetics and *ex-vivo* studies**

Daniel Barr,<sup>a#</sup> Eszter Szennyés,<sup>b#</sup> Éva Bokor,<sup>b#</sup> Ziad H. Al-Oanzi,<sup>c</sup> Colin Moffatt,<sup>d</sup> Sándor Kun,<sup>b</sup> Tibor Docsa,<sup>e</sup> Ádám Sipos,<sup>e</sup> Matthew P. Davies,<sup>a</sup> Rachel T. Mathomes,<sup>f</sup> Timothy J. Snape,<sup>f</sup> Lorraine Agius,<sup>\*c</sup> László Somsák,<sup>\*b</sup> Joseph M. Hayes<sup>\*f</sup>

<sup>a</sup> *School of Physical Sciences & Computing, University of Central Lancashire, Preston PR1 2HE, United Kingdom*

<sup>b</sup> *Department of Organic Chemistry, University of Debrecen, POB 400, H-4002 Debrecen, Hungary*

<sup>c</sup> *Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne NE2 4HH, United Kingdom*

<sup>d</sup> *Health & Life Sciences, De Montfort University, Gateway House, Leicester. LE1 9BH, United Kingdom*

<sup>e</sup> *Department of Medical Chemistry, Medical and Health Science Centre, University of Debrecen, Egyetem tér 1, H-4032 Debrecen, Hungary*

<sup>f</sup> *School of Pharmacy & Biomedical Sciences, University of Central Lancashire, Preston PR1 2HE, United Kingdom*

\* Corresponding authors

tel: +441912087033 (LA); email: [loranne.agius@newcastle.ac.uk](mailto:loranne.agius@newcastle.ac.uk)

tel: +3652512900 ext 22348 (LS); fax: +3652512744; e-mail: [somsak.laszlo@science.unideb.hu](mailto:somsak.laszlo@science.unideb.hu)

tel: +441772894334 (JMH); fax: +441772894981; email: [jhayes@uclan.ac.uk](mailto:jhayes@uclan.ac.uk)

#Equal contribution

## ABSTRACT

Several C- $\beta$ -D-glucopyranosyl azoles have recently been unravelled among the most potent glycogen phosphorylase (GP) catalytic site inhibitors discovered to date. Towards further exploring their translational potential, *ex-vivo* experiments have been performed for their effectiveness in reduction of glycogenolysis in hepatocytes. New compounds for these experiments were predicted *in silico* where, for the first time, effective ranking of GP catalytic site inhibitor potencies using the MM-GBSA method has been demonstrated. For a congeneric training set of 27 ligands, excellent statistics in terms of Pearson ( $R_p$ ) and Spearman ( $R_s$ ) correlations (both 0.98), predictive index ( $PI = 0.99$ ) and AU-ROC (0.99) for predicted versus experimental binding affinities were obtained, with ligand tautomeric/ionization states additionally considered using DFT calculations. Seven 2-aryl-4(5)-( $\beta$ -D-glucopyranosyl)-imidazoles and 2-aryl-4-( $\beta$ -D-glucopyranosyl)-thiazoles were subsequently synthesized and kinetics experiments against rabbit muscle GPb revealed new potent inhibitors with best  $K_i$ -s in the low  $\mu$ M range (**5c** = 1.97  $\mu$ M; **13b** = 4.58  $\mu$ M). Ten C- $\beta$ -D-glucopyranosyl azoles were then tested *ex-vivo* in mouse primary hepatocytes. Four of these (**5a-c** and **9d**) demonstrated significant reduction of glucagon stimulated glycogenolysis ( $IC_{50} = 30$ -60  $\mu$ M). Structural and predicted physicochemical properties associated with their effectiveness was analysed, with permeability related parameters identified as crucial factors. The most effective ligand series **5** contained an imidazole ring and the calculated  $pK_a$  (Epik: 6.2; Jaguar 5.5) for protonated imidazole suggests that cellular permeation through the neutral state is favoured, while within the cell there is predicted more favourable binding to GP in the protonated form.

**Keywords:** C-glycoside, imidazole, thiazole, glycogen phosphorylase inhibitor, permeability, MM-GBSA

## INTRODUCTION

Diabetes mellitus, characterized by hyperglycemia, is a chronic severe multiform disease with considerable socio-economic consequences. Recent WHO figures reveal an increase of cases from 108 million people in 1980 to 422 million in 2014, affecting ~8.5% of the global adult population.<sup>1</sup> This number may even be an underestimate due to methodological uncertainties as well as undiagnosed cases.<sup>2</sup> Diabetes has become one of the largest contributors to mortality due to long term complications such as retinopathy, neuropathy, nephropathy, but particularly cardiovascular diseases. Type 2 diabetes (T2D) is the most common form and involves abnormal insulin secretion and/or insulin resistance. Therapy for lowering blood glucose in T2D is by diet, exercise, and oral anti-hyperglycemic agents. In spite of the availability of different classes of anti-hyperglycaemic drugs,<sup>3</sup> treatments are often unable to achieve the required degree of glucose control for a large number of patients.<sup>4</sup>

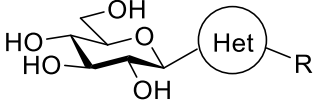
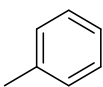
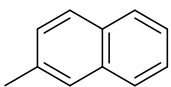
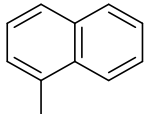
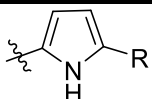
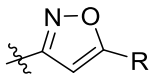
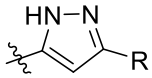
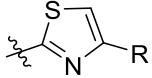
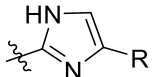
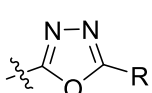
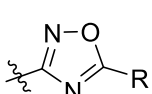
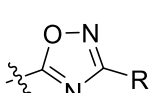
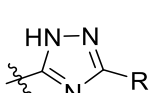
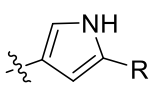
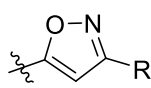
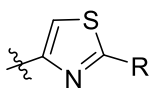
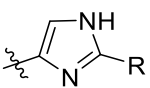
In T2D, glucose production by the liver is elevated and is contributing to the high blood glucose levels. Liver glycogenolysis releases glucose from glycogen into the bloodstream, with glycogen phosphorylase (GP; EC 2.4.1.1) the rate-limiting enzyme in this pathway.<sup>5</sup> Due to its key role in glycogen metabolism, GP inhibition has become a validated approach in the development of new T2D treatments.<sup>6</sup> Importantly, GP inhibitors (GPIs) also have revealed potential for treatment of other conditions such as myocardial and cerebral ischemias, and cancer.<sup>7, 8</sup>

Although several structural classes of GPIs have been reported,<sup>9-11</sup> glucose analogues binding at the catalytic site (competitive reversible inhibitors) have been the most explored<sup>10, 12</sup> and structure based inhibitor design efforts are greatly facilitated by the many solved crystallographic protein-ligand complexes.<sup>13</sup> Recently reported C- $\beta$ -D-glucopyranosyl azoles (Table 1, 1-9) comprise some of the most effective GPIs.<sup>14</sup> Their aromatic R groups,

particularly 2-naphthyl, enhance inhibitor potency by exploiting favourable interactions in the  $\beta$ -cavity of the catalytic site.<sup>10, 12</sup> Of particular interest, imidazoles<sup>14</sup> **5** and 1,2,4-triazoles<sup>14</sup> **9** have revealed inhibitors such as **5a,b** and **9b** with  $K_i$ -s in the nanomolar range. Compound **5b** is the most potent catalytic site inhibitor discovered to date. The efficacy of GPIs on glycogenolysis in cellular models<sup>15-18</sup> and blood glucose control *in vivo*<sup>18-21</sup> has been demonstrated for some inhibitors. However, the lack of such data for potent GPIs of the C- $\beta$ -D-glucopyranosyl azole type limits the potential for eventual translation into the clinic and the actual benefit of the patient, and has motivated the current work.

In this study, *in silico* screening has been performed for the new C- $\beta$ -D-glucopyranosyl azoles **10-13**, as well as ligands **3b,c** and **5c** also previously not synthesized or evaluated as GP inhibitors (Table 1). Initial docking experiments were followed by post-docking molecular mechanics – generalized Born surface area (MM-GBSA)<sup>22</sup> binding free energy (BFE) predictions, with the approach first validated for a training set of 27 ligands. Seven candidates were selected for synthesis and kinetics experiments against GP performed to determine the true potencies. Ten C- $\beta$ -D-glucopyranosyl azole GPIs with four different heterocycles and different R groups then underwent *ex-vivo* experiments in hepatocytes. These inhibitors were selected to analyse structural features and predicted physicochemical properties that may govern their efficacy for reduction of glycogenolysis at the cellular level.

**Table 1.** C-β-D-Glucopyranosyl azole type inhibitors (ref.<sup>14</sup> and references within) of rabbit muscle glycogen phosphorylase b ( $K_i$  [ $\mu$ M]) and the target compounds for this study.

|                                                                                          |                                                                                                  | Aromatic R groups                                                                                |                                                                                                   |                                                                                                    |
|------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------|
| <br>Het |                                                                                                  | <br><b>a</b>    | <br><b>b</b>    | <br><b>c</b>    |
| <b>1</b>                                                                                 |                 | No inh. at 625 $\mu\text{M}^a$                                                                   | No inh. at 625 $\mu\text{M}^a$                                                                    | NS <sup>c</sup>                                                                                    |
| <b>2</b>                                                                                 |                 | No inh. at 625 $\mu\text{M}^a$                                                                   | No inh. at 625 $\mu\text{M}^a$                                                                    | NS <sup>c</sup>                                                                                    |
| <b>3</b>                                                                                 |                 | 400 <sup>a</sup>                                                                                 | <sub>-</sub> <sup>b</sup>                                                                         | <sub>-</sub> <sup>b</sup>                                                                          |
| <b>4</b>                                                                                 |                 | 310 <sup>a</sup>                                                                                 | 158 <sup>a</sup>                                                                                  | NS <sup>c</sup>                                                                                    |
| <b>5</b>                                                                                 |                 | 0.28 <sup>a</sup>                                                                                | 0.031 <sup>a</sup>                                                                                | <sub>-</sub> <sup>b</sup>                                                                          |
| <b>6</b>                                                                                 |                | 10% inh. at 625 $\mu\text{M}^a$                                                                  | 10% inh. at 625 $\mu\text{M}^a$                                                                   | 10% inh. at 625 $\mu\text{M}^a$                                                                    |
| <b>7</b>                                                                                 |               | 10% inh. at 625 $\mu\text{M}^a$                                                                  | 38 <sup>a</sup>                                                                                   | No inh. at 625 $\mu\text{M}^a$                                                                     |
| <b>8</b>                                                                                 |               | 64 <sup>a</sup>                                                                                  | 11.6 <sup>a</sup>                                                                                 | 19 <sup>a</sup>                                                                                    |
| <b>9</b>                                                                                 |               | 7 <sup>a</sup>                                                                                   | 0.41 <sup>a</sup>                                                                                 | 11.5 <sup>23</sup>                                                                                 |
|                                                                                          | <br><b>10</b> | <br><b>11</b> | <br><b>12</b> | <br><b>13</b> |

Target compounds of this study with R groups **a-c**.

<sup>a</sup> Member of training set ligands for calculations. For full list of ligands see Table S1.

<sup>b</sup> Existing heterocyclic linkers with the indicated additional R substituent studied in this work.

<sup>c</sup> Not studied.

## RESULTS AND DISCUSSION

### *In Silico* Screening Studies

### Model Validation

A training set of 27 previously studied  $\beta$ -D-glucopyranosyl derivatives with twelve different 5-membered heterocyclic linkers (Table S1) was first used to validate an *in silico* screening protocol. These GP inhibitors had a good range of  $K_i$  values, from 0.031  $\mu$ M to no inhibition. ‘Activity’ was defined in terms of inhibitor potency throughout. The performance of docking (Glide-SP and -XP<sup>24</sup>, GOLD-ChemPLP<sup>25</sup>) and post-docking using MM-GBSA (Prime<sup>24</sup>) with respect to accurate prediction of relative ligand activities for the training set ligands was analyzed in a statistical manner.<sup>23</sup> Correlation between predicted and experimental activities was measured using Pearson ( $R_p$ ) and Spearman ( $R_s$ ) correlation coefficients.  $R_p$  takes into account the absolute differences between activity values within each set (predicted and experimental), whereas for  $R_s$ , only the activity rank is important since in each set of data the values are converted to a rank order. The ‘predictive index’ (PI)<sup>26</sup> was considered as a measure of how well a model differentiates inhibitors that have larger differences in potencies. A PI value (range -1 to +1) closer to +1 indicates better performance. Finally, the area under the ROC curve (AU-ROC) values are reported as the probability (range 0 to 1) of active compounds being ranked higher than inactive compounds. For this, a threshold  $K_i < 20 \mu$ M was used to define ‘actives’, leading to  $n = 7$  actives in the set of  $N = 27$  compounds.

The results of this statistical analysis are shown in Table 2. Comparing the different docking methods for predicted (GlideScores for Glide and ChemPLP fitness scores for GOLD) versus experimental (BFEs) binding affinities, only modest statistical performance was observed. Glide-SP performed best ( $R_p = 0.70$ ,  $R_s = 0.61$ ,  $PI = 0.63$ ), with an AU-ROC value of 0.83 indicating reasonably good potential to recognise inhibitors with  $K_i$ ’s  $< 20 \mu$ M threshold value. Overall, however, the statistics reflect inferior performance of docking methods for these ligands compared to a recent GP *in silico* screening study of 3-( $\beta$ -D-glucopyranosyl)-5-substituted-1,2,4-triazole derivatives.<sup>23</sup>

Significantly enhanced performance was achieved by refining the docking output using post-docking Prime MM-GBSA calculation of BFEs ( $\Delta G'_{bind}$ ):

$$\Delta G'_{bind} = \Delta E_{MM} + \Delta G_{solv} \quad \text{Eq. (1)}$$

$\Delta E_{MM}$  represents the molecular mechanics energy difference (internal, electrostatic and van der Waals) between the protein-ligand bound and unbound states calculated using the OPLS3 forcefield.  $\Delta G_{solv}$  is the corresponding solvation free energy change on binding calculated using a variable-dielectric generalized Born solvation model. The resulting BFEs were calculated both *with* and *without* incorporation of ligand reorganisation (strain) energy on binding (protein constrained to be rigid). Best statistical performance was obtained using the Glide-SP poses as input. Good statistics were obtained *with* ligand strain included ( $R_P = 0.85$ ;  $R_S = 0.74$ ;  $PI = 0.85$  and  $AU-ROC = 0.95$ ) but even better statistics *without* ligand strain were observed ( $R_P = 0.93$ ;  $R_S = 0.84$ ;  $PI = 0.92$  and  $AU-ROC = 0.98$ ). This best performing Prime MM-GBSA model (*with* and *without* strain) using Glide-SP poses is referred to as *Model 1* (Table 2).

**Table 2.** For the training set of twenty-seven  $\beta$ -D-glucopyranosyl azoles (Table S1), statistical analysis of the agreement between predicted and experimental activities as described in the text.<sup>a</sup>

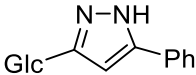
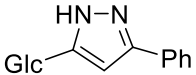
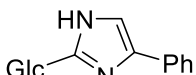
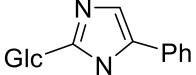
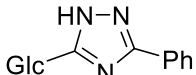
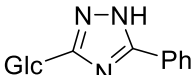
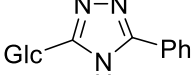
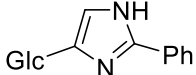
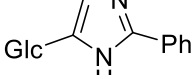
|                             | Pearson Correlation   |                         |                       | Spearman Correlation  |                         |                       |                       |                           |                       |                           |
|-----------------------------|-----------------------|-------------------------|-----------------------|-----------------------|-------------------------|-----------------------|-----------------------|---------------------------|-----------------------|---------------------------|
| Method                      | R <sub>P</sub>        | p-value                 | t-value               | R <sub>S</sub>        | p-value                 | t-value               | PI                    | p-value                   | AU-ROC                | p-value                   |
| Docking                     |                       |                         |                       |                       |                         |                       |                       |                           |                       |                           |
| Glide-SP                    | <b>0.70</b>           | <b>0.003</b>            | <b>3.62</b>           | <b>0.61</b>           | <b>0.012</b>            | <b>2.89</b>           | 0.63                  | 0.0084                    | <b>0.83</b>           | <b>0.0047</b>             |
| Glide-XP                    | 0.51                  | 0.043                   | 2.23                  | 0.44                  | 0.087                   | 1.83                  | <b>0.65</b>           | <b>0.0075</b>             | 0.73                  | 0.0406                    |
| GOLD-ChemPLP                | 0.64                  | 0.007                   | 3.13                  | 0.56                  | 0.024                   | 2.52                  | 0.63                  | 0.0091                    | 0.79                  | 0.0132                    |
| Prime                       |                       |                         |                       |                       |                         |                       |                       |                           |                       |                           |
| MM-GBSA <sup>b,c</sup>      |                       |                         |                       |                       |                         |                       |                       |                           |                       |                           |
| Glide-SP poses<br>(Model 1) | 0.85<br><b>(0.93)</b> | 0.000<br><b>(0.000)</b> | 6.07<br><b>(9.27)</b> | 0.74<br><b>(0.84)</b> | 0.001<br><b>(0.000)</b> | 4.06<br><b>(5.82)</b> | 0.85<br><b>(0.92)</b> | 0.0000<br><b>(0.0000)</b> | 0.95<br><b>(0.98)</b> | 0.0001<br><b>(0.0000)</b> |
| Glide-XP poses              | 0.76<br>(0.71)        | 0.001<br>(0.002)        | 4.38<br>(3.81)        | 0.59<br>(0.54)        | 0.016<br>(0.032)        | 2.74<br>(2.39)        | 0.76<br>(0.70)        | 0.0012<br>(0.0029)        | 0.86<br>(0.85)        | 0.0017<br>(0.0026)        |
| GOLD-ChemPLP poses          | 0.80<br>(0.77)        | 0.000<br>(0.001)        | 5.01<br>(4.53)        | 0.59<br>(0.57)        | 0.016<br>(0.021)        | 2.74<br>(2.60)        | 0.76<br>(0.73)        | 0.0019<br>(0.0012)        | 0.86<br>(0.86)        | 0.0017<br>(0.0021)        |
| Refined                     |                       |                         |                       |                       |                         |                       |                       |                           |                       |                           |
| MM-GBSA <sup>c,d</sup>      |                       |                         |                       |                       |                         |                       |                       |                           |                       |                           |

|                            |                       |                         |                        |                       |                         |                        |                       |                           |                       |                           |
|----------------------------|-----------------------|-------------------------|------------------------|-----------------------|-------------------------|------------------------|-----------------------|---------------------------|-----------------------|---------------------------|
| <i>Model 2<sup>e</sup></i> | 0.93<br><b>(0.98)</b> | 0.000<br><b>(0.000)</b> | 9.50<br><b>(16.61)</b> | 0.91<br><b>(0.98)</b> | 0.000<br><b>(0.000)</b> | 8.47<br><b>(16.94)</b> | 0.95<br><b>(0.99)</b> | 0.0000<br><b>(0.0000)</b> | 0.97<br><b>(0.99)</b> | 0.0000<br><b>(0.0000)</b> |
| <i>Model 3<sup>f</sup></i> | 0.84<br>(0.91)        | 0.000<br>(0.000)        | 5.73<br>(8.03)         | 0.70<br>(0.80)        | 0.003<br>(0.000)        | 3.67<br>(5.04)         | 0.83<br>(0.89)        | 0.0001<br>(0.0001)        | 0.94<br>(0.96)        | 0.0001<br>(0.0000)        |
| <i>Model 4<sup>g</sup></i> | 0.93<br>(0.97)        | 0.000<br>(0.000)        | 9.16<br>(14.95)        | 0.91<br>(0.97)        | 0.000<br>(0.000)        | 8.15<br>(15.08)        | 0.95<br>(0.98)        | 0.0000<br>(0.0000)        | 0.96<br>(0.97)        | 0.0000<br>(0.0000)        |

<sup>a</sup> Best statistics for docking, MM-GBSA and refined MM-GBSA methods are highlighted in ***bold italics***. For the correlation coefficients Rp and Rs, p-values are derived from the t-distribution (t-values) for n-2 degrees of freedom. P-values for PI were derived from a null distribution created from bootstrapping (n = 1,000,000) 'active' group allocation, and p-values for AU-ROC were derived from the Wilcoxon-Rank-Sum distribution since  $W = AU-ROC \times n \times (N - n)$ .<sup>27</sup> W is the Wilcoxon statistic (also known as Mann-Whitney U), N is the number of ligands (27) and n is the number of actives (7). P-values < 0.05 in all cases indicate statistical significance. All statistics were calculated using the R program.<sup>28</sup> <sup>b</sup> Calculated using Prime and Eq. (1). <sup>c</sup> Values are *with* ligand strain included but with values *without* ligand strain given in parentheses. <sup>d</sup> Using the best Prime MM-GBSA model with Eq.(1) from the Glide-SP poses (*Model 1*) as a benchmark for improvement. <sup>e</sup> *Model 1* data with tautomeric correction (Eq. (1)). <sup>f</sup> *Model 1* data with entropy included (Eq. (2)). <sup>g</sup> *Model 1* data with tautomeric correction & entropy included (Eq. (2)).

As predicted by LigPrep<sup>24</sup>, there were three heterocyclic linkers (pyrazole **3**, imidazole **5**, and 1,2,4-triazole **9**) in the training set and one heterocycle (imidazole **13**) among the target compounds with the potential to form tautomers (Table 3). The imidazole based ligands **5** and **13** were also predicted to have potential to be protonated. Until now, we have taken the best predicted  $\Delta G'_{bind}$  value of a ligand as the predicted value from the different possible ionization/tautomeric states. However, to consider more accurately the potential tautomeric effects on binding, DFT gas and solution phase calculations were performed to establish the most favourable tautomeric states of **3**, **5**, **9** and **13** ligands in the unbound state (Table 3). Additionally, the potential for protonation (+1 charge) of the imidazole based ligands **5** and **13** was probed using Epik and Jaguar pK<sub>a</sub> calculations.<sup>24</sup>

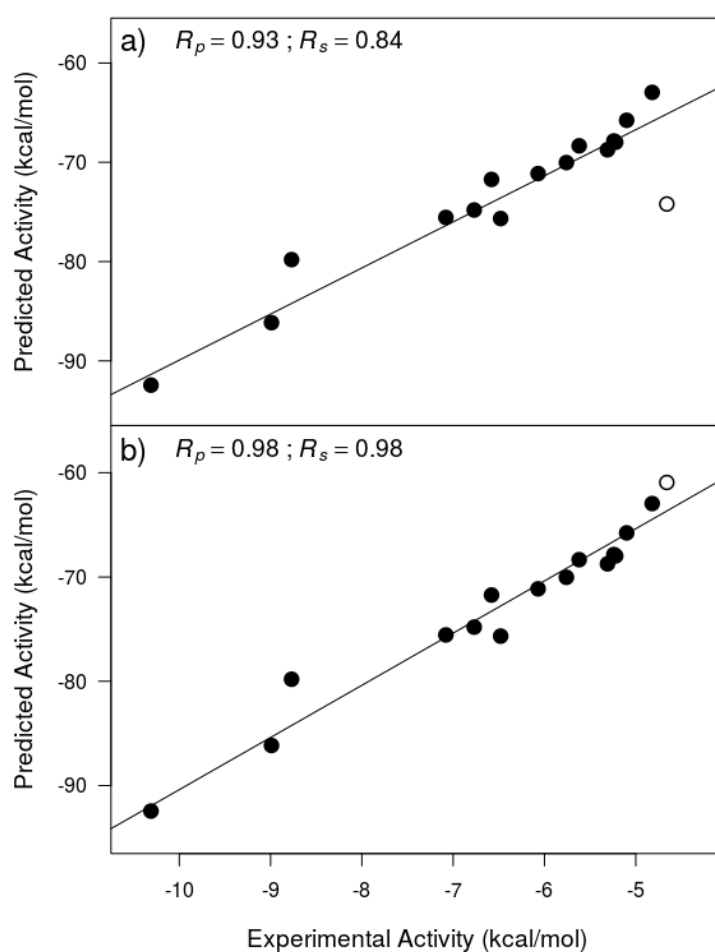
**Table 3.** Relative gas and solution phase energies (in parentheses) of the tautomeric forms of relevant ligands calculated using DFT at the M06-2X/6-31+G\* level of theory with water solvation effects included via a Solvation Model 8 (SM8) continuum treatment.<sup>a</sup>

| Heterocyclic compound  | Tautomer 1                                                                                      | Tautomer 2                                                                                      | Tautomer 3                                                                                       |
|------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|
| <b>3<sup>b,c</sup></b> | <br>0.0 (0.0)  | <br>3.0 (2.3)  | -                                                                                                |
| <b>5<sup>b,c</sup></b> | <br>0.0 (0.0)  | <br>0.6 (2.1)  | -                                                                                                |
| <b>9<sup>b</sup></b>   | <br>0.0 (0.0)  | <br>3.2 (1.6)  | <br>6.1 (6.2) |
| <b>13<sup>c</sup></b>  | <br>0.0 (0.0) | <br>4.8 (4.6) | -                                                                                                |

<sup>a</sup> As a model for the aromatic R groups (Table 1), a phenyl group was used. <sup>b</sup> Training set ligand heterocycle. <sup>c</sup> Predicted target ligand heterocycle.

The most favourable binding state of the imidazole **5** ligands was consistently predicted as the protonated (+1) state (*c.f.* **5b** binding in Figure 2(A)). In fact, considering only neutral states of the **5** ligands, the  $R_P$  value of 0.93 for *Model 1 without strain*, for example, decreases to 0.74. The calculated  $pK_a$  value of protonated **5** with R = phenyl was  $\sim 6$  (Epik: 6.2; Jaguar: 5.5) implying that these ligands will be mainly neutral in solution, but partial protonation means that this favourable binding will be facilitated. For the 1,2,4-triazole linker (**9**), only neutral state ligands were predicted by LigPrep and the most stable free state tautomer as calculated by DFT (tautomer 1, Table 3) was also the most favourable in terms of  $\Delta G'_{bind}$ . For pyrazole **3**, however, binding ( $\Delta G'_{bind}$ ) through tautomer 2 was preferred but the most stable free state was tautomer 1 ( $\sim 2$ -3 kcal/mol). Significantly, the best performing Prime MM-GBSA model using Glide-SP poses (*Model 1*) had ligand **3a** (noting that **3b** and **3c** were targets for

prediction) as an outlier in the plots of  $\Delta G'_{bind}$  against the experimental BFEs (Figure 1(a)). Taking instead the  $\Delta G'_{bind}$  value for tautomer 1 led to further improvement in performance statistics (*Model 2* – data from *Model 1* but including tautomeric correction, Table 2). For *Model 2*, excellent values for each metric *with* ligand strain ( $R_P = 0.93$ ;  $R_S = 0.91$ ;  $PI = 0.95$  and  $AU-ROC = 0.97$ ) and particularly *without* ligand strain correction ( $R_P = 0.98$ ;  $R_S = 0.98$ ;  $PI = 0.99$  and  $AU-ROC = 0.99$ ) were now obtained (Figure 1(b)).



**Figure 1.** Scatter plots of predicted  $\Delta G'_{bind}$  values (Eq. (1)) against experimental binding free energies obtained using MM-GBSA *without* ligand strain **(a)** *before* tautomeric correction (*Model 1*) and **(b)** *after* tautomeric correction (*Model 2*) (c.f. Table 2). Tautomeric correction for **3a** is displayed via an open circle. The lines drawn on the plots are from a Type II (standard major axis) regression model which allows error in both variables<sup>29</sup> and is appropriate in this case.

We additionally probed the effect of inclusion of entropy effects ( $\Delta S_{MM}$ ) as per  $\Delta G_{bind}$  (Eq. (2)), accounting for the loss of ligand vibrational, rotational and translational entropy on binding calculated using the Rigid Rotor Harmonic Oscillator approximation with the OPLS3 forcefield.

$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{solv} - T\Delta S_{MM} \quad \text{Eq. (2)}$$

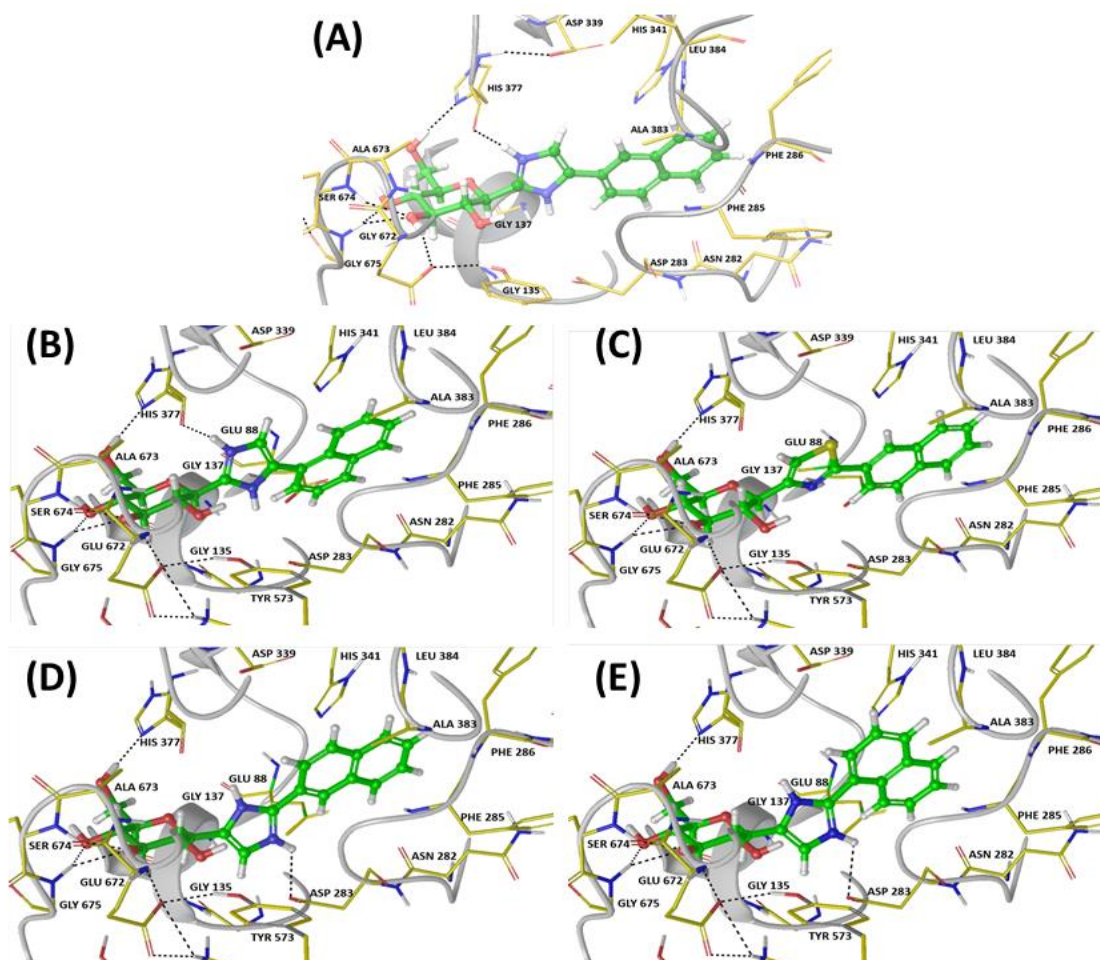
The effect of its inclusion on the performance statistics was minimal (*Models 3* and *4*, Table 2), which could be attributed to the similarity among this congeneric series of ligands. Entropy inclusion is often a source of uncertainty in calculations of this type but in cases does improve quality of results.<sup>22, 30, 31</sup> Based on the statistics (Table 2), we considered the MM-GBSA *Model 2* (Glide-SP poses as input; tautomeric correction; neglectation of entropy effects) the most efficient for predictions on the target compounds (Table 1). This is the first time that the endpoint MM-GBSA method has been demonstrated as an effective method for ranking of congeneric series of ligands for the GP catalytic site. Previous comparisons of the effectiveness of MM-GBSA and MM-PBSA revealed that the prediction accuracies can be quite different for different protein families.<sup>32</sup>

### *Screening of Target Compounds*

The full set of predicted  $\Delta G'_{bind}$  values for the target compounds using *Model 2* are reported in Table S2. As benchmarks for comparison, the predicted  $\Delta G'_{bind}$  value ranges for the training set active compounds ( $K_i$ 's < 20  $\mu$ M) were -66.5 – -87.9 kcal/mol and -71.7 – -92.4 kcal/mol *with* and *without* ligand strain corrections included, respectively.

The 1-naphthyl substituted imidazole linker (**5c**) was predicted best among the target compounds with very good  $\Delta G'_{bind}$  values of -85.2 kcal/mol and -93.8 kcal/mol *with* and *without* ligand strain included, respectively, and was immediately a synthetic candidate. It is predicted to bind in the protonated state of the heterocycle (Figure 2(B)). While none of the

new heterocyclic scaffolds (**10-13**) from Table 1 were predicted to produce more potent inhibitors than the pre-existing most effective heterocycles (imidazole **5** and 1,2,4-triazole **9**), the imidazole **13** did yield reasonably good  $\Delta G'_{bind}$  values. Compounds **13a**, **13b** and **13c** had  $\Delta G'_{bind}$  values of  $-72.4$ ,  $-78.5$  and  $-79.5$  kcal/mol *with* ligand strain, respectively, and  $-78.9$ ,  $-86.7$  and  $-88.9$  kcal/mol *without* ligand strain included, respectively. The predicted binding modes of **13b** and **13c** to GPb are shown in Figure 2(D) and 2(E), respectively, where, like the isomeric imidazole **5c**, the predicted most favourable binding state of the heterocycle is protonated. The calculated  $pK_a$  value of protonated **13** with R = phenyl (Epik: 6.1; Jaguar: 5.4) supports partial protonation for binding in this state. In terms of the other ligands screened, the tautomeric corrected  $\Delta G'_{bind}$  values were poor for **3b** and **3c** (e.g.  $\sim -55$  kcal/mol *with* ligand strain), possibly in line with the existing  $K_i$  of 400  $\mu$ M for **3a**<sup>33</sup>, and the other heterocycles (**10-12** ligands) were predicted only borderline close to active ( $K_i$ 's  $< 20$   $\mu$ M). We did, however, consider the thiazole (**12**) based analogues interesting given the potential of sulphur to improve cell permeability and due to sulphur atom occurrence in a large number of marketed drugs, particularly in the form of heterocycles.<sup>34</sup> The predicted binding of **12b** is shown in Figure 2(C). All things considered, imidazole **5c**, and imidazoles **13b** and **13c** appeared the most attractive candidates from the *in silico* screening and selected for synthesis, together with **12a-c** despite their lower predicted potency rankings (Table S2) to study the effects of a sulphur based heterocycle. For comparative purposes, the R = phenyl analogue **13a** was also chosen.



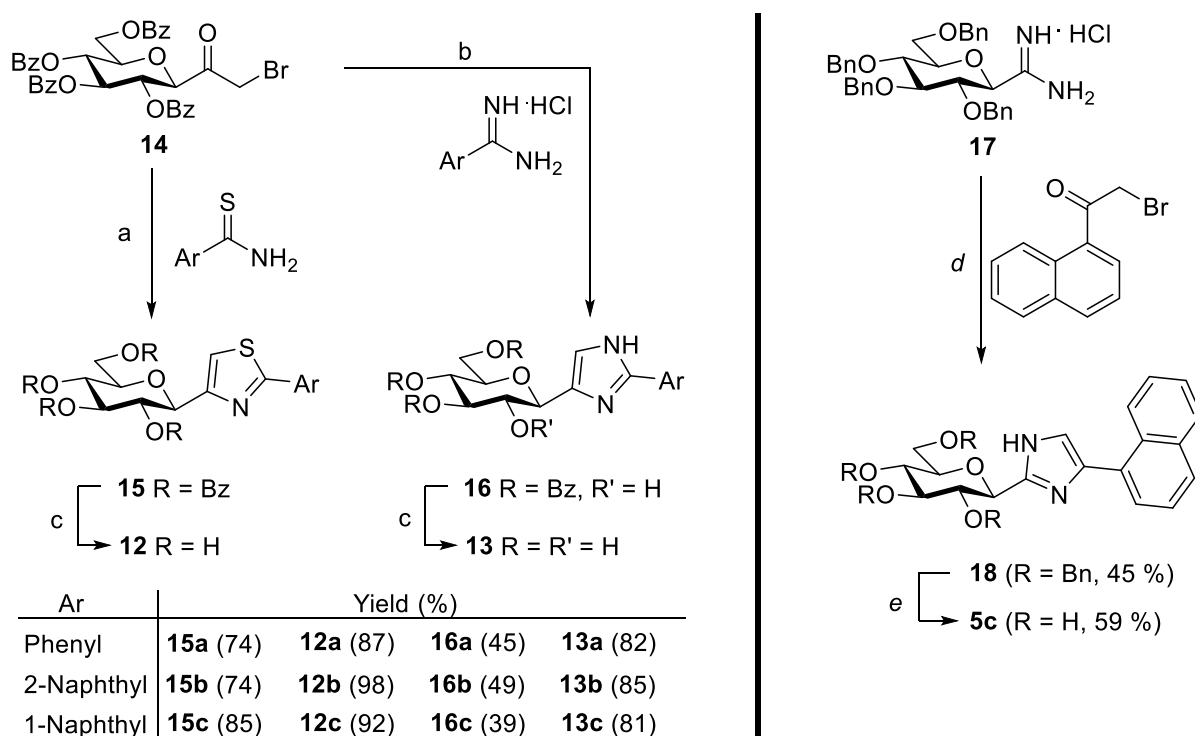
**Figure 2.** Predicted binding interactions of (A) **5b**, (B) **5c**, (C) **12b**, (D) **13b** and (E) **13c** with glycogen phosphorylase b as calculated using MM-GBSA. For **5b** and **5c**, the protonated heterocycle exploits a favourable hydrogen bond interaction with His377 O. The other NH is close to but not within hydrogen bonding distance of the Asp283 sidechain carboxylate but exploits ion-ion interactions. RMSD (heavy atoms) of the predicted **5b** position compared to its native crystallographic complex (PDB code: 5JTU) was just 0.250 Å. For **13b** and **13c**, compared to isomeric **5b** and **5c**, the heterocycles and NH interactions with His377 O are weaker (distances 3.4–3.6 Å). The ligands instead form hydrogen bond and ion-ion interactions with the Asp283 sidechain carboxylate. Also, the aromatic groups occupy different positions in the  $\beta$ -cavity to imidazole 5 type ligands. For **12b**, the thiazole does not exploit any direct hydrogen bond interactions with GP, only a weak heterocycle CH to His377 O interaction, but the 2-naphthyl moiety extends deep into the  $\beta$ -cavity.

## Syntheses

For the preparation of 2-aryl-4-( $\beta$ -D-glucopyranosyl)-thiazoles **12** (Scheme 1) glucosyl bromomethyl ketone **14**<sup>35</sup> was cyclized with thiobenzamide and naphthalene-2- and -1-

thiocarboxamide in anhydrous DMF to afford the protected thiazoles **15a-c**, respectively. Debenzoylation of **15a-c** was carried out by the Zemplén protocol providing the deprotected derivatives **12a-c** in very good yields.

When compound **14** was treated with arene carboxamidines in a THF-H<sub>2</sub>O solvent mixture in the presence of K<sub>2</sub>CO<sub>3</sub> the cyclisation was accompanied by hydrolytic cleavage of the 2'-O-benzoyl group resulting in partially protected derivatives **16a-c** in moderate yields. Subsequently, complete debenzoylation of **16a-c** was accomplished by the Zemplén method to give excellent yields of the final products **13a-c**.



**Scheme 1.** Reagents and conditions: a) dry DMF, 140 °C; b) 4 equiv. K<sub>2</sub>CO<sub>3</sub>, THF-H<sub>2</sub>O (4 : 1), rt; c) ~ 1M NaOMe in MeOH, rt.; d) 2 equiv. K<sub>2</sub>CO<sub>3</sub>, THF-H<sub>2</sub>O (4 : 1), rt; e) 40 equiv. EtSH, 20 equiv. BF<sub>3</sub>·Et<sub>2</sub>O, dry CH<sub>2</sub>Cl<sub>2</sub>, rt.

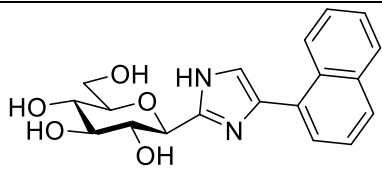
The synthesis of 4-(1-naphthyl)-imidazole derivative **5c** was performed by the cyclization of formamidinium salt<sup>36,37</sup> **17** with the corresponding  $\alpha$ -bromo-ketone under basic conditions to give

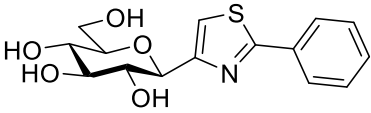
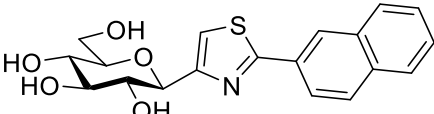
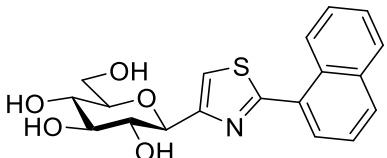
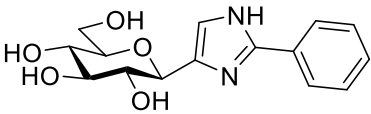
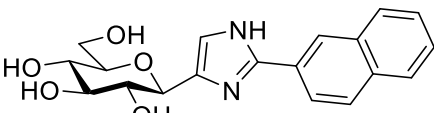
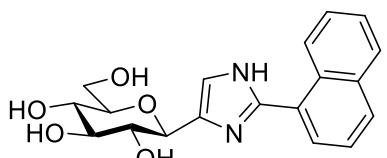
the protected imidazole **18** in acceptable yield. Cleavage of the *O*-benzyl groups in **18** was effected by treatment with ethanethiol in the presence of  $\text{BF}_3 \cdot \text{Et}_2\text{O}$  to obtain the test compound **5c** in good yield.

## Kinetics Results

The new compounds were assayed against rmGPb as described earlier<sup>35</sup> and the inhibition constants are summarized in Table 4. Consistent with glucose-derived inhibitors studied so far, these new compounds also proved competitive catalytic site inhibitors as indicated by the Dixon plots of the primary data (see Supporting Information); that **12b** and **13a** may also bind at a secondary site cannot be confirmed or fully discarded from Dixon plots. The most potent predicted target compound **5c** was also the most potent experimentally ( $K_i = 1.5 \mu\text{M}$ ), although it is not as potent as its previously reported analogues, **5a** ( $K_i = 0.28 \mu\text{M}$ ) and **5b** ( $K_i = 0.031 \mu\text{M}$ ). The potencies of the isomeric imidazole **13b** ( $K_i = 4.58 \mu\text{M}$ ) was in line with expectations from the predictions being a low micromolar inhibitor, but **13a** ( $K_i = 68.6 \mu\text{M}$ ) and **13c** ( $K_i = 71.1 \mu\text{M}$ ) were less potent than expected. Overprediction of 1-naphthyl analogues inhibitor potencies was therefore observed for this particular R group that was less well represented in the model validation training set ligands (Table S1). Thiazole **12b** ( $K_i = 26.2 \mu\text{M}$ ) was borderline close to our defined threshold  $K_i < 20 \mu\text{M}$  for activity, while its phenyl **12a** ( $K_i = 326 \mu\text{M}$ ) and 1-naphthyl **12c** ( $K_i = 540 \mu\text{M}$ ) analogues were much weaker inhibitors.

**Table 4.** Inhibitory effect of the new C- $\beta$ -D-glucopyranosyl heterocycles against rabbit muscle GPb

| Compound                                                                                             | $K_i$ [ $\mu\text{M}$ ] |
|------------------------------------------------------------------------------------------------------|-------------------------|
|  <p><b>5c</b></p> | 1.97                    |

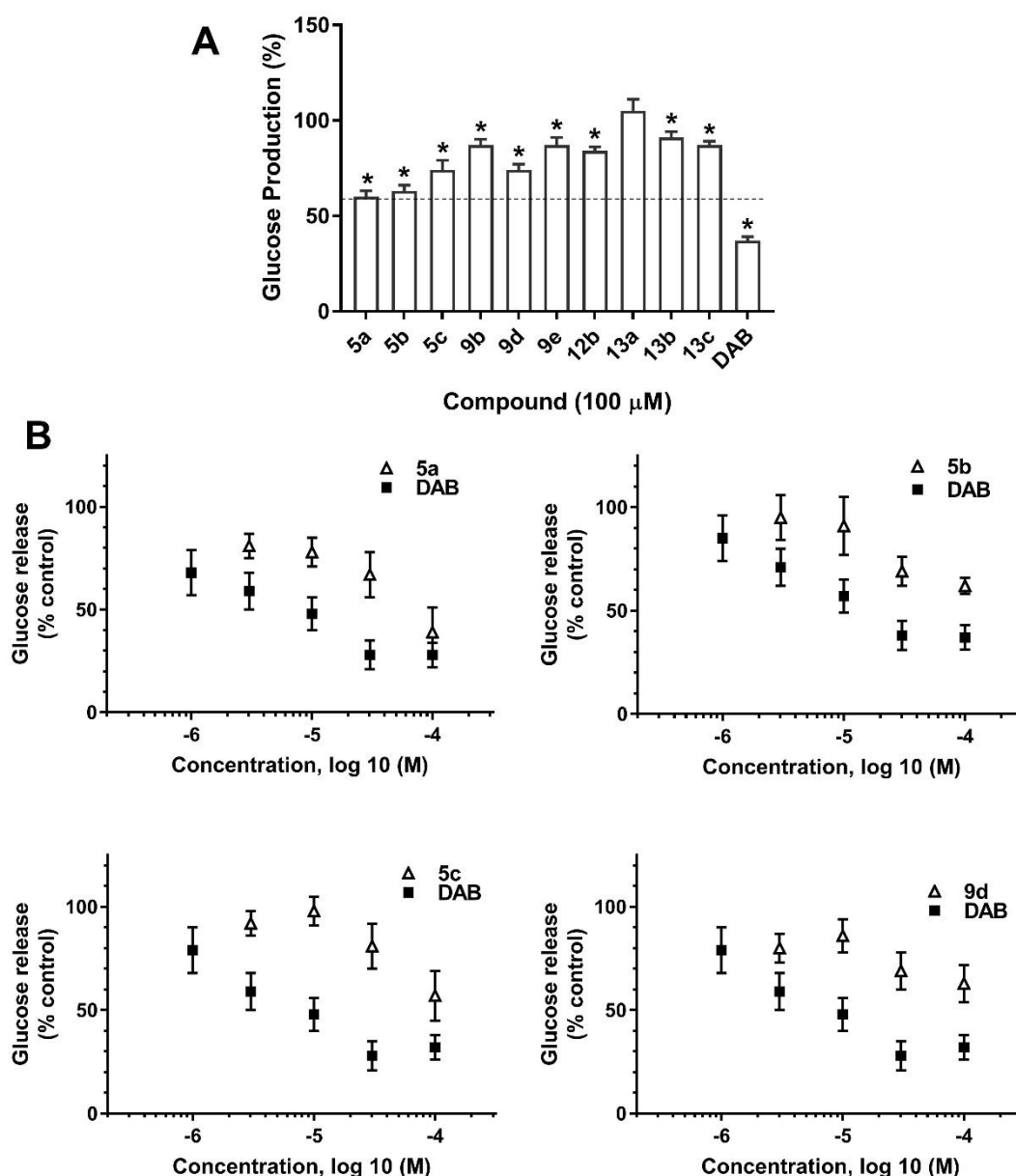
|                                                                                     |      |
|-------------------------------------------------------------------------------------|------|
|    | 326* |
| <b>12a</b>                                                                          |      |
|    | 26.2 |
| <b>12b</b>                                                                          |      |
|    | 540* |
| <b>12c</b>                                                                          |      |
|    | 68.6 |
| <b>13a</b>                                                                          |      |
|   | 4.58 |
| <b>13b</b>                                                                          |      |
|  | 71.1 |
| <b>13c</b>                                                                          |      |
| *Calculated from the IC <sub>50</sub> value by a web-based tool. <sup>38</sup>      |      |

## Cellular and Pharmacokinetic Evaluation

Based on the kinetics results and to allow analysis of inhibitor structural features that may influence efficacy at the cellular level, compounds **5a**, **5b**, and **9b** from Table 1, **5c**, **12b**, **13a**, **13b** and **13c** from Table 4, and **9d** and **9e** shown in Table 5 were chosen ( $K_i$ -s in the range of 0.03 – 71  $\mu$ M) to study their effects on glycogenolysis in hepatocytes. The pharmacokinetic profiles of the compounds were also predicted using QikProp 4.9.<sup>24</sup>

The IC<sub>50</sub> values for cellular inhibition of glycogenolysis of the 10 compounds are shown in Table 5 together with the predicted absorption, distribution, metabolism and excretion

(ADME) properties. Glycogenolysis was determined from glucose release in the presence of glucagon to promote phosphorylation (activation) of GP and fractional inhibition of glucose production by the compounds (at 100  $\mu$ M) is shown in Figure 3. 1,4-Dideoxy-1,4-imino-D-arabinitol (DAB) was used as a reference compound because of its high potency ( $IC_{50}$  of  $\leq 1$   $\mu$ M) at inhibiting phosphorylated and non-phosphorylated forms of GP and known efficacy at the cellular level (Figure 3).<sup>39, 40</sup> Unlike glucose analogues which stabilize the T-conformation, DAB stabilizes the R-conformation and favours conversion of GPb to GPa,<sup>41, 42</sup> however, it inhibits GPa to GPb similarly.<sup>43</sup> DAB also inhibits glucosidase including the glycogen debranching enzyme but with 10-fold lower affinity compared with GPa and GPb.<sup>43, 44</sup> Compounds **5a** and **5b** showed the highest cellular efficacy ( $IC_{50}$  of 30-40  $\mu$ M), followed by **5c** and **9d** ( $IC_{50}$  of  $\sim 60$   $\mu$ M). The other compounds (**9b**, **9e**, **12b**, **13b**, **13c**) except **13a** caused significant inhibition of glucose release at 100  $\mu$ M but with an  $IC_{50} > 100$   $\mu$ M. Some compounds including **9d** caused release of lactate dehydrogenase at concentrations  $> 100$   $\mu$ M, which is suggestive of possible counter-effects of compromised calcium homeostasis on glycogenolysis at high concentrations  $\geq 100$   $\mu$ M.



**Figure 3.** Glucose release from mouse hepatocytes incubated with glucagon and the compounds indicated in comparison with DAB used as the reference compound. **A.** Fractional inhibition by compounds determined at 100  $\mu\text{M}$ , expressed as percentage of control incubations with glucagon, dashed line indicates glucose production in the absence of glucagon. **B.** Comparison with DAB at varying compound concentration. Glucose release in the absence of glucagon was  $58 \pm 3\%$ . Means  $\pm$  SEM,  $n=6-8$ .

Ligand series 5 were therefore the most effective for both purified enzyme GP inhibition ( $K_i$ 's = 0.031 – 1.97  $\mu\text{M}$ ) and also in the cellular assays ( $\text{IC}_{50}$ 's  $\sim$  30 – 60  $\mu\text{M}$ ). These ligands will be predominantly neutral (predicted  $pK_a \sim 6$  for protonated imidazole **5a**) in the free state and

favour permeability. Within the cell, the protonated form was predicted to bind strongest at GP. Considering Lipinski's rules of five and Jorgensen's rule of three for oral bioavailability, calculated properties for **5a-c** are consistent (no violations) with favourable absorption and permeability (assuming the inhibitors are not substrates for biological transporters). Veber's rules<sup>45</sup> consider other parameters for description of drug-likeness with too many rotatable bonds ( $>10$ ) and too large a polar surface area ( $\text{PSA} > 140 \text{ \AA}^2$ ) bad for absorption and permeation. These rules were also passed for the **5** series of compounds. For the isomeric imidazole **13** series,  $\text{IC}_{50}$ 's were  $> 100 \text{ }\mu\text{M}$  in the cellular assays; however, the purified enzyme inhibitions ( $K_i$ 's =  $4.58 - 71 \text{ }\mu\text{M}$ ) were also less for **13a-c** (compared to **5a-c**). For **13a**, the not significant cellular inhibition could in part stem from its lower log P of  $-0.09$  (theoretical log D of  $-0.12$  at pH 7).

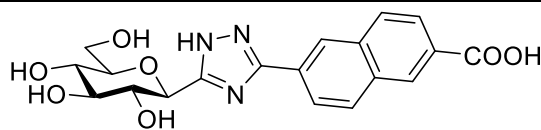
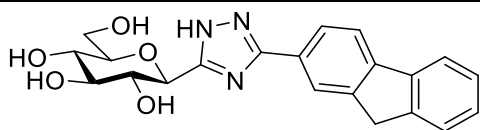
Of the **9** series, both **9b** and **9e** had  $\text{IC}_{50} > 100 \text{ }\mu\text{M}$  despite their potent inhibition of the free GP enzyme ( $K_i = 0.41$  and  $4.42 \text{ }\mu\text{M}$ , respectively). However, **9d** ( $K_i = 1.19 \text{ }\mu\text{M}$ ) was one of the most effective compounds at the cellular level ( $\text{IC}_{50} = 57 \text{ }\mu\text{M}$ ), indicating that differences in compound physicochemical properties arising from the R group substitutions could be important. Comparing **9b** ( $K_i = 0.41 \text{ }\mu\text{M}$ ) first with **5b** ( $K_i = 0.031 \text{ }\mu\text{M}$ ), both of which have R = 2-naphthyl and are two of the most potent GPIs known to date, the lack of cellular efficacy for **9b** ( $\text{IC}_{50} > 100 \text{ }\mu\text{M}$ ) compared to **5b** ( $\text{IC}_{50} = 28 \text{ }\mu\text{M}$ ) could be in part attributed to its less favorable lipophilicity (log P =  $-0.01$  versus  $0.63$ ) and PSA ( $136 \text{ \AA}^2$  versus  $121 \text{ \AA}^2$ ), that is, properties important for cellular permeability. On the other hand, the cellular efficient **9d** ( $\text{IC}_{50} = 57 \text{ }\mu\text{M}$ ) with R = 2-fluorenyl has a predicted log P of  $0.57$  close to that of ligand **5b**, reinforcing permeability as a key parameter. In further agreement, the other cellular inefficient **9e** ligand ( $\text{IC}_{50} > 100 \text{ }\mu\text{M}$ ) contains an anionic (carboxyl) 2-naphthyl substituent, and had one violation each of Lipinski's ( $\text{HBD} = 6$ ), Jorgensen's ( $\text{Caco-2} = 1.5 \text{ nm s}^{-1}$ ) and Veber's (PSA

= 187 Å<sup>2</sup>) rules, violations consistent with poor permeability. The cellular performance of thiazole derivative **12b** (IC<sub>50</sub> > 100 μM) was disappointing despite its predicted parameters favorable for permeability. There was one violation of Jorgensen's rules (NPM = 7) and it is possible that metabolism could be the more important factor in this case.

**Table 5.** IC<sub>50</sub> results for glucagon-stimulated glycogenolysis in mouse hepatocytes together with ADME property predictions for the studied glycogen phosphorylase inhibitors. Purified enzyme inhibition constants (*K<sub>i</sub>*'s) are also shown for comparison.<sup>[a]</sup>

| Ligand                         | Lipinski's Rule of Five and Violations (V) <sup>[b]</sup> |                    |                    |                               |   | Jorgensen's Rule of Three & Violations (V) <sup>[b]</sup> |                 |                    |   | Veber's Rules & Violations (V) <sup>[b]</sup> |                     |   |                             |                                     |
|--------------------------------|-----------------------------------------------------------|--------------------|--------------------|-------------------------------|---|-----------------------------------------------------------|-----------------|--------------------|---|-----------------------------------------------|---------------------|---|-----------------------------|-------------------------------------|
|                                | <i>M<sub>r</sub></i><br>[Da]                              | HBD <sup>[c]</sup> | HBA <sup>[d]</sup> | log <i>P</i> <sub>(o/w)</sub> | V | Caco-2<br>[nm s <sup>-1</sup> ] <sup>[e]</sup>            | log <i>S</i>    | NPM <sup>[f]</sup> | V | PSA [Å <sup>2</sup> ] <sup>[g]</sup>          | NROT <sup>[h]</sup> | V | <i>K<sub>i</sub></i><br>GPb | <i>IC</i> <sub>50</sub><br>Cellular |
|                                | (<500)                                                    | (≤5)               | (≤10)              | (<5)                          |   | (>22)                                                     | (>-5.7)         | (<7)               |   | (<140 Å <sup>2</sup> )                        | (≤10)               |   | (μM)                        | (μM)                                |
| <b>5a</b>                      | 306.3                                                     | 5                  | 7                  | -0.15(-0.16) <sup>[i]</sup>   | 0 | 117.0                                                     | -2.8            | 6                  | 0 | 121                                           | 5                   | 0 | 0.28                        | 37±14                               |
| <b>5b</b>                      | 356.4                                                     | 5                  | 7                  | 0.63 (0.62) <sup>[i]</sup>    | 0 | 112.9                                                     | -3.7            | 6                  | 0 | 121                                           | 5                   | 0 | 0.031                       | 28 ± 8                              |
| <b>5c</b>                      | 356.4                                                     | 5                  | 7                  | 0.57 (0.56) <sup>[i]</sup>    | 0 | 116.5                                                     | -3.5            | 6                  | 0 | 122                                           | 5                   | 0 | 1.544                       | 63±6                                |
| <b>9b</b>                      | 357.4                                                     | 5                  | 8                  | -0.01                         | 0 | 58.8                                                      | -3.5            | 6                  | 0 | 136                                           | 5                   | 0 | 0.41                        | >100                                |
| <b>9d</b>                      | 395.4                                                     | 5                  | 8                  | 0.57                          | 0 | 58.8                                                      | -4.3            | <b>7</b>           | 1 | 137                                           | 5                   | 0 | 1.19                        | 57±12                               |
| <b>9e</b>                      | 401.4                                                     | <b>6</b>           | 10                 | -0.67                         | 1 | <b>1.5*</b>                                               | -3.5            | 6                  | 1 | <b>187</b>                                    | 6                   | 1 | 4.42                        | >100                                |
| <b>12b</b>                     | 373.4                                                     | 4                  | 6                  | 1.21                          | 0 | 182.8                                                     | -4.1            | <b>7</b>           | 1 | 106                                           | 5                   | 0 | 23                          | >100                                |
| <b>13a</b>                     | 306.3                                                     | 5                  | 7                  | -0.09 (-0.12) <sup>[i]</sup>  | 0 | 131.1                                                     | -2.9            | 6                  | 0 | 120                                           | 5                   | 0 | 37                          | NS <sup>[j]</sup>                   |
| <b>13b</b>                     | 356.4                                                     | 5                  | 7                  | 0.71 (0.68) <sup>[i]</sup>    | 0 | 130.8                                                     | -3.8            | 6                  | 0 | 120                                           | 5                   | 0 | 5.4                         | >100                                |
| <b>13c</b>                     | 356.4                                                     | 5                  | 7                  | 0.68 (0.65) <sup>[i]</sup>    | 0 | 135.3                                                     | -3.7            | 6                  | 0 | 120                                           | 5                   | 0 | 93                          | >100                                |
| <b>Range</b><br><sup>[k]</sup> | <b>130-725</b>                                            | <b>0-6</b>         | <b>02-20</b>       | <b>-2-6.5'</b>                | - | <b>&lt;25 poor;<br/>&gt; 500<br/>great</b>                | <b>-6.5-0.5</b> | <b>1-8</b>         | - | <b>7-200</b>                                  | <b>0-15</b>         | - | -                           |                                     |

[a] ADME properties were calculated using Qikprop 4.9; predicted properties outside the range for 95% of known drugs are indicated with an asterisk (\*). [b] Rules as listed in the columns, with any violations of the rules highlighted in bold italics. [c] Number of hydrogen bond donors. [d] Number of hydrogen bond acceptors. [e] Caco-2 cell permeability. [f] Number of primary metabolites. [g] PSA represents the van der Waals (polar) surface areas of N and O atoms. [h] Number of rotatable bonds. [i] Theoretical log D at pH 7 based on QM calculated Jaguar pK<sub>a</sub> is given in parentheses. [j] Not significant. [k] Range for 95% of known drugs - QikProp User's Manual.



Potential permeability issues for glucose analogues have been highlighted in the past.<sup>15, 23, 46, 47</sup>

Here, log P in particular was identified as an important benchmark, but with exceptions such

as for **12b**. The log P range of the ten compounds tested was -0.67 – 1.21 which is outside the proposed log P (~2-4) sweet spot range in medicinal chemistry,<sup>48</sup> and is proposed as a target range in future inhibitor design efforts. Further, a recent study of D-glucopyranosylidene-spiro-isoxazolines revealed that the *O*-peracetylated form of its glucose unit in the 2-naphthyl analogue showed a lower IC<sub>50</sub> by a factor of 8 in primary cell cultures compared to its unprotected hydroxyl equivalent.<sup>18</sup> Although hydrogen bonds increase solubility, they must be broken for a compound to permeate through the membrane lipid bilayer. Finally, we highlight that a further consideration when assessing the effects of glucose analogues in hepatocytes in comparison with physiological effectors is that the efficacy of high glucose on GP is in part due to the raised cell glucose 6-P which acts synergistically with glucose in stabilizing the T-state.<sup>49-51</sup>

## Conclusions

C-β-D-Glucopyranosyl azoles are potent inhibitors of GP and have potential for treatment of T2D and inhibition of glycogenolysis in other conditions. New inhibitors have been designed and synthesised motivated by *in silico* predictions. Post-docking MM-GBSA binding affinity calculations with tautomeric correction produced excellent performance statistics for a congeneric trainings set of 27 ligands. For the first time, MM-GBSA has been demonstrated as an effective approach to study GP catalytic site inhibitors. As a result, four new micromolar GPIs (**5c**, **12b**, **13a** and **13b**) have been identified, with **5c** the most potent ( $K_i = 1.97 \mu\text{M}$ ). Ten C-β-D-glucopyranosyl azoles, consisting of some of the most potent GP inhibitors known to date, were tested for inhibition of glucagon-stimulated glycogenolysis in hepatocytes. Four of these, **5a**, **5b**, **5c** and **9d** were effective at inhibiting glycogenolysis at low micromolar concentrations with **5a** ( $IC_{50} = 37 \mu\text{M}$ ) and **5b** ( $IC_{50} = 28 \mu\text{M}$ ) the most efficient. Therefore, the most potent purified enzyme GPIs (**5a** and **5b**) were also the most potent inhibitors of

glycogenolysis. The calculated  $pK_a$  (Epik: 6.2; Jaguar 5.5) of the **5a** protonated imidazole ligand is consistent with passive permeation of the neutral form but with preferred binding at GP predicted as the protonated state of the ligand. To summarize, it has been demonstrated that linking the structural features and related physicochemical properties of glucopyranosyl type GPIs with their potential efficacy at the cellular level can help guide the design of more drug-like analogues in the future. Log P in particular has been identified as a simple predictable parameter that can be readily considered in this regard. Most importantly, the benefits of integrating cellular efficacy studies into GP inhibitor design efforts has been demonstrated towards the more efficient identification of drug-like compounds.

## ACKNOWLEDGEMENTS

This work was supported by the National Research, Development and Innovation Office of Hungary (Grant FK125067) and the EU co-financed by the European Regional Development Fund under the projects GINOP-2.3.2-15-2016-00008 and GINOP-2.3.3-15-2016-00004.

*Supporting Information Available:* This material is available free of charge via the internet at <http://pubs.acs.org>. Included in this document are Table S1 and S2, and the full experimental methods.

## REFERENCES

- [1] World Health Organisation, [www.who.int](http://www.who.int) [Accessed 20.02.2019].
- [2] Green, A., Hirsch, N. C., and Pramming, S. K. (2003) The changing world demography of type 2 diabetes, *Diabetes/Metab. Res. Rev.* 19, 3-7.
- [3] Israili, Z. H. (2011) Advances in the Treatment of Type 2 Diabetes Mellitus, *Am. J. Ther.* 18, 117-152.
- [4] Wagman, A. S., and Nuss, J. M. (2001) Current therapies and emerging targets for the treatment of diabetes, *Curr. Pharm. Des.* 7, 417-450.
- [5] Bollen, M., Keppens, S., and Stalmans, W. (1998) Specific features of glycogen metabolism in the liver, *Biochem. J.* 336, 19-31.

- [6] Oikonomakos, N. G. (2002) Glycogen phosphorylase as a molecular target for type 2 diabetes therapy, *Curr. Protein Pept. Sci.* 3, 561-586.
- [7] Xu, L., and Sun, H. (2010) Pharmacological Manipulation of Brain Glycogenolysis as a Therapeutic Approach to Cerebral Ischemia, *Mini-Rev. Med. Chem.* 10, 1188-1193.
- [8] Zois, C. E., and Harris, A. L. (2016) Glycogen metabolism has a key role in the cancer microenvironment and provides new targets for cancer therapy, *J. Mol. Med.* 94, 137-154.
- [9] Somsák, L., Czifrák, K., Tóth, M., Bokor, É., Chrysina, E. D., Alexacou, K. M., Hayes, J. M., Tiraidis, C., Lazoura, E., Leonidas, D. D., Zographos, S. E., and Oikonomakos, N. G. (2008) New inhibitors of glycogen phosphorylase as potential antidiabetic agents, *Curr. Med. Chem.* 15, 2933-2983.
- [10] Hayes, J. M., Kantsadi, A. L., and Leonidas, D. D. (2014) Natural products and their derivatives as inhibitors of glycogen phosphorylase: potential treatment for type 2 diabetes, *Phytochem. Rev.* 13, 471-498.
- [11] Stravodimos, G. A., Chetter, B. A., Kyriakis, E., Kantsadi, A. L., Chatzileontiadou, D. S., Skamnaki, V. T., Kato, A., Hayes, J. M., and Leonidas, D. D. (2017) Phytogetic Polyphenols as Glycogen Phosphorylase Inhibitors: The Potential of Triterpenes and Flavonoids for Glycaemic Control in Type 2 Diabetes, *Curr. Med. Chem.* 24, 384-403.
- [12] Somsák, L. (2011) Glucose derived inhibitors of glycogen phosphorylase, *C. R. Chim.* 14, 211-223.
- [13] Hayes, J. M. (2016) Computer-Aided Discovery of Glycogen Phosphorylase Inhibitors Exploiting Natural Products, In *Discovery and Development of Antidiabetic Agents from Natural Products: Natural Product Drug Discovery* (Brahmachari, G., Ed.), pp 29-62, Elsevier.
- [14] Bokor, É., Kun, S., Goyard, D., Tóth, M., Praly, J. P., Vidal, S., and Somsák, L. (2017) C-Glycopyranosyl Arenes and Hetarenes: Synthetic Methods and Bioactivity Focused on Antidiabetic Potential, *Chem. Rev.* 117, 1687-1764.
- [15] Kantsadi, A. L., Hayes, J. M., Manta, S., Skamnaki, V. T., Kiritsis, C., Psarra, A.-M. G., Koutsogiannis, Z., Dimopoulou, A., Theofanous, S., Nikoleousakos, N., Zoumpoulakis, P., Kontou, M., Papadopoulos, G., Zographos, S. E., Komiotis, D., and Leonidas, D. D. (2012) The  $\sigma$ -Hole Phenomenon of Halogen Atoms Forms the Structural Basis of the Strong Inhibitory Potency of C5 Halogen Substituted Glucopyranosyl Nucleosides towards Glycogen Phosphorylase b, *ChemMedChem* 7, 722-732.
- [16] Parmenopoulou, V., Kantsadi, A. L., Tsirkone, V. G., Chatzileontiadou, D. S. M., Manta, S., Zographos, S. E., Molfeta, C., Archontis, G., Agius, L., Hayes, J. M., Leonidas, D. D., and Komiotis, D. (2014) Structure based inhibitor design targeting glycogen phosphorylase b. Virtual screening, synthesis, biochemical and biological assessment of novel *N*-acyl- $\beta$ -D-glucopyranosylamines, *Bioorg. Med. Chem.* 22, 4810-4825.
- [17] Docsa, T., Czifrák, K., Hüse, C., Somsák, L., and Gergely, P. (2011) The effect of glucopyranosylidene-spiro-thiohydantoin on the glycogen metabolism in liver tissues of streptozotocin-induced and obese diabetic rats, *Mol. Med. Rep.* 4, 477-481.
- [18] Goyard, D., Kónya, B., Chajistamatiou, A. S., Chrysina, E. D., Leroy, J., Balzarín, S., Tournier, M., Tósch, D., Petit, P., Duret, C., Maurel, P., Somsák, L., Docsa, T., Gergely, P., Praly, J. P., Azay-Milhou, J., and Vidal, S. (2016) Glucose-derived spiro-isoxazolines are anti-hyperglycemic agents against type 2 diabetes through glycogen phosphorylase inhibition, *Eur. J. Med. Chem.* 108, 444-454.
- [19] Nagy, L., Docsa, T., Brunyánszki, A., Szántó, M., Hegedűs, C., Márton, J., Kónya, B., Virág, L., Somsák, L., Gergely, P., and Bai, P. (2013) Glycogen phosphorylase inhibitor *N*-(3,5-dimethyl-benzoyl)-*N'*-( $\beta$ -D-glucopyranosyl) urea improves glucose tolerance

- under normoglycemic and diabetic conditions through rearranging hepatic metabolism, *PLoS ONE* 8, e69420.
- [20] Docsa, T., Marics, B., Németh, J., Hüse, C., Somsák, L., Gergely, P., and Peitl, B. (2015) Insulin Sensitivity is Modified by a Glycogen Phosphorylase Inhibitor: Glucopyranosylidene-Spiro-Thiohydantoin in Streptozotocin-Induced Diabetic Rats, *Curr. Top. Med. Chem.* 15, 2390-2394.
- [21] Nagy, L., Márton, J., Vida, A., Kis, G., Bokor, É., Kun, S., Gonczi, M., Docsa, T., Tóth, A., Antal, M., Gergely, P., Csoka, B., Pacher, P., Somsák, L., and Bai, P. (2018) Glycogen phosphorylase inhibition improves beta cell function, *Br. J. Pharmacol.* 175, 301-319.
- [22] Hayes, J. M., and Archontis, G. (2012) MM-GB(PB)SA calculations of protein-ligand binding free energies, In *Molecular Dynamics - Studies of Synthetic and Biological Macromolecules* (Wang, L., Ed.), Intech Open.
- [23] Kun, S., Begum, J., Kyriakis, E., Stamati, E. C. V., Barkas, T. A., Szennyés, E., Bokor, É., Szabó, K. E., Stravodimos, G. A., Sipos, Á., Docsa, T., Gergely, P., Moffatt, C., Patraskaki, M. S., Kokolaki, M. C., Gkerdi, A., Skamnaki, V. T., Leonidas, D. D., Somsák, L., and Hayes, J. M. (2018) A multidisciplinary study of 3-(b-D-glucopyranosyl)-5-substituted-1,2,4-triazole derivatives as glycogen phosphorylase inhibitors: Computation, synthesis, crystallography and kinetics reveal new potent inhibitors, *Eur. J. Med. Chem.* 147, 266-278.
- [24] Small-Molecule Drug Discovery Suite 2016-3, Schrödinger, LLC, New York, NY, 2016.
- [25] Jones, G., Willett, P., Glen, R. C., Leach, A. R., and Taylor, R. (1997) Development and validation of a genetic algorithm for flexible docking, *J. Mol. Biol.* 267, 727-748.
- [26] Pearlman, D. A., and Charifson, P. S. (2001) Are free energy calculations useful in practice? A comparison with rapid scoring functions for the p38 MAP kinase protein system, *J. Med. Chem.* 44, 3417-3423.
- [27] Hanley, J. A., and Mcneil, B. J. (1982) The Meaning and Use of the Area under a Receiver Operating Characteristic (Roc) Curve, *Radiology* 143, 29-36.
- [28] Team, R. C. (2018) R: a language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.
- [29] Sokal, R. R., and Rohlf, F. J. (1995) *Biometry – The principles and practice of statistics in biological research*, 3rd ed., W. H. Freeman, New York.
- [30] Hayes, J. M., Skamnaki, V. T., Archontis, G., Lamprakis, C., Sarrou, J., Bischler, N., Skaltsounis, A. L., Zographos, S. E., and Oikonomakos, N. G. (2011) Kinetics, in silico docking, molecular dynamics, and MM-GBSA binding studies on prototype indirubins, KT5720, and staurosporine as phosphorylase kinase ATP-binding site inhibitors: The role of water molecules examined, *Proteins: Struct., Funct., Bioinf.* 79, 703-719.
- [31] Sun, H. Y., Duan, L. L., Chen, F., Liu, H., Wang, Z., Pan, P. C., Zhu, F., Zhang, J. Z. H., and Hou, T. J. (2018) Assessing the performance of MM/PBSA and MM/GBSA methods. 7. Entropy effects on the performance of end-point binding free energy calculation approaches, *Phys. Chem. Chem. Phys.* 20, 14450-14460.
- [32] Sun, H. Y., Li, Y. Y., Tian, S., Xu, L., and Hou, T. J. (2014) Assessing the performance of MM/PBSA and MM/GBSA methods. 4. Accuracies of MM/PBSA and MM/GBSA methodologies evaluated by various simulation protocols using PDBbind data set, *Phys. Chem. Chem. Phys.* 16, 16719-16729.
- [33] Bokor, É., Kun, S., Docsa, T., Gergely, P., and Somsák, L. (2015) 4(5)-Aryl-2-C-glucopyranosyl-imidazoles as New Nanomolar Glucose Analogue Inhibitors of Glycogen Phosphorylase, *ACS Med. Chem. Lett.* 6, 1215-1219.
- [34] Mirza, A., Desai, R., and Reynisson, J. (2009) Known drug space as a metric in exploring the boundaries of drug-like chemical space, *Eur. J. Med. Chem.* 44, 5006-5011.

- [35] Szennyés, E., Bokor, É., Docsa, T., Sipos, Á., and Somsák, L. (2019) Synthesis of C- $\beta$ -D-glucopyranosyl derivatives of some fused azoles for the inhibition of glycogen phosphorylase, *Carbohydr. Res.* 472, 33-41.
- [36] Szennyés, E., Bokor, É., Batta, G., Docsa, T., Gergely, P., and Somsák, L. (2016) Improved preparation of 4(5)-aryl-2-( $\beta$ -D-glucopyranosyl)-imidazoles, the most efficient glucose analogue inhibitors of glycogen phosphorylase, *RSC Adv.* 6, 94787-94794.
- [37] Szennyés, E., Bokor, É., Kiss, A., Somsák, L., and Pascal, Y. (2017) Preparation of 2,6-anhydro-3,4,5,7-tetra-O-benzyl-D-glycero-D-gulo-heptonimidamide, In *Carbohydrate Chemistry: Proven Synthetic Methods* (Vogel, C., and Murphy, P. V., Eds.), pp 323-332, CRC Press, Boca Raton.
- [38] Cer, R. Z., Mudunuri, U., Stephens, R., and Lebeda, F. J. (2009) IC<sub>50</sub>-to-K<sub>i</sub>: a web-based tool for converting IC<sub>50</sub> to K<sub>i</sub> values for inhibitors of enzyme activity and ligand binding, *Nucleic Acids Res.* 37, W441-W445.
- [39] Andersen, B., Rassov, A., Westergaard, N., and Lundgren, K. (1999) Inhibition of glycogenolysis in primary rat hepatocytes by 1,4-dideoxy-1,4-imino-arabinitol, *Biochem. J.* 342, 545-550.
- [40] Fosgerau, K., Westergaard, N., Quistorff, B., Grunnet, N., Kristiansen, M., and Lundgren, K. (2000) Kinetic and Functional Characterization of 1,4-Dideoxy-1,4-imino-D-arabinitol: A Potent Inhibitor of Glycogen Phosphorylase with Anti-hyperglycemic Effect in ob/ob Mice, *Arch. Biochem. Biophys.* 380, 274-284.
- [41] Latsis, T., Andersen, B., and Agius, L. (2002) Diverse effects of two allosteric inhibitors on the phosphorylation state of glycogen phosphorylase in hepatocytes, *Biochem. J.* 368, 309-316.
- [42] Oikonomakos, N. G., Tiraidis, C., Leonidas, D. D., Zographos, S. E., Kristiansen, M., Jessen, C. U., Nørskov-Lauritsen, L., and Agius, L. (2006) Iminosugars as potential inhibitors of glycogenolysis: Structural insights into the molecular basis of glycogen phosphorylase inhibition, *J. Med. Chem.* 49, 5687-5701.
- [43] Andersen, B., Rassov, A., Westergaard, N., and Lundgren, K. (1999) Inhibition of glycogenolysis in primary rat hepatocytes by 1,4-dideoxy-1,4-imino-d-arabinitol, *Biochem J* 342, 545-550.
- [44] Minami, Y., Kurlyarna, C., Ikeda, K., Kato, A., Takebayashi, K., Adachi, I., Fleet, G. W. J., Kettawan, A., Karnoto, T., and Asano, N. (2008) Effect of five-membered sugar mimics on mammalian glycogen-degrading enzymes and various glucosidases, *Bioorg. Med. Chem.* 16, 2734-2740.
- [45] Veber, D. F., Johnson, S. R., Cheng, H. Y., Smith, B. R., Ward, K. W., and Kopple, K. D. (2002) Molecular properties that influence the oral bioavailability of drug candidates, *J. Med. Chem.* 45, 2615-2623.
- [46] Begum, J., Varga, G., Docsa, T., Gergely, P., Hayes, J. M., Juhász, L., and Somsák, L. (2015) Computationally motivated synthesis and enzyme kinetic evaluation of N-( $\beta$ -D-glucopyranosyl)-1,2,4-triazolecarboxamides as glycogen phosphorylase inhibitors, *MedChemComm* 6, 80-89.
- [47] Polyák, M., Varga, G., Szilágyi, B., Juhász, L., Docsa, T., Gergely, P., Begum, J., Hayes, J. M., and Somsák, L. (2013) Synthesis, enzyme kinetics and computational evaluation of N-( $\beta$ -D-glucopyranosyl) oxadiazolecarboxamides as glycogen phosphorylase inhibitors, *Bioorg. Med. Chem.* 21, 5738-5747.
- [48] Hann, M. M., and Keseru, G. M. (2012) Finding the sweet spot: the role of nature and nurture in medicinal chemistry, *Nat. Rev. Drug Discovery* 11, 355-365.
- [49] Agius, L. (2015) Role of glycogen phosphorylase in liver glycogen metabolism, *Mol. Aspects Med.* 46, 34-45.

- [50] Aiston, S., Andersen, B., and Agius, L. (2003) Glucose 6-phosphate regulates hepatic glycogenolysis through inactivation of phosphorylase, *Diabetes* 52, 1333-1339.
- [51] Aiston, S., Green, A., Mukhtar, M., and Agius, L. (2004) Glucose 6-phosphate causes translocation of phosphorylase in hepatocytes and inactivates the enzyme synergistically with glucose, *Biochem. J.* 377, 195-204.