



Outliers in SAR and QSAR: 4. effects of allosteric protein–ligand interactions on the classical quantitative structure–activity relationships

Ki Hwan Kim¹

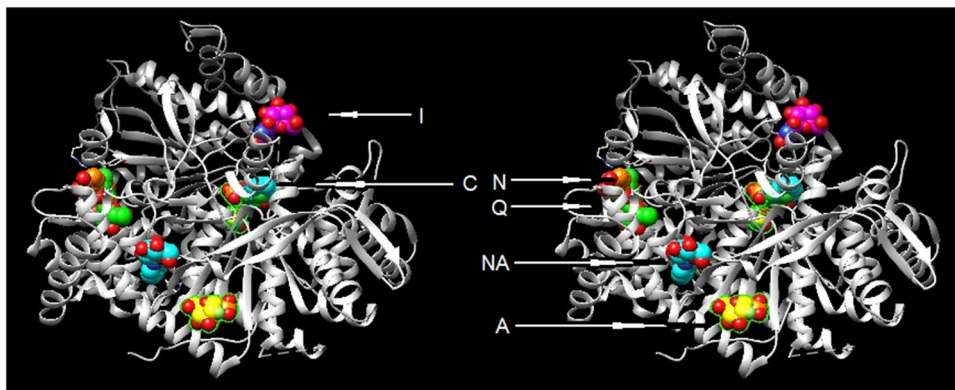
Received: 22 September 2021 / Accepted: 10 December 2021 / Published online: 22 February 2022
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022, corrected publication 2022

Abstract

Effects of allosteric interactions on the classical structure–activity relationship (SAR) and quantitative SAR (QSAR) have been investigated. Apprehending the outliers in SAR and QSAR studies can improve the quality, predictability, and use of QSAR in designing unknown compounds in drug discovery research. We explored allosteric protein–ligand interactions as a possible source of outliers in SAR/QSAR. We used glycogen phosphorylase as an example of a protein that has an allosteric site. Examination of the ligand-bound x-ray crystal structures of glycogen phosphorylase revealed that many inhibitors bound at more than one binding site. The results of QSAR analyses of the inhibitors included a QSAR that recognized an outlier bound at a distinctive allosteric binding site. The case provided an example of constructive use of QSAR identifying outliers with alternative binding modes. Other allosteric QSARs that captured our attention were the inverted parabola/bilinear QSARs. The x-ray crystal structures and the QSAR analyses indicated that the inverted parabola QSARs could be associated with the conformational changes in the allosteric interactions. Our results showed that the normal parabola, as well as the inverted parabola QSARs, can describe the allosteric interactions.

Graphical abstract

Examination of the ligand-bound X-ray crystal structures of glycogen phosphorylase revealed that many inhibitors bound at more than one binding site. The results of QSAR analyses of the inhibitors included a QSAR that recognized an outlier bound at a distinctive allosteric binding site.



Keywords Quantitative structure–activity relationships (QSAR) · X-ray crystal structure · Source of outliers · Glycogen phosphorylase · Allosteric interactions · Dual binding modes · Inverted parabola/bilinear QSAR

✉ Ki Hwan Kim
pkhkim@gmail.com

¹ San Marcos, CA, USA

Introduction

Structure–activity relationship (SAR) and quantitative structure–activity relationship (QSAR) studies play a significant role in drug discovery and development. SAR/QSAR developments often yield observations of outliers. A decade after our initial attention to the outliers in SAR and QSAR [1], many were detected in developing QSARs [2], as noted in previous literature and the C-QSAR database [3]. We suggested that outliers in SAR and QSAR could result from the distinctive binding modes or flexibilities of the binding site even if the compounds involved were structural congeners [1, 4]. We also discussed the importance of considering the role of water molecules in protein–ligand interactions and QSAR studies [2]. In these studies, “outliers” referred to those compounds that possessed the unexpected biological activity. They were unable to fit in a derived QSAR model, as described by Verma and Hansch [5].

In the course of these studies, we searched the C-QSAR database with the query ‘carbonic anhydrase inhibitor,’ ‘elastase inhibitor,’ and ‘rhinovirus inhibitor.’ Among the 260 equations retrieved from the searches, 19 QSAR equations flagged our attention (Table S1, Supplementary Material). These 19 QSARs showed an inverted parabolic or bilinear relationship with ClogP or CMR. The majority of the QSARs in the C-QSAR database had a normal parabolic or bilinear relationship. Such inverted parabolic or bilinear QSARs were deemed atypical correlations.

Inverted parabolic or bilinear equations with ClogP, CMR (or other such) indicated that the biological activity initially decreased as ClogP or CMR increased. At the inversion point, however, the biological activity stopped decreasing and began increasing as ClogP or CMR increased further. It was suggested in the literature that such inverted relationships were due to allosteric interactions involving conformational change [6–8].

In this study, we examined whether such inverted parabolic or bilinear correlations were associated with the allosteric interactions. In addition, we examined whether allosteric interactions can be a source of outliers in some SAR and QSARs. Only the inverted parabola/bilinear QSARs were suggested as “allosteric QSARs” by Verma and Hansch [5–10]. We investigated whether linear and normal parabola/bilinear QSARs could equally describe allosteric interactions.

Material and method

RCSB protein data bank searching

The RCSB protein data bank [11, 12] was searched with the query ‘glycogen phosphorylase.’ The multiple sequence alignments for protein structure comparison were completed utilizing the Clustal Omega program of the EMBL-EBI [13] and/or the UCSF Chimera molecular modeling program (version 1.14) [14].

Molecular graphics

All the figures were generated using the UCSF Chimera molecular modeling program using the multiple sequence alignments obtained from the Clustal Omega of EMBL-EBI described above or the structure comparison tool of Chimera. All the figures included in this paper were generated from the corresponding X-ray crystal structures after multiple sequence alignments.

Multiple regression analysis for QSAR derivatization

The QSAR equations presented in this paper were executed using the C-QSAR program of Biobyte [15]. Most of the physicochemical parameters and structural descriptors were auto-loaded utilizing the C-QSAR program. CPI was the calculated hydrophobic parameter for the substituents. CMR was the calculated molar refractivity for the molecule or substituents and depended on the volume and polarizability. MgVol was the molar volume for the molecule. The indicator variable was also assigned the value of one or zero for special features with special effects that could not be parameterized. Further details for the indicator variable used were explained whenever it was used. Each regression equation included 95% confidence limits for each terms in parentheses.

In these QSAR equations, n was the number of data points, r^2 was the squared correlation coefficient that showed the goodness of fit, while q^2 was the goodness of leave-one-out prediction. Finally, s was the standard deviation.

In this paper, all the QSAR equations except Eqs. 4g and 4h were developed from only those compounds whose ligand–protein X-ray crystal structures were available in the RCSB PDB protein data bank.

Results and discussion

Different binding sites of ligands in a protein

In SAR and/or QSAR, the general assumption was that all the structural congeners bind at the same binding site in an essentially identical binding mode [16]. In our previous studies, it was demonstrated that even if some compounds were structural congeners, their binding modes could be different and thus showed up as outliers in SAR/QSAR, despite the fact that their binding sites were the same [1].

Subsequently, it would be reasonable to expect some congeneric compounds not to fit the same SAR/QSAR when they bound at a different binding site. Such situations would likewise have included compounds binding in orthosteric and allosteric binding sites of the protein.

After searching the ligand-bound protein structures in the RCSB protein data bank, we observed numerous examples that showed even very close structural analogs had bound at different binding sites. Such X-ray crystal structures provided clues for possible sources of outliers in SAR/QSARs. These findings are summarized below.

We chose the glycogen phosphorylase enzyme as an example for a thorough examination. A list of an extensive number of known allosteric receptors or enzymes are available (<http://mdl.shsmu.edu.cn/ASD>) [17].

Inverted parabola/bilinear correlations and allosteric interactions

In addition to the multiple binding sites, another aspect to be considered regarding the allosteric modulation of the protein is the inverted parabolic/bilinear allosteric QSAR correlations.

Besides those equations listed in Table S1, a number of additional inverted parabolic or bilinear QSAR have been reported [3]. Hansch and his co-workers [5] suggested the rationale behind the inverted parabola/bilinear relationships from allosteric interactions. Works on the allosteric interactions appeared in the literature as early as 1958 [7]. However, QSAR on the allosteric interactions began in 2001. Since then, Hansch's group published several QSAR papers on the allosteric interactions [6, 18–20]. They initially observed that some QSARs correlated the data by an inverted parabolic relationship with ClogP, CMR, and molar volume (MgVol) [8]. Such inverted parabolic relationships showed that the activity at first decreased as the value of the related parameter increased. However, at a specific point, it turned around and increased. They attributed such behavior to a change in the structure of the receptor that occurred with ligand binding as in the allosteric interactions. They suggested that a change in the reaction mechanism occurred at the inversion point [7].

Furthermore, Hansch and co-workers [5–8, 10, 18, 21] proposed allosteric QSARs could be used to uncover an allosteric interaction. The classic means for uncovering allosteric reactions was to carefully evaluate a particular molecule at a time enzymatically, and eventually use X-ray crystallography to confirm it. Allosteric QSAR correlations can be illustrated by Eqs. i–iv for Clog P or CMR [5].

$$\log(1/C) = -a \text{Clog P} + b \text{Clog P}^2 + \text{constant} \quad (\text{i})$$

$$\log(1/C) = -a \text{Clog P} + b \log(\beta \times 10^{\text{Clog P}} + 1) + \text{constant} \quad (\text{ii})$$

$$\log(1/C) = -a \text{CMR} + b \text{CMR}^2 + \text{constant} \quad (\text{iii})$$

$$\log(1/C) = -a \text{CMR} + b \log(\beta \times 10^{\text{CMR}} + 1) + \text{constant} \quad (\text{iv})$$

These equations were inverted parabola or bilinear correlations. They implied that as Clog P or CMR increased, the activity decreased. However, at the inversion point, the exponential term took over and the activity increased with further increase in Clog P or CMR value. Similarly, in the bilinear model, activity first decreased linearly up to the inversion point and then increased linearly [7]. They proposed another way of explaining the inverted parabola/bilinear correlations which suggested that there could be an additional binding site. As molecules became larger in ClogP or CMR, they were limited in binding to the 'typical' site, and forced to bind in the secondary site [7]. Hansch and others reported several QSARs regarding such possible allosteric interactions [5–10].

One does not typically recognize exactly what the receptor structure is in a cell, much less a whole animal. Nonetheless, it was suggested that QSAR can serve as a valuable tool in gaining an indirect view of what one might learn about its in situ properties [7].

Glycogen phosphorylase and allosteric inhibitors

Glycogen phosphorylase (GP) comprises a family of three isozymes: muscle GP (mGP), liver GP (IGP), and brain GP (bGP) [22]. GP is an allosteric enzyme that catalyzes the first step of glycogenolysis in the liver, muscle, and brain to produce glucose-1-phosphate (G-1-P) from glycogen.

GP is a homodimer that exists in two interconvertible forms, GPb and GPa. GPb is the 'closed' nonphosphorylated form, low activity, low substrate affinity, and predominantly T-state in equilibrium. GPa is the 'open' phosphorylated form, high activity, high substrate affinity, and predominantly R-state in equilibrium [23–25]. Phosphorylation

at Ser14 at the N-terminus converts the enzyme from the T-state to the R-state. The allosteric transition (T- to R-state) of GP is affected by allosteric modulators that bind to GP and stabilize or promote either T-state or R-state conformation [24]. In the “closed” T-state conformation, the active site is blocked, preventing the entrance of the substrate. In the “open” R-state conformation, the catalytic site becomes accessible to the substrate. Potent GP inhibitors stabilize the inactive T state conformation [26]. Allosteric inhibitors can alter the equilibrium between T- and R-state [23].

Allosteric inhibitors interact with binding sites on the enzyme that are distinct from the binding site (the orthosteric site) for the endogenous agonist [27]. Allosteric sites allow inhibitors to bind to the enzyme and often result in a conformational change. Verma and Hansch suggested that allosteric effects occur when the interaction between protein and ligand results in a structural change of the protein [5].

X-ray crystal structures of ligand-bound glycogen phosphorylase

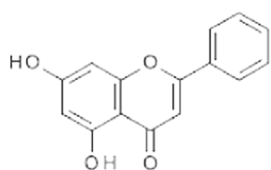
There are over 200 glycogen phosphorylase crystal structures reported in the PDB database (Table S3). They are from the organisms of human, rabbit, and Baker’s yeast,

along with the muscle, liver, and brain forms [17]. Most of these structures are utilized in the “allosteric QSARs” presented below.

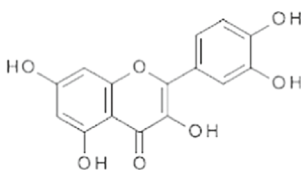
Allosteric inhibitors, their different binding modes and their effects on SAR/QSAR

The general assumption in SAR/QSAR is that all the structurally similar compounds, especially structural congeners, bind in a similar binding mode at the same binding site. Therefore, if some compounds bind at other binding sites, it would be reasonable to expect that they do not fit to the same SAR/QSAR, thus becoming outliers.

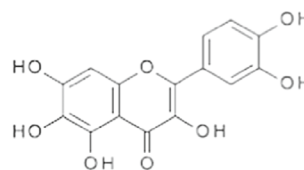
Such various binding modes are exactly the case of the GP inhibitor flavonoid derivatives that Chetter et al. reported [28]. Flavonoid analogs chrysin, quercetin, and quercetagenin are structurally similar. Ordinarily, these compounds would be considered as congeners and included in the same set of data for QSAR development. However, the binding sites of these compounds at GP were reported to be completely different: the inhibitor binding site for chrysin, the quercetin binding site for quercetin, and the allosteric binding site for quercetagenin (Please see Table 6 and Fig. 6 and relevant discussion below).



Chrysin (PDB ID: 3EBO)
Binding site: inhibitor site



Quercetin (PDB ID: 4MRA)
Binding site: quercetin



Quercetagenin
Binding site: allosteric

We investigated whether such differences in the binding site of the allosteric inhibitors among the structural congeneric series are common. For this, we used the X-ray crystal structures of the inhibitor-bound GP complexes listed in Table S2 (Supplementary Material). We also examined whether any unusual allosteric binding of ligands can yield outliers in QSAR. The results are summarized here.

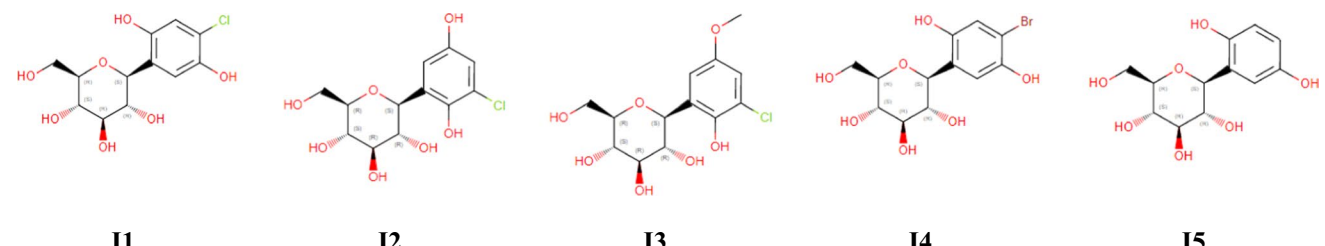
(C-β-D-glucopyranosyl)-hydroquinone derivatives

A substantial number of glucose derivatives have been shown to inhibit GP. Most often these compounds bound to the catalytic site of the enzyme. Three main groups of glucose derivatives that exhibited potent inhibitory activities

were C-glucopyranosyl heterocycles, N-acyl-N'-glucopyranosyl urea, and glucopyranosylidene-spiro-heterocycles [29].

Alexacou et al. [30] reported the inhibitory potencies of glucopyranosyl-hydroquinone regioisomers (**I1–I4**) listed in Table 1. These compounds were competitive inhibitors of GPb with respect to α-D-glucose-1-phosphate (Glc-1-P). The X-ray crystal structures of these compounds revealed that they bound at the catalytic site (Fig. 1) and stabilize the T conformation of the enzyme. The X-ray structure of similar compound **I5**, that He et al. [31] reported, showed it was also bound to the catalytic site.

Alexacou et al. [30] described that **I1** bound to the novel allosteric binding site as well as the catalytic site in the GPb

Table 1 (C- β -D-glucopyranosyl)-hydroquinone regioisomers and their X-ray crystal structure information


Compd	PDB	Ligand	Ki (nM)	pKi			CPI ^b	Binding Site ^c	Ref.
				Obs ^a	Cal	Dev			
I1	3S0J	Z15	169,900	3.77	3.81	-0.04	1.17	C	[30]
	3NP7	Z15	(140,000) ^d					C, N	[30]
I2	3NP7	Z16	95,000 (140,000) ^d	4.02	3.81	0.21	1.17	C	[30]
I3	3NP9	Z2T	39,800	4.40	4.40	0.00	1.77	C	[30]
I4	3NPA	Z57	136,400	3.87	4.01	-0.14	1.37	C	[30]
I5	2FET	H53	977,000	3.01	3.05	-0.04	0.40	C	[31]

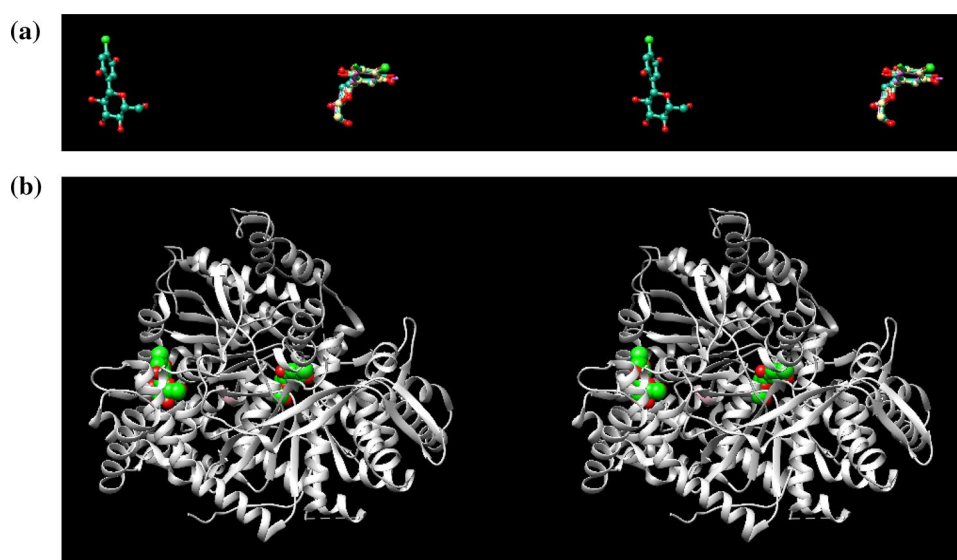
^aCalculated using Eq. 1

^bThe CPI values were auto-loaded utilizing the C-QSAR program

^cBinding site: C = catalytic site, N = novel allosteric binding site

^dKi values are for the mixture of **I1** and **I2**

Fig. 1 a Binding mode of **I1** (cyan) at the catalytic site (right) and the novel allosteric binding site (left). Other analogs are also shown at the catalytic site. **b** Location of the two binding sites of **I1**: the catalytic site (C, right) and the novel allosteric binding site (N, left)



complex structure. On the other hand, **I3–I5** did not bind at this allosteric binding site.

Compound **I1** bound at the new allosteric binding site only when GPb crystals were soaked with a mixture of **I1** and **I2** and not when soaked with **I1** alone. The two experimental conditions were similar when soaking GPb native crystals with either a solution of a mixture of compounds **I1** and **I2** (100 mM, 21 h) or a solution of **I3** (70 mM, 20 h).

They were unsure whether this new allosteric binding site represented a genuine new binding site with a regulatory function or if it was an artifact of the experimental conditions. Nonetheless, Alexacou et al. stressed the new allosteric binding site displayed some specificity toward **I1**, since only **I1** (which is the weaker inhibitor) bound to this site from the mixture of compounds **I1** and **I2** [30].

Figure 1 is a stereo-pair picture of the compound **II** bound to the catalytic site as well as to the novel allosteric binding site. The crystal structures of the ligand-GP complexes showed that the inhibitors were accommodated at the catalytic site without any significant conformational change of the protein structure.

Even though surrounded by such a complex protein environment, the inhibitory potency expressed as pKi of these compounds significantly correlated with the hydrophobic parameter CPI of the substituent (Eq. 1). Equation 1 suggests that when the hydrophobicity of the substituents increases, the inhibitory potency will increase as well. Table 1 lists the calculated pKi values using Eq. 1.

$$\text{pKi} = 0.99(\pm 0.47) \text{CPI} + 2.65(\pm 0.60)$$

$$n = 5, r^2 = 0.94, q^2 = 0.86, s = 0.149 \quad (1)$$

Because of the limited number of compounds involved in this case, it was not possible to examine any two-parameter equations. Since the coefficient of CPI is positive, it would likely become a normal parabola/bilinear correlation rather than an inverted correlation, even if CPI could be

extended. Even so, Eq. 1 accounts for 94% of the variance in these inhibitory potency data. No additional parameter was required to explain the observed behavior of **II** binding at the novel allosteric binding site. (Please see further discussions below.) Because **II** bound at the novel allosteric binding site only when GPb crystals were soaked with a mixture of **II** and **I2**, the result was not unexpected.

Phosphorylated glucose derivatives

Martin et al. [32] reported several glucose analogs binding to the catalytic site of T-state GPb: a T-state-stabilizing inhibitor α -D-glucose (**II1**; synergistic with binding of AMP, IMP, and caffeine to the inhibitor site), R-state-stabilizing phosphorylated ligands α -D-glucose 1-phosphate (**II2**), 2-deoxy-2-fluoro- α -D-glucose 1-phosphate (**II3**), and α -D-glucose 1-methylenephosphonate (**II4**). They are listed in Table 2.

Martin et al. described that the phosphorylated ligands **II2**, **II3**, and **II4** were bound at the allosteric activator (AMP) site (A site, also sometimes called N site in the literature) in addition to the catalytic site. The binding of the

Table 2 Binding sites of α -D-glucose (**II1**, GLC) and its structural analogs **II2–II6** and their X-ray crystal structure information

Compound	Structure	PDB	Ligand	Ki (nM)	pKi	Binding Site ^a	Ref.
II1		2GPB	GLC	2,000,000	2.70	C	[32]
II2		3GPB	G1P	9,500,000	2.02	C, A	[32]
II3		4GPB	GFP	–	–	C, A	[32]
II4		5GPB	GPM	–	–	C, A	[32]
II5		6GPB	H2P	14,000	4.84	C	[33]
II6		1XC7	GL6	5,900,000	2.23	C	[34]

^aBinding Site: C=catalytic site, A=allosteric (AMP) binding site

phosphorylated inhibitors was accompanied by the movement of catalytic site residues, especially a shift of a loop out of the catalytic site toward the exterior of the enzyme.

Table 2 includes heptulose 2-phosphate (**II5**) reported by Johnson et al. for comparison [33].

Even though all the compounds in Table 2 are structurally similar to glucose, their binding modes are different as seen in their crystal structures. **II1** and **II2** are bound at the catalytic site, whereas the other three analogs (**II2–II4**) are bound at two separate binding sites: the catalytic site and the allosteric activator (AMP) site. In each binding site, their binding conformations are essentially identical as seen in Fig. 2a. Furthermore, Fig. 2b shows the location of these two binding sites in GPb. Compound **II6** (phosphoramidate) reported by Chrysina et al. [34] also binds at the catalytic site.

Compounds **III1–III6** in Table 2 provide another example that shows structurally close analogs bind at different binding sites. When studying SAR/QSAR, one should carefully consider their binding site as well as their binding mode. Because of their structural diversities and lack of p*K*_i values of some compounds, no QSAR was developed from these compounds.

Spiro-glucose derivatives

Table 3 lists 17 spiro-glucose analogs with their X-ray crystal structures information. Bentlifa et al. [35] reported

III1–III5 as inhibitors of rmGPb and showed that the inhibitors bound preferentially at the catalytic site of the enzyme retaining the less active T-state conformation. Watson et al. [36] also reported **III6–III9** and described that they bound at the catalytic site. Czifrak et al. [37], Oikonomaos et al. [38], and Gregoriou et al. [39] described **III10–III14** for their inhibition of GPb. They reported that all five compounds bound at the catalytic site of T-state GPb with very little change in the tertiary structure. Szabo et al. [29] reported the inhibitory activities of **III15–III17** and their binding modes. **III15** was unique in this series because the compound was bound at two binding sites: the catalytic site (C) and the new allosteric (indole) binding site (NA).

The entire compounds listed in Table 3 did not yield any statistically sound QSAR. Since the GP is an allosteric enzyme, the ‘splitting QSAR’ approach suggested by Verma and Hansch [5] was utilized to develop Eqs. 2a–2d. The *K*_i value of one compound (**III9**) was not available and omitted. From the first set of nine compounds, Eq. 2a was derived, which is a normal parabola correlation with MolVol. A statistically slightly inferior correlation was obtained with CMR (Eq. 2b). The results were not surprising because there was high collinearity between CMR and MolVol with this set of compounds (Eq. 2c). The remaining seven outlier compounds of Eq. 2a yielded Eq. 2d. Interestingly, Eq. 2d is an inverted parabola correlation with CMR. Compound

Fig. 2 **a** Binding modes of **II2–II4** at the catalytic site (top) and the allosteric activator (AMP) site (bottom). **b** Location of the two binding sites of **II2** (G1P, PDB ID: 3GPB): the catalytic site (C, top) and the allosteric activator site (A, bottom)

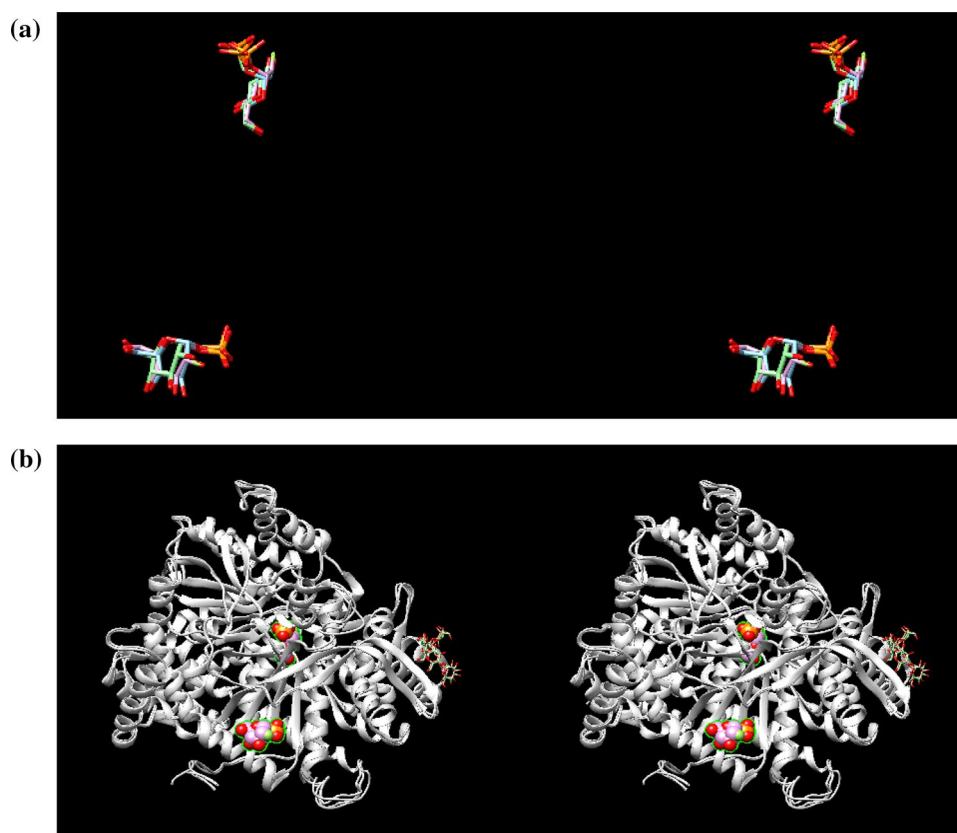


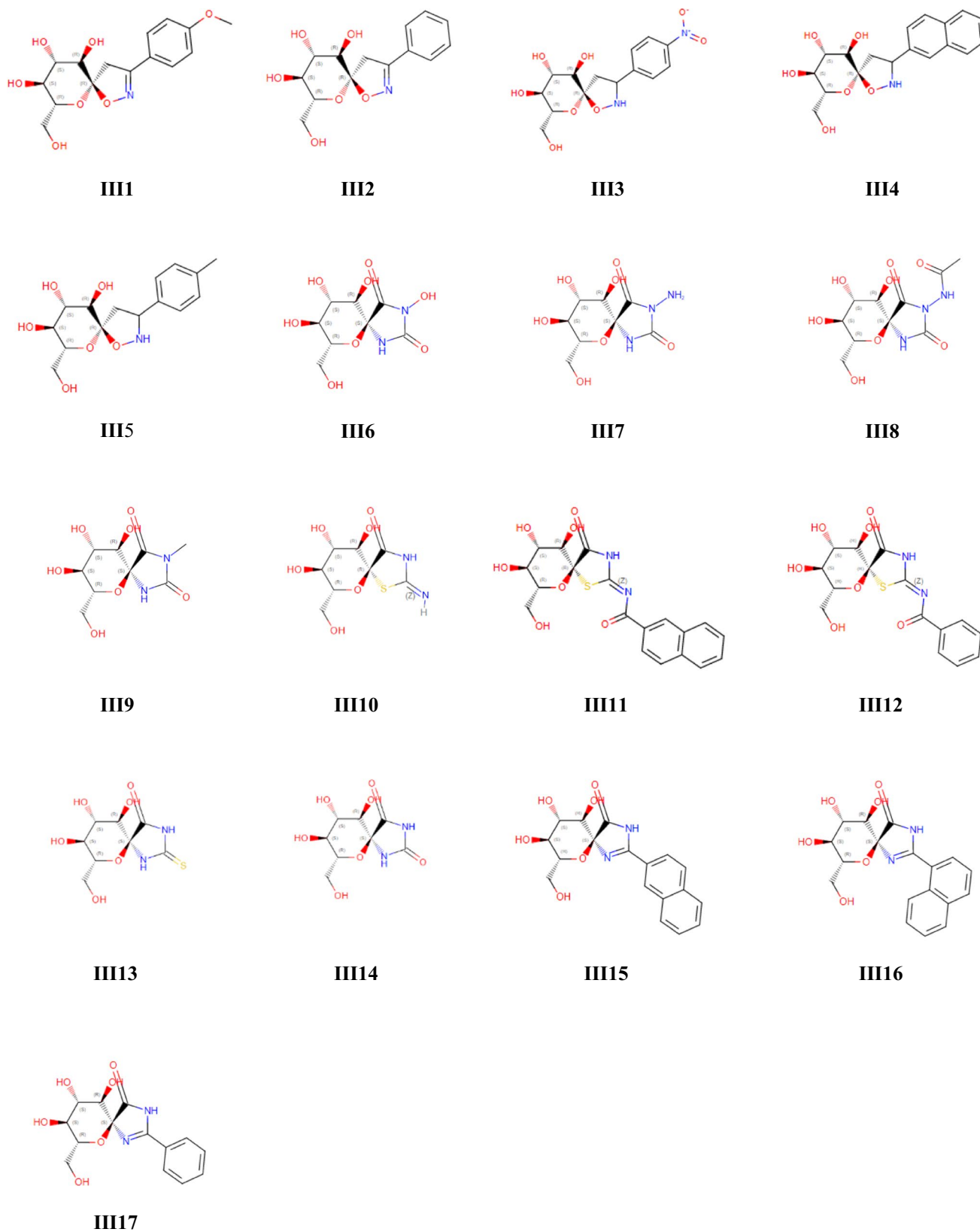
Table 3 Spiro-glucose analogs (**III1–III17**) and their X-ray crystal structure information

Table 3 (continued)

Compd	PDB	Ligand	Ki (nM)	pKi			ClogP ^b	CMR ^b	MolVol ^b	Bind Site ^c	Ref.
				Obs	Cal ^a	Dev					
III1 ^d	2QRG	M07	6600	5.18	5.03	0.15	0.54	8.08	5.02	C	[35]
III2 ^d	2QRH	M08	19,600	4.71	4.93	−0.22	0.29	7.46	4.16	C	[35]
III3 ^e	2QRM	M09	92,500	4.03	3.83	0.20	0.33	8.07	4.88	C	[35]
III4 ^e	2QRP	S06	630	6.20	5.82	0.38	1.46	9.15	5.76	C	[35]
III5 ^d	2QRQ	S13	7900	5.10	5.00	0.10	0.78	7.92	4.75	C	[35]
III6 ^d	1FTW	GL5	38,905	4.41	4.47	−0.06	−2.63	5.26	2.56	C	[36]
III7 ^e	1FTQ	GL2	144,544	3.84	4.37	−0.53	−2.54	5.48	2.69	C	[36]
III8 ^e	1FU4	GL9	549,541	3.26	3.30	−0.04	−2.03	6.44	3.76	C	[36]
III9 ^d	1FTY	GL7	–	–	–	–	−1.62	5.57	2.82	C	[36]
III10 ^d	4CTM	MIF	24,000	4.62	4.54	0.08	−3.23	5.70	2.72	C	[37]
III11 ^d	4CTN	M8P	10,000	5.00	4.95	0.05	0.86	10.55	7.73	C	[37]
III12 ^d	4CTO	M7C	9000	5.05	5.06	−0.01	−0.31	8.86	5.81	C	[37]
III13 ^f	1HLF	GL4	5100	5.29	4.54	0.75	−2.39	5.96	2.72	C	[38]
III14 ^e	1GGN	GLS	3100	5.51	5.07	0.44	−2.40	5.11	2.37	C	[38]
	1A8I	GLS								C	[39]
III15 ^e	6QA6	HT8	2110	5.68	6.13	−0.45	0.21	9.27	6.25	C, NA ^g	[29]
III16 ^d	6QA7	HTW	13,000	4.89	5.05	−0.16	0.21	9.27	6.25	C	[29]
III17 ^d	6QA8	HTE	9000	5.05	4.99	0.06	−0.96	7.58	4.58	C	[29]

^aCalculated using Eq. 2a

^bThe ClogP values were calculated utilizing the BioLoom program. The CMR and MolVol values were auto-loaded from the C-QSAR program

^cBinding Site: C = catalytic site, NA = new allosteric (indole) binding site

^dUsed to derive Eq. 2d

^eUsed to derive Eq. 2a

^fTreated as outliers when Eq. 2b was developed

^gSzabo reported that the S epimer of **III15** did not bind at the active site but binds at the new allosteric site

III13 was treated as a final outlier. Equation 2d shows that the inhibitory potency of these compounds first decreases with an increase in molar refractivity (CMR) up to the inversion point for CMR = 6.96 and then increases. Equations 2a and 2d explain 77% and 88% of the variance in the inhibitory activity data of the molecules, respectively.

$$\begin{aligned} \text{pKi} &= 4.15(\pm 3.55) \text{ MolVol} - 0.85(\pm 0.82) \text{ MolVol}^2 \\ &+ 0.01(\pm 3.76) \\ n &= 9, r^2 = 0.77, q^2 = 0.45, s = 0.143 \quad (2a) \\ \text{outlier: } &\mathbf{III13}, \mathbf{III14}, \mathbf{III17}, \mathbf{III18}, \mathbf{III14}, \mathbf{III15} \\ \text{optimum MolVol: } &2.43(\pm 4.04). \end{aligned}$$

$$\begin{aligned} \text{pKi} &= 0.79(\pm 0.68) \text{ CMR} - 0.04(\pm 0.04) \text{ CMR}^2 \\ &+ 1.51(\pm 2.58) \\ n &= 9, r^2 = 0.75, q^2 = 0.42, s = 0.148 \quad (2b) \\ \text{outlier: } &\mathbf{III3}, \mathbf{III4}, \mathbf{III7}, \mathbf{III8}, \mathbf{III13}, \mathbf{III14}, \mathbf{III15} \\ \text{optimum CMR} &= 9.88(\pm 60.60). \end{aligned}$$

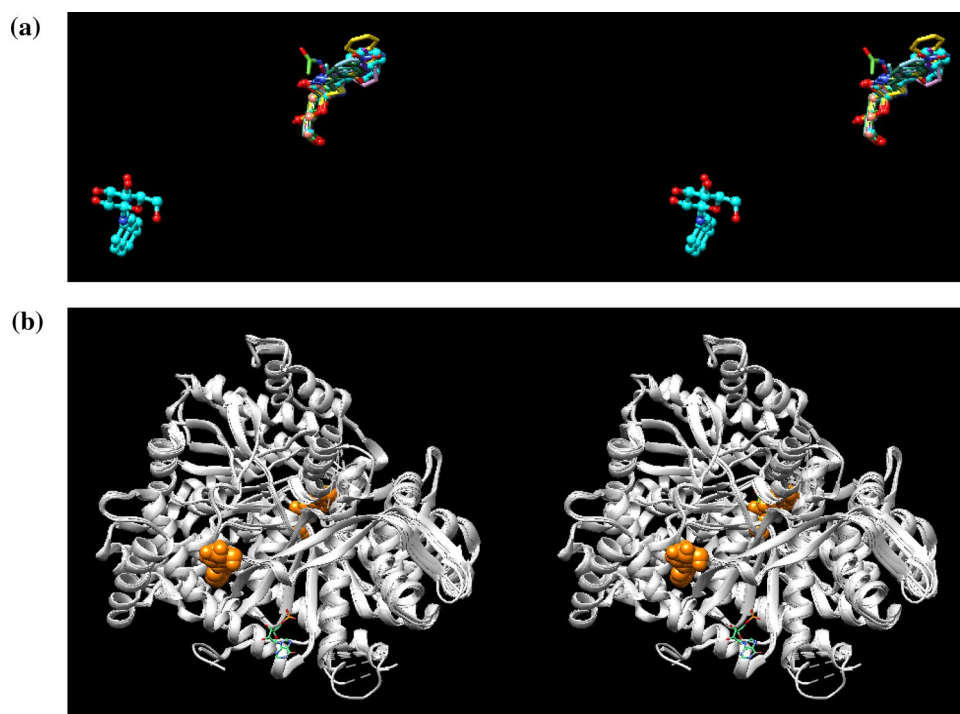
$$\begin{aligned} \text{CMR} &= 4.40(\pm 0.22) \text{ MolVol} - 1.70(\pm 0.48) \\ n &= 10, r^2 = 1.00, q^2 = 0.99, s = 0.113 \quad (2c) \end{aligned}$$

$$\begin{aligned} \text{pKi} &= -7.80(\pm 6.47) \text{ CMR} + 0.56(\pm 0.45) \text{ CMR}^2 \\ &+ 30.30(\pm 22.35) \\ n &= 6, r^2 = 0.88, q^2 = 0.46, s = 0.537 \quad (2d) \\ \text{outlier: } &\mathbf{III13} \\ \text{inversion point for CMR: } &6.96(\pm 0.86). \end{aligned}$$

Among the 17 compounds listed in Table 3, only **III15** was reported to bind at the two binding sites. Equation 2d includes **III15**, and no other parameter was required to account for any effects due to the dual binding of **III15**. Figure 3 shows the binding modes and binding sites of **III1–III17**. The crystal structures show that all compounds are bound only to the catalytic site except **III15**. Compound **III15** is bound to the new allosteric binding site as well as the catalytic site.

Szabo et al. [29] reported that the crystal structures showed only the R epimers of **III16** and **III17** bound preferentially at the catalytic site. The R epimer of **III15** was

Fig. 3 **a** Binding modes of 17 compounds (**III1–III17**) including **III15** bound at two separate binding sites: the catalytic site (right) and the new allosteric (indole) binding site (left). **b** Location of the binding sites of **III15** at the catalytic site (C, right) and the new allosteric binding site (NA, left)



bound at both the catalytic and the new allosteric (indole) binding sites. They suggested the catalytic site was the primary binding site for this inhibitor, and the new allosteric (indole) binding site (NA) was the secondary binding site. On the other hand, the S epimer of **III15** did not bind at the catalytic site but bound at the new allosteric (indole) binding site. Other glucose-derived inhibitors discussed later were also bound at this new allosteric site [29].

Upon binding at the new allosteric site, **III15** participated in five hydrogen bond interactions with several protein residues. Besides that, the imidazolinone ring formed a hydrogen bond with the side chain of the enzyme. The binding of **III15** at the new allosteric site triggered a shift of the side chain of Arg60 by about 3.0 Å. This shift caused a small translocation of the helix (residues 60–64) [29].

Equation 2a (and 2b) is a normal parabola correlation, and 2d is an inverted parabola correlation. The results suggest that there are (at least) two different modes of interactions among these analogs affecting the inhibitory potencies expressed as pKi even though all bind at the catalytic site. Relatively low r^2 values of Eqs. 2a (or 2b) and 2d indicate other factors that have not yet been accounted for. However, a smaller number of compounds available for analyses especially Eq. 2d prevented further investigation. **III13** is shown to be an outlier in Eq. 2d. This is the only thiohydantoin compound in Table 3. The thio atom of **III13** interacts with the carboxyl oxygen atom of Asp339 residue through a water molecule (W1009). Such hydrogen bonding interaction with a thiocarbonyl group is not present in other analogs. There is a corresponding carbonyl derivative (**III6**)

used in Eq. 2d. The X-ray crystal structure of **III6** (1FTW) lacks such hydrogen bonding with Asp339. Furthermore, a thiocarbonyl group is generally different and more basic than the corresponding carbonyl group [40]. The observed pKi value of **III13** is more potent than the calculated value from Eq. 2d by 0.75, which is the largest deviation in this series. This deviation could be due to the effects of hydrogen bonding interactions that are not accounted for in Eq. 2d.

No additional parameter was needed to account for the binding at two different binding sites of **III15**. The results indicate such effects are minor in the current situation.

Glucopyranosyl nucleoside derivatives

Numerous researchers reported the inhibitory potencies and their X-ray crystal structures of several D-glucopyranosyl nucleoside analogs (**IV1–IV22**) bound to GPb (Table 4) [41–46]. Three structurally similar furanosyl analogs are additionally included (**IV23–IV25**) in the table for comparison [46]. The crystal structures demonstrated that most of these inhibitors were competitive inhibitors (with the substrate Glc-1-P) and preferentially bound at the catalytic site which promoted the less active T state conformation of the enzyme.

No formal paper has been published about **IV19–IV25**, but the crystal structures of these structures revealed fascinating information regarding their binding sites. There are two sub-groups of these structures including **IV18**: five 6-membered pyranosyl compounds (**IV18–IV22**) and three 5-membered furanosyl compounds (**IV23–IV25**). Most of the inhibitors are bound to the catalytic site, but there are

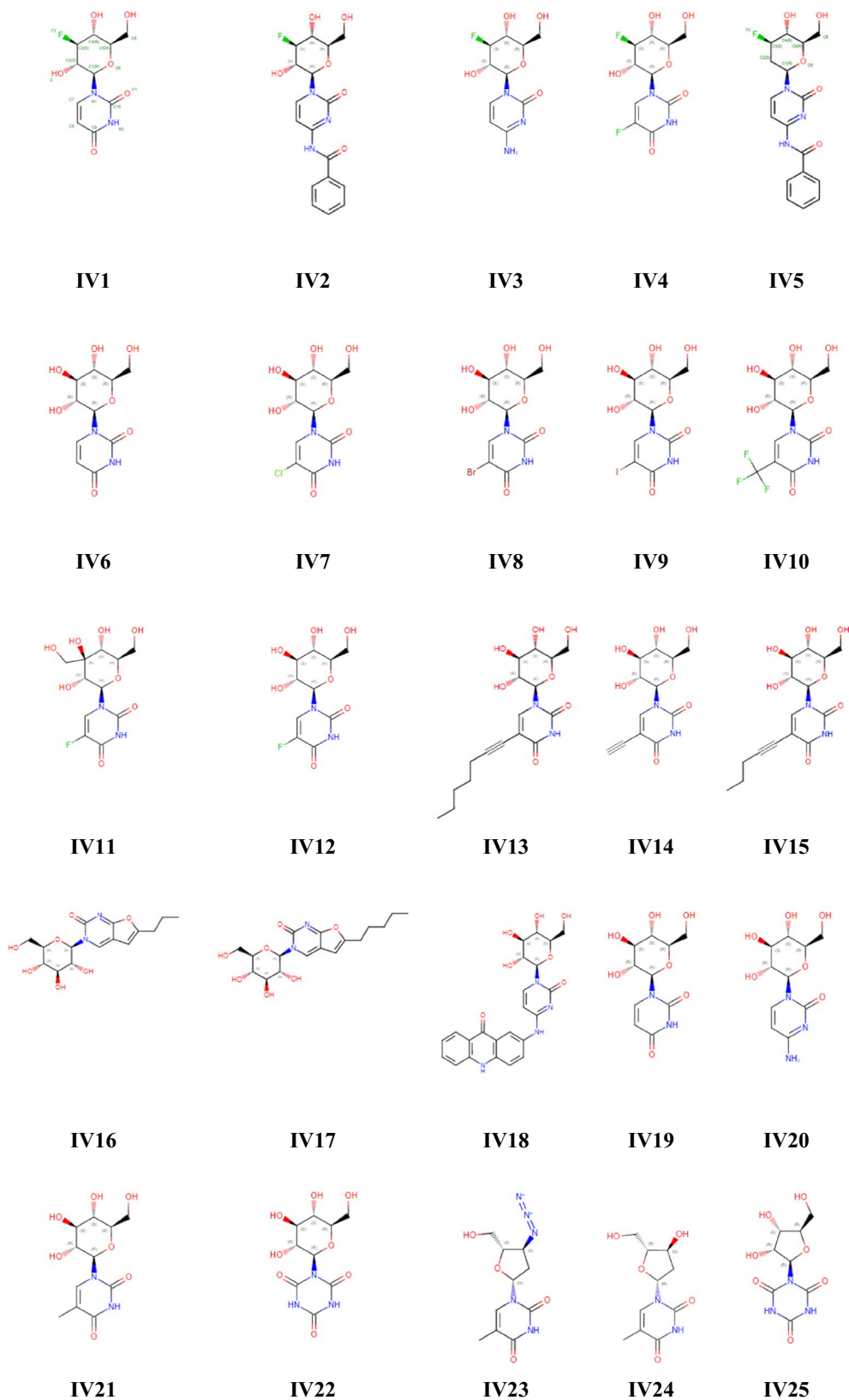
Table 4 (C- β -D-glucopyranosyl)-hydroquinone regioisomers and their X-ray crystal structure information

Table 4 (continued)

Compd	PDB	Ligand	Ki (nM)	pKi			PI-2 ^b	I ^c	Bind Site	Ref.
				Obs ^a	Cal	Dev				
IV1	3L79	DKX	3,460,000	2.46	2.37	0.09	−0.72	1	C	[41]
IV2 ^d	3L7A	DKY	46,420	4.33	5.07	−0.74	0.62	1	C	[41]
IV3	3L7B	DKZ	4,010,000	2.40	2.33	0.07	−0.99	1	C	[41]
IV4	3L7C	DK4	3,670,000	2.44	2.40	0.04	−0.24	1	C	[41]
IV5	3L7D	DK5	6,550,000	2.18	2.38	−0.20	0.62	1	C	[41]
IV6	3T3D	CJB	12,390	4.91	5.20	−0.29	−0.72	0	C	[42]
IV7 ^d	3T3E	GPQ	1020	5.99	5.26	0.73	−0.24	0	C	[42]
IV8	3T3G	GPU	3270	5.49	5.22	0.27	0.48	0	C	[42]
IV9 ^d	3T3H	GPV	1940	5.71	4.98	0.73	0.74	0	C	[42]
IV10	3T3I	GPW	1700	5.77	5.19	0.52	0.75	0	C	[42]
IV11 ^d	3SYM	GP0	27,100	4.57	5.26	−0.69	−0.24	0	C	[43]
IV12	3SYR	GPK	7900	5.10	5.23	−0.13	−0.24	0	C	[43]
IV13 ^d	4EJ2	D1F	3,204,000	2.49	2.51	−0.02	2.19	0	C	[44]
IV14	4EL5	D1M	4700	5.33	5.22	0.11	−0.45	0	C	[44]
IV15 ^e	4EKY	D1J	303,000	3.52	4.55	−1.03	1.13	0	C	[44]
IV16	4ELO	D1K	32,400	4.49	5.05	−0.56	1.56	0	C	[44]
IV17	4EKE	D1I	33,400	4.48	4.38	0.10	3.40	0	C	[44]
IV18 ^e	5MEM	7LS	71	7.15	4.09	3.06	1.44	0	C	[45]
IV19	3BCS	CJB	6100	5.22	5.20	0.02	−0.72	0	C	[46]
IV20	3BD8	C3B	7700	5.11	5.16	−0.05	−0.99	0	C	[46]
IV21	3BD7	CKB	6600	5.18	5.23	−0.05	−0.22	0	C, I	[46]
IV22	3BDA	C4B	–	–	–	–	–	–	C, I	[46]
IV23	3BCR	AZZ	–	–	–	–	–	–	I	[46]
IV24	3BCU	THM	–	–	–	–	–	–	I	[46]
IV25	3BD6	RDD	–	–	–	–	–	–	A	[46]

^aCalculated pKi values using Eq. 3a for all except those in (c)

^bCPI values were auto-loaded utilizing the C-QSAR program

^cIndicator variable: 1.0 was assigned for 3-fluoro substituted analogs and 0.0 for all others

^dCalculate pKi values using Eq. 3b

^eFinal outliers after developing Eq. 3a and 3b

^fBinding site: A = allosteric (AMP) binding site, I = inhibitor (purine) binding site, C = catalytic site

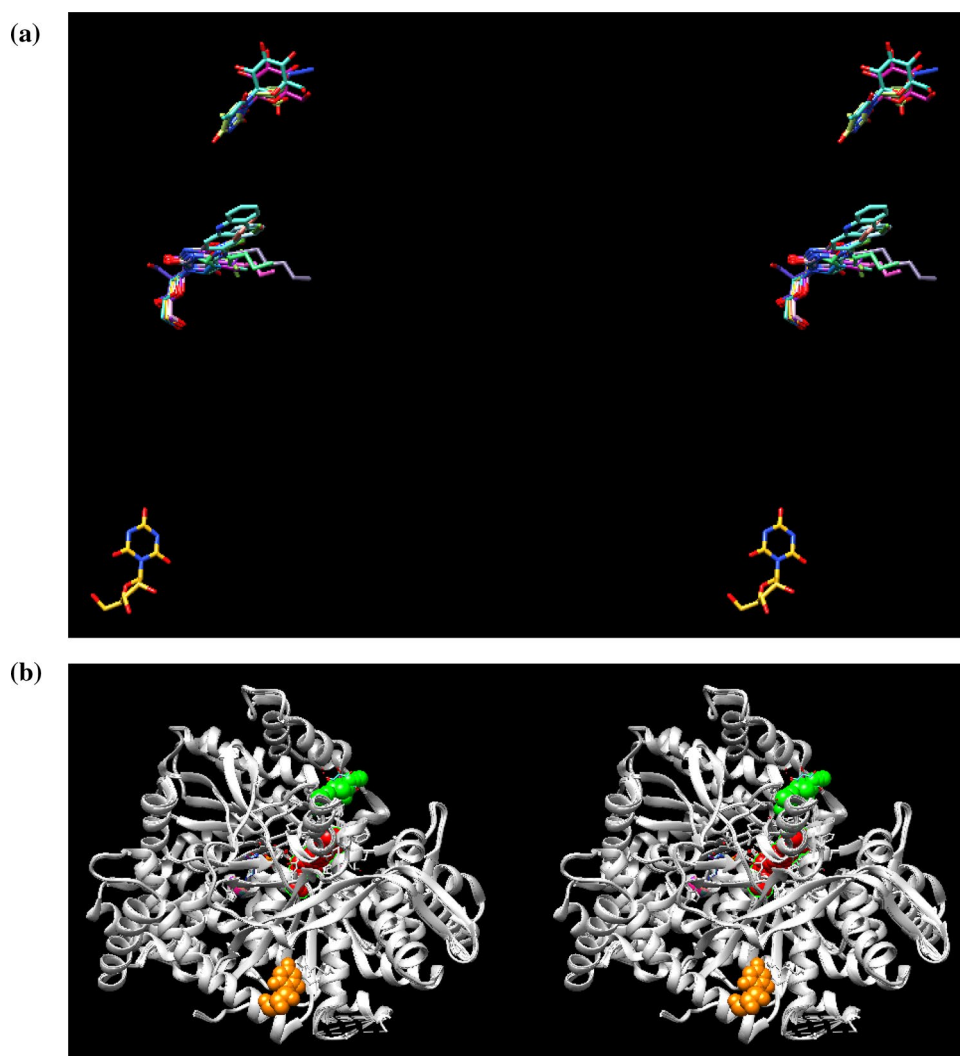
two other binding sites for some compounds. Compound **IV21** is a structural analog of **IV1–IV20**, but it binds at the inhibitor binding site. Compound **IV22** is similar to both **IV21** and **IV25**, but it binds to both the catalytic site and the inhibitor binding site. Besides, the binding sites of the three furanosyl compounds (**IV23**, **IV24**, and **IV25**) are different from **IV1–IV20**. Compounds **IV23** and **IV24** bind to the inhibitor (purine) binding site, whereas **IV25** binds to the allosteric (AMP) binding site.

Figure 4 is a stereo pair picture of compounds **IV1–IV25** at the binding site of GP. Figure 4a shows the binding modes at the three binding sites of 20 compounds (**IV1–IV20**): the catalytic site (C, middle), the inhibitor (purine) binding site (I, top), and the allosteric (AMP) binding site (A, bottom). Figure 4b shows the locations of these three binding sites of GPb.

Despite different structural analogs (spiro vs. hydroquinones), the overall binding situations are similar to the compounds listed in Tables 3 and 4: the primary binding site is the catalytic site for most compounds. Only a limited number of compounds have been reported to bind at the allosteric binding site (new allosteric binding site NA, inhibitor binding site I, or allosteric AMP binding site A). In addition, their QSARs were also developed from two subgroups even though their parameters were the same in each group.

Equation 3a and 3b was developed from the compounds listed in Table 4 using the ‘splitting QSAR’ approach. Because only five compounds were involved in deriving Eq. 3b, two-term QSAR correlations were not considered.

Fig. 4 **a** Binding modes of **IV1–IV25** in the catalytic site (C, middle), the inhibitor (purine) binding site (I, top), and the allosteric (AMP) binding site (A, top). **IV25** is shown at the bottom. **b** **IV18** (red), **IV23** (green), **IV25** (orange) are shown in the CPK model showing the locations of their binding sites



$$\begin{aligned} \text{pKi} &= -0.07(\pm 0.06) \text{CPI}^2 - 2.83(\pm 0.38)I + 5.23(\pm 0.22) \\ n &= 14, r^2 = 0.96, q^2 = 0.68, s = 0.287 \quad (3a) \\ \text{outlier: } &\mathbf{IV2, IV7, IV9, IV11, IV13, IV15, IV18} \end{aligned}$$

$$\begin{aligned} \text{pKi} &= -0.58(\pm 0.65) \text{CPI}^2 + 5.30(\pm 1.41) \\ n &= 5, r^2 = 0.73, q^2 = -0.47, s = 0.835 \quad (3b) \\ \text{outlier: } &\mathbf{IV15, IV18} \end{aligned}$$

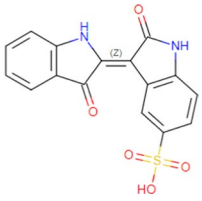
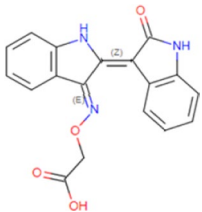
Even though the inhibitory potencies of these compounds correlated well with the same parameter CPI^2 in Eqs. 3a and 3b, the size of the coefficients of the CPI^2 term were significantly different, indicating the two sets of compounds behaved differently. A similar phenomenon was observed with the compounds in Table 3 (Eqs. 2b and 2d). In Eq. 3a, an indicator variable was used for the five compounds of 3-deoxy-3-fluoro- β -D-glucopyranosyl derivatives (**IV1–IV5**). The negative coefficient of the indicator variable

showed that the 3-deoxy-3-fluoro derivatives yielded significantly weaker potency than the other compounds. This result is consistent with the suggestion of Tsirkone et al. [41]. They indicated the 3-hydroxyl group of the glucose moiety was a good hydrogen bond donor and acceptor, but the corresponding 3-fluorine was not as good and did not improve the potency as much.

Equation 3a and 3b explains 96% and 73% of the variance in the inhibitory activity data, respectively. **VIII15** and **VIII18** are outliers in Eq. 3b. Relatively low correlation coefficient and high standard deviation of Eq. 3b both indicate that there are other effects that have not been accounted for. However, the limited number of data points prevented further investigation.

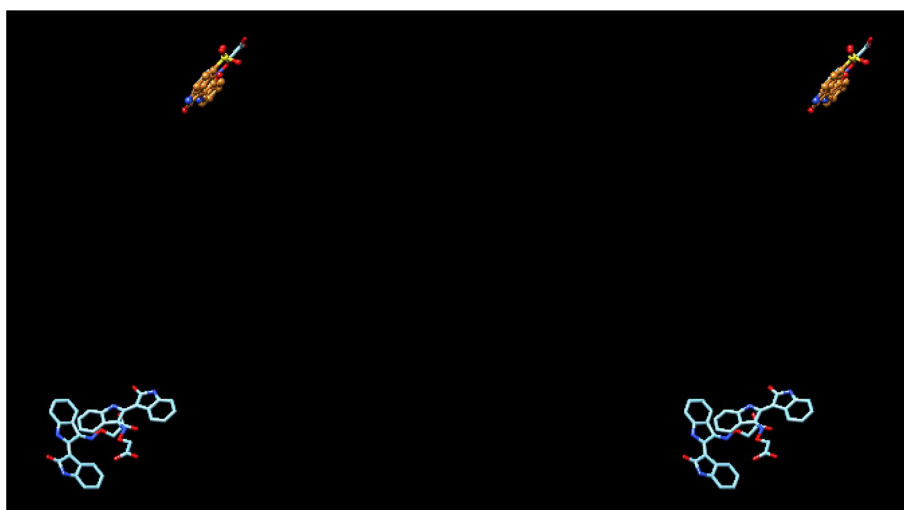
Kantsadi et al. [44] reported that an extended C5-alkynyl group exploited interactions with the β -pocket of the active site and induced significant conformational changes of the 280 s loop. **IV13** induced significant conformational changes in the 280 s loop, while other compounds also induced such

Table 5 Indirubin analogs and their X-ray crystal structure information

Compound	Structure	PDB	Ligand	Ki (nM)	pKi	Bind Site ^a	Ref.
V1	 Indirubin-5-sulphonate	1UZU	1NR (E226)	13,800	4.86	I	[47]
V2	 Indirubin-3'-aminooxy-acetate	1Z62	IAA (E243)	16,000	4.80	I, A	[48]

^aBinding site: I = inhibitor (purine) binding site, A = allosteric (AMP) binding site

Fig. 5 Binding modes and binding sites of indirubin analogs in Table 5. **V1** bound at the purine inhibitor site (I). **V2** (E243, IAA, PDB ID: 1Z62, cyan) bound at the purine inhibitor site (I) as well. Two **V2** molecules are bound at the allosteric activator AMP binding site (A) and a new subsite in the vicinity of the allosteric site



conformational change of the 280 s loop, though to a lesser extent.

IV15 and **IV18** are final outliers of Eq. 3b. Their observed K_i values cannot be explained based on Eq. 3b: The observed K_i value of **IV15** is about tenfold weaker than the calculated value, and that of **IV18** is more than 1000-fold stronger than the calculated value. Such unusual binding potency of **IV15** was also noted by Kantsadi et al. [44]. They suggested that the alkynyl group of **IV15** was pointing toward the side chains of Asp339 and His341. Thus, restructuring of the 339–341 loop and the change of water structure were suggested for the increase in K_i value. Such changes were in contrast to a significant conformational change of the 280 s loop structure upon binding of a similar compound such as **IV13**. On the other hand, Mamais et al. [45] explained that the tight binding of **IV18** was because of the increased hydrogen bonding network and van der Waals

interactions due to the conformational changes in the side chains of the 280 loop residues observed in the **IV18**-GP complex. Therefore, it could be concluded that the outliers **IV15** and **IV18** in Eq. 3b were not due to the binding at two different binding sites.

Among the 25 compounds listed in Table 4, two compounds (**IV21** and **IV22**) were reported to bind at the two binding sites. Equation 3a includes **IV21**, and no other parameter was required to account for any effects appropriate to the dual binding of **IV21**, indicating that the binding at the two binding sites does not cause significant effects on its binding.

Compounds **IV22** and **IV23–IV25** that only bound at the allosteric binding site were not included in deriving Eqs. 3a or 3b, because their K_i values were unavailable. Consequently, the effects of two-site binders on QSAR were not examined with allosteric correlations. Nevertheless, these

compounds provide other examples that show structurally close analogs bind at various binding sites.

Indirubin derivatives

Kosmopoulou et al. [47, 48] reported the binding mode of indirubin-5-sulphonate (**V1**, E226) and indirubin-3-aminoxy-acetate (**V2**, E243) to GPb by kinetic and crystallographic experiments (Table 5). They showed **V1** was a competitive inhibitor with respect to ATP and **V2** was a competitive inhibitor with respect to both Glc-1-P and AMP. The X-ray crystal structures showed that **V1** bound at the inhibitor (purine) binding site (Fig. 5). Only one molecule was bound at this site. On the other hand, two additional **V2** molecules were bound at the allosteric (AMP) binding site and a new subsite in the vicinity of the allosteric site, respectively.

Even though there are only two crystal structures of indirubin analogs, this case represents another example that structural analogs bind at various binding sites of the same allosteric enzyme.

Flavonoids

Chetter et al. [28] reported that flavonoids are novel inhibitors of GP, but their mode of action is unspecific in regard to the GP binding sites involved. The crystal structures show **V11**, **V12**, and **V13** bound exclusively at the inhibitor binding site. Tsitsanou et al. [49] reported that **V14** (chrysin) and **V16** were accommodated at the inhibitor site, whereas flavonoid **V18** (quercetagenin) was bound at the allosteric site [50]. Kantsadi et al. [51] reported that **V15** bound at the novel binding site. Anderka et al. [52] described that quinolone class **V17** (AVE9423) was bound to the allosteric AMP site. In addition, Kato et al. [53] showed **V18** bound at the GP allosteric site. The binding site of **V19** was reported to be unknown [51]. The inhibitory potencies and X-ray crystal structure information (Fig. 6) of **V11–V17** are listed in Table 6.

The inhibitory potency determined from rmGPb of **V11–V15** in Table 6 correlates well with CPI (Eq. 4a), explaining 91% of the variance in the biological data. Equation 4a indicates that the inhibitory potency (pKi) of these compounds can be accounted for with their hydrophobicity. Statistically slightly less satisfactory correlation was obtained with CMR ($r^2=0.88$, $s=0.207$). There are significant correlations between the inhibitory potency determined from rmGPb, rmGPa, and hlGPa as shown in Eqs. 4b–4d.

Compound **V15** bound at the quercetin binding site, a completely independent binding site of **V11–V14**. The binding potency of **V15** was about tenfold weaker than **V11–V14**. Upon critical examination of Eq. 4a, we realized that this one point (**V15**) greatly influenced the correlation, yielding

a statistical artifact. Without **V15**, a statistically less strong Eq. 4e was obtained. The pKi values from rmGPa and hlGPa gave statistically similar but weaker correlations than Eq. 4e. Therefore, Eq. 4a was considered as a preliminary QSAR.

An optimistic point of Eq. 4a and 4e was that these allosteric QSARs indicated that there may be something unusual about **V15** from the rest of the analogs, even if these compounds were structurally similar and could typically be considered as an analog for SAR/QSAR studies. An interesting correlation obtained excluding **V15** was Eq. 4f with CMR^2 . Equation 4f was statistically superior to Eq. 4e. Including **V15**, the correlation with CMR^2 was statistically much inferior ($r^2=0.21$, $s=0.723$).

$$pKi_{(rmGPb)} = 0.27(\pm 0.16) CPI^2 + 3.89(\pm 0.84) \\ n = 5, r^2 = 0.91, q^2 = 0.67, s = 0.251 \quad (4a)$$

$$pKi_{(rmGPb)} = 1.29(\pm 0.71) pKi_{(rmGPa)} - 1.46(\pm 3.66) \\ n = 5, r^2 = 0.92, q^2 = 0.78, s = 0.233 \quad (4b)$$

$$pKi_{(rmGPb)} = 1.29(\pm 0.58) pKi_{(hlGPa)} - 1.36(\pm 2.95) \\ n = 5, r^2 = 0.94, q^2 = 0.85, s = 0.194 \quad (4c)$$

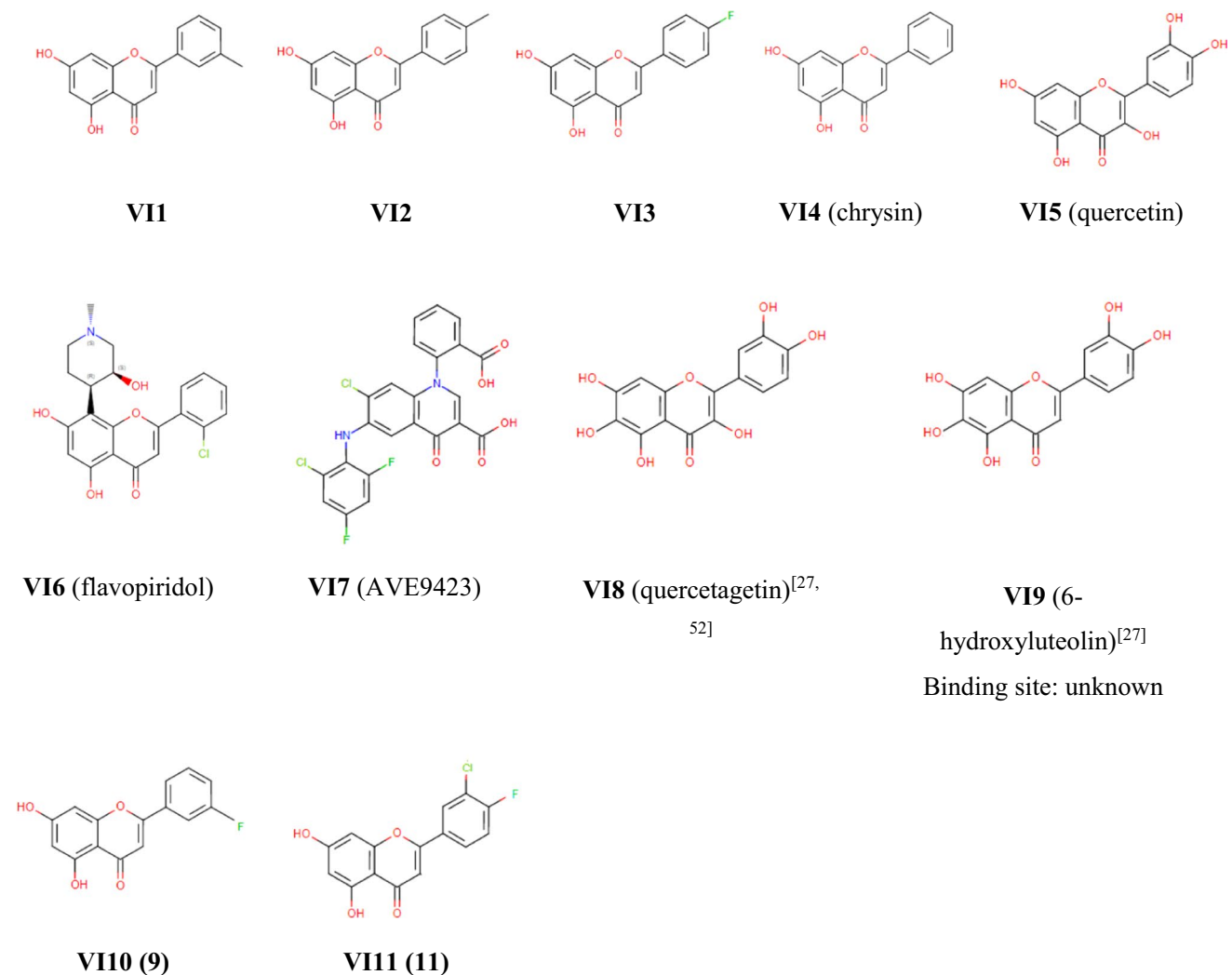
$$pKi_{(rmGPa)} = 0.99(\pm 0.13) pKi_{(hlGPa)} + 0.16(\pm 0.67) \\ n = 5, r^2 = 1.00, q^2 = 0.04, s = 0.194 \quad (4d)$$

$$pKi_{(rmGPb)} = 0.36(\pm 0.57) CPI^2 + 3.37(\pm 3.29) \\ n = 4, r^2 = 0.79, q^2 = 0.23, s = 0.273 \quad (4e)$$

$$pKi_{(rmGPb)} = 0.31(\pm 0.36) CMR^2 + 3.04(\pm 2.74) \\ n = 4, r^2 = 0.88, q^2 = 0.51, s = 0.207 \quad (4f)$$

Among the compounds that Chetter et al. [28] reported were two structural analogs, **V110** and **V111**. Assuming their binding site was the same as **V11–V14**, these two compounds were added to derive fresh QSARs. Equation 4g was the result. Equation 4g indicates that **V110** and **V111** behave in the same way as **V11–V14**, suggesting that they bind at the catalytic site. **V15** became an outlier in this case: the difference between the observed and the calculated pKi values is 1.35. The result is not surprising because **V15** is bound at a separate binding site from **V11–V14**. Equation 4f is provided for comparison.

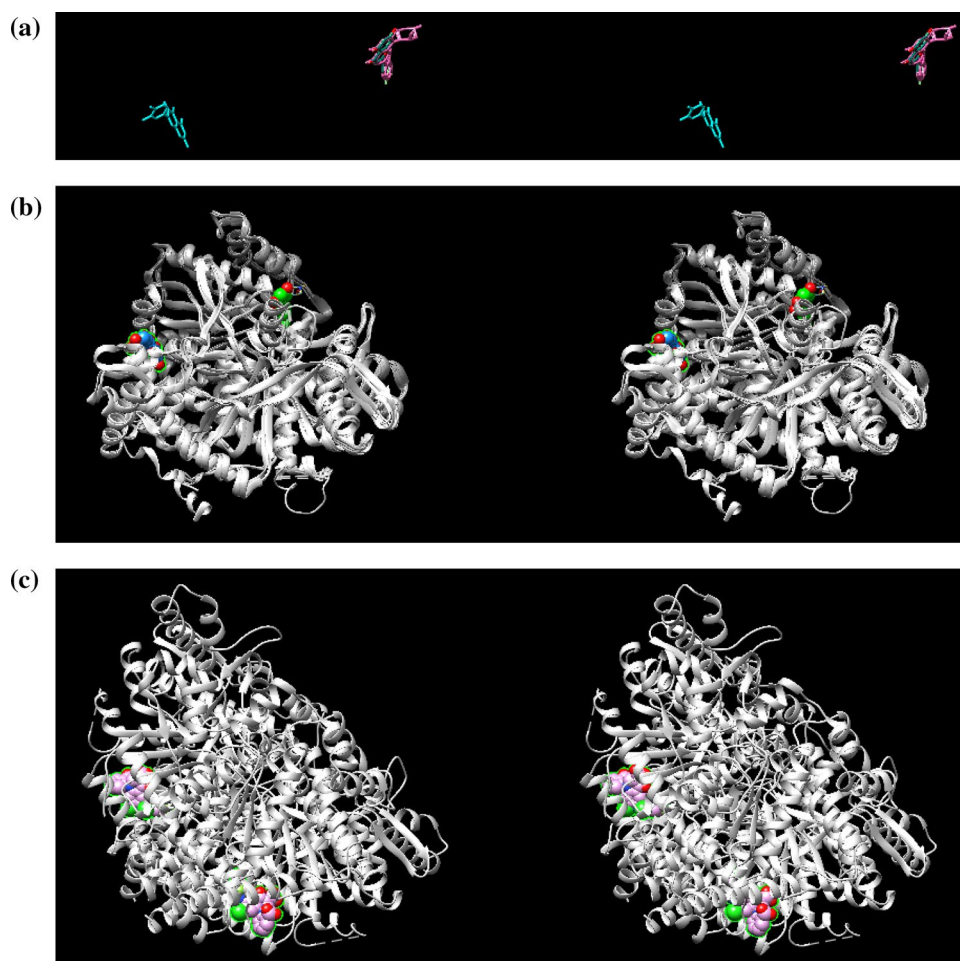
$$pKi_{(rmGPb)} = 0.28(\pm 0.15) CMR^2 + 3.32(\pm 1.14) \\ n = 6, r^2 = 0.87, q^2 = 0.72, s = 0.167 \quad (4g) \\ \text{outlier : IV5}$$

Table 6 Flavonoid analogs and their X-ray crystal structure information. (VI8 and VI9 are additionally included for comparison.)

Compd	PDB	Lig	Ki (nM)			pKi			CPI ^b	CMR ^b	Bind Site ^c	Ref.		
			rmGPb	rmGPa	hlGPa	rmGPb	cal ^a	dev						
VI1	6Y55	O9Q	1000	1750	2230	6.00	5.76	0.24	5.76	5.65	2.60	2.97	I	[28]
VI2	6Y5C	O9T	2310	3,460	3890	5.64	5.76	-0.12	5.46	5.41	2.60	2.97	I	[28]
VI3	6Y5O	O9Z	11,400	17,790	21,360	4.94	5.08	-0.14	4.75	4.67	2.24	2.53	I	[28]
VI4	3EBO	57D	7560	5140	7280	5.12	5.06	0.06	5.29	5.14	2.10	2.51	I	[49] [28]
VI5 ^d	4MRA	QUE	69,120	32,930	43,520	4.16	5.51	-1.35	4.48	4.36	0.85	2.82	Q	[51]
VI6	1E1Y 1C8K 3EBP	CPB CPB CPB	2340	-	-	5.91							I I I	[50] [50] [49]
VI7 ^e	3CEM	AVD	-	-	-	-							A	[52]
VI10	-	9	7260	8550	7390	5.14	5.08	0.06	5.07	5.13	2.24	2.53		[28]
VI11	-	11	1890	2270	3390	5.72	5.83	-0.11	5.64	5.47	2.95	3.02		[28]

^aCalculated using Eq. 4g^bAuto-loaded utilizing the C-QSAR program^cBinding site: Q = quercetin binding site, I = inhibitor (purine) binding site, A = allosteric (AMP) binding site^dNot included in Eq. 4g^eBound to hGPa

Fig. 6 **a** Binding modes of seven compounds of **VII–VII7** at the inhibitor binding site (right) and the quercetin (**VI5**) binding site (left, cyan). **b** Location of the binding sites: the inhibitor (purine) binding site (I, **VII1**, green) and the quercetin binding site (Q, **VI5**, blue). **c** Location of the allosteric (AMP) binding site (A) of **VII7** (pink). (The one on the left is from the allosteric (AMP) binding site of the other dimer.)



$$pK_{i(\text{rmGPb})} = 0.20(\pm 0.22) \text{CPI}^2 + 4.18(\pm 1.41)$$

$$n = 6, r^2 = 0.61, q^2 = -0.23, s = 0.291 \quad (4h)$$

outlier : **IV5**

Equations **4g** and **4f** are essentially identical within the confidential limits: the same coefficient of CMR, intercept, and r^2 , but q^2 and s values improved.

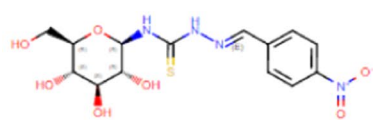
Even though **VII7** was not included in Eqs. **4a–4h** because of the lack of K_i value, this compound was also expected to become an outlier in Eq. **4g** because its binding site was also different.

Hansch and his co-workers attributed allosteric inverted parabolic (or bilinear) QSARs to a change in the structure of the receptor and/or a change in the reaction mechanism upon ligand binding to the protein [5–10]. They also suggested that such inverted parabolic correlations could be due to the presence of another binding site [5]. The allosteric QSAR Eq. **4g** supports Hansch's suggestions and indicates that the allosteric inverted parabola correlation could be due to the conformational change in the protein. Besides the inverted parabola QSAR for the allosteric interactions, the binding of **IV5** at the new allosteric site stood out as an outlier.

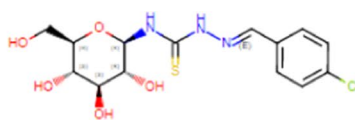
β -D-Glucopyranosyl-thiosemicarbazone derivatives

Alexacou et al. [54] reported 15 aromatic aldehyde 4-(β -D-glucopyranosyl)thiosemicarbazones listed in Table 7 as inhibitors of rabbit muscle GPb. They described that these compounds were competitive inhibitors of GPb with respect to α -D-glucose-1-phosphate and revealed the inhibitors were accommodated at the catalytic site with the glucopyranosyl moiety at approximately the same position as α -D-glucose.

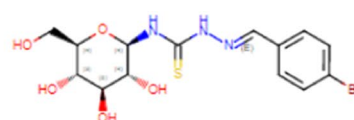
These inhibitors fit tightly into the β -pocket, a side-channel from the catalytic site with no access to the bulk solvent. Alexacou et al. reported that 14 out of the 15 inhibitors bound at the new allosteric site of the enzyme as well as the catalytic site [54]. They described that the binding of several compounds (**VII2**, **VIII3**, **VIII4**, **VII5**, **VIII8**, and **VIII11**) triggered a significant shift of the 280 s loop. On the other hand, the orthonitro-substituted compound **VIII13** (PDB ID: 3MSC, 24S) was described to bind only at the catalytic site and not bound at the new allosteric site [54]. However, to our surprise, examination of the corresponding crystal structure revealed that this compound was equally bound at the catalytic site as well as the new allosteric binding site

Table 7 Aromatic aldehyde 4-(β -D-glucopyranosyl)thiosemicarbazones reported by Alexacou et al. [54] and their X-ray crystal structure information

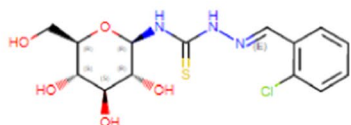
VII1



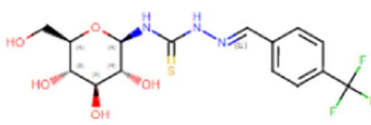
VII2



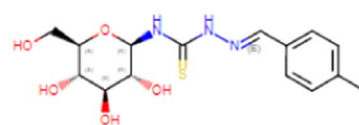
VII3



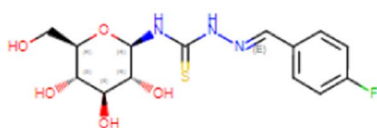
VII4



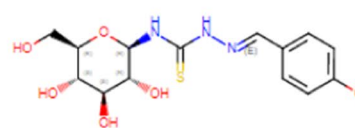
VII5



VII6



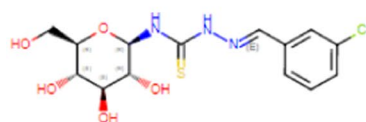
VII7



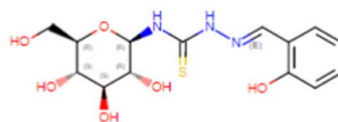
VII8



VII9



VII10



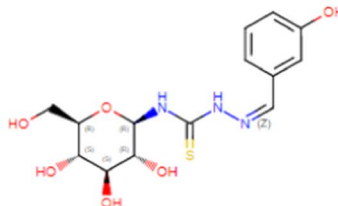
VII11



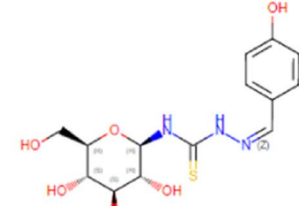
VII12



VII13



VII14



VII15

Table 7 (continued)

Compd	PDB	Ligand	IC ₅₀ (nM)	pIC ₅₀			CPI ^b	CMR ^b	Iz	Bind Site ^c	Ref.
				obs ^a	cal	dev					
VIII1	3MT9	18O	25,700	4.59	4.78	−0.19	1.26	6.26	0	C, NA	[54]
VII2	3MT8	17 T	28,300	4.55	4.47	0.08	2.23	6.14	0	C, NA	[54]
VII3	3MT7	16O	93,200	4.03	4.30	−0.27	2.38	6.43	0	C, NA	[54]
VII4 ^d	3MS7	22S	370,000	3.43	3.45	−0.02	2.23	6.14	0	C, NA	[54]
VII5 ^d	3MS4	21 N	524,300	3.28	3.27	0.01	2.40	6.16	0	C, NA	[54]
VII6 ^d	3MS2	18S	192,400	3.72	3.71	0.01	2.02	8.12	0	C, NA	[54]
VII7	3MQF	20X	5700	5.24	4.82	0.42	1.66	5.67	0	C, NA	[54]
VII8 ^d	3MRX	17S	406,500	3.39	2.25	1.14	1.79	6.27	0	C, NA	[54]
VII9	3MTA	22O	50,400	4.30	4.30	0.00	2.38	6.43	0	C, NA	[54]
VII10	3MTB	23 V	23,200	4.64	4.47	0.17	2.23	6.14	0	C, NA	[54]
VII11	3NC4	26O	26,600	4.58	4.83	−0.25	1.59	5.81	0	C, NA	[54]
VII12	3MRT	12E	200,000	3.70	3.64	0.06	0.23	5.44	0	C, NA	[54]
VII13 ^e	3MSC	24S	484,200	3.32	3.48	−0.16	1.26	6.26	1	C, NA	[54]
VII14	3MRV	16F	180,000	3.75	3.53	0.22	1.59	5.81	1	C, NA	[54]
VII15	3MTD	25E	340,500	3.47	3.53	−0.06	1.59	5.81	1	C, NA	[54]

^aEq. 5a was used to calculate the pIC₅₀ values of compounds VIII–VII15 except for compounds VII4, VII5, VII6, and VII8

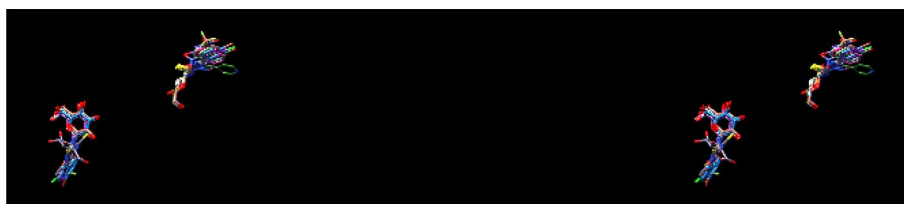
^bThe parameter values were auto-loaded utilizing the C-QSAR program

^cBinding site: C = catalytic site, NA = new allosteric site

^dEq. 5b was used to calculate the pIC₅₀ values of compounds VII4, VII5, VII6, and VII8

^eAlexacou et al. [54] reported that VII13 bound only at the catalytic site. However, the PDB file shows VII13 also bound at the two separate binding sites

Fig. 7 Binding modes of VIII–VII15 at the two binding sites



(Fig. 7). The orthonitro compound VII13 is a Z isomer to the N=C double bond, but so are the two other compounds VII14 and VII15.

Equations 5a and 5b was developed using the ‘split QSAR’ method suggested by Verma and Hansch [5]. The indicator variable Iz was assigned the value of one for the three Z-isomers (VII13–VII15) and zero for all others. The inhibitory potencies (pIC₅₀) of these compounds correlated parabolically with CPI. It is a normal parabola, not an inverted one. The negative coefficient of Iz indicates that the three Z-isomers are about 20-fold weaker than the others. Four compounds (VII4–VII6, VII8) were not used in Eq. 5a (QSAR1). An inverted parabolic relationship for Eq. 5b or 5c could be observed when their pIC₅₀ values were plotted against CPI or CMR values. However, because not enough data points were available, two-parameter equations were not

considered. Excluding VII8, the remaining three compounds yielded Eq. 5b or 5c (QSAR2). There is a high collinearity between CPI and CMR for these compounds (Eq. 5d). Thus, Eq. 5b or 5c should be considered preliminary.

$$\begin{aligned} \text{pIC}_{50} = & 2.18(\pm 1.30) \text{CPI} - 0.72(\pm 0.46) \text{CPI}^2 \\ & - 1.30(\pm 0.50) \text{Iz} + 3.17(\pm 0.82) \\ n = 11, r^2 = 0.87, q^2 = -1.45, s = 0.259 \end{aligned} \quad (5a)$$

outlier : VII4, VII5, VII6, VII8

optimum CPI = 1.52(± 0.32)

$$\begin{aligned} \text{pIC}_{50} = & -1.16(\pm 1.62) \text{CPI} + 6.04(\pm 3.60) \\ n = 3, r^2 = 0.99, q^2 = 0.81, s = 0.035 \end{aligned} \quad (5b)$$

outlier : VII8

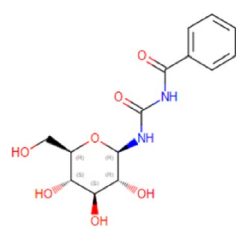
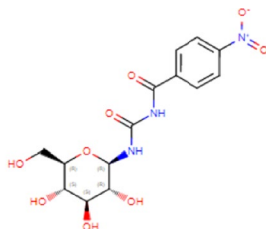
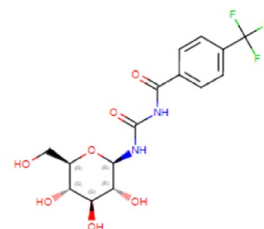
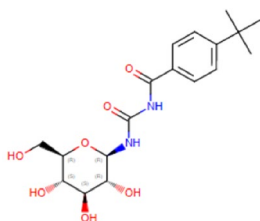
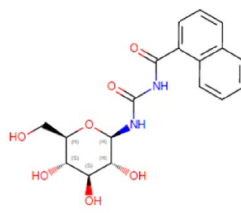
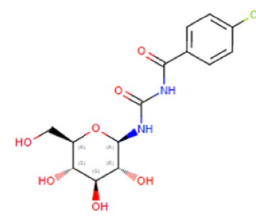
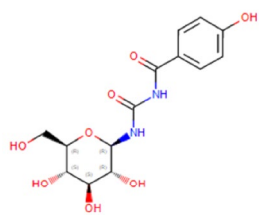
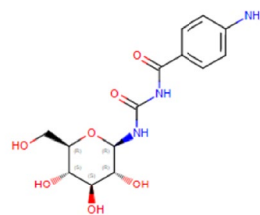
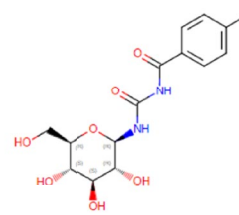
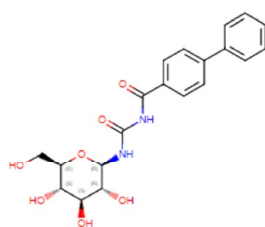
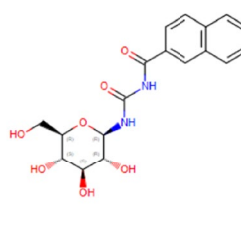
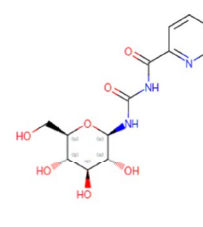
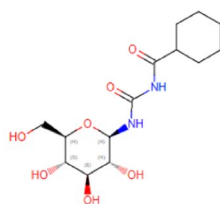
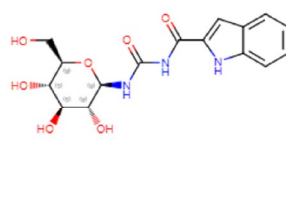
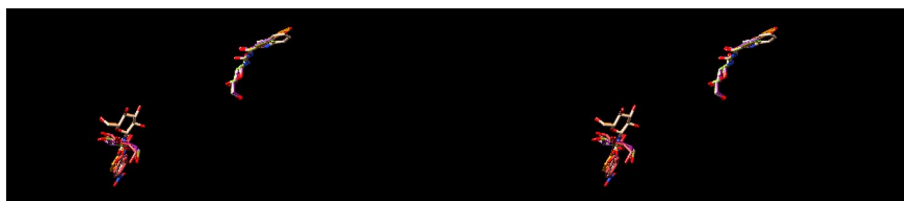
Table 8 N-substituted-N-β-D-Glucopyranosyl)urea derivatives reported by Chrygina et al. [55] and others [56, 57] and their X-ray crystal structure information**VII16****VII17****VII18****VII19****VII20****VII21****VII22****VII23****VII24****VII25****VII26****VII27****VII28****VII29**

Table 8 (continued)

Compd	PDB	Ligand	Ki (nM)	pKi			cal ^b	dev ^b	CPI ^c	CMR ^c	Bind Site ^d	Ref.
				obs	cal ^a	dev ^a						
VII16 ^e	1K06 1K08 2QNB	BZD	4600	5.34	4.67	0.67	5.10	0.24	1.19	4.34	C, NA C, NA C, NA	[56] [57] [57]
VII17	2QN8	NBY	3300	5.48	5.73	−0.25	5.22	0.26	1.36	4.95	C, NA	[57]
VII18	3ZCQ	62 N	1800	5.75	5.59	0.16	5.20	0.55	2.42	4.85	C	[55]
VII19	3ZCR	F85	700	6.16	6.18	−0.02	5.47	0.69	3.01	6.19	C	[55]
VII20 ^e	3ZCS	CAW	5000	5.30	6.25	−0.95	5.43	−0.13	2.36	6.02	C	[55]
VII21	2QN3	F55	4400	5.36	5.56	−0.20	5.20	0.16	3.01	6.19	C	[57]
VII22	2QN7	HBZ	6300	5.20	4.98	0.22	5.13	0.07	0.86	4.49	C	[57]
VII23	2QN9	NBX	6000	5.22	5.37	−0.15	5.17	0.06	0.28	4.71	C	[57]
VII24	2QLM	F68	2300	5.64	5.52	0.12	5.19	0.45	1.69	4.80	C	[57]
VII25	2QLN	F59	3700	5.43	5.50	−0.07	5.60	−0.17	3.08	6.80	C	[57]
VII26 ^f	3ZCT	VMP	350	6.46	6.25	0.21	5.43	1.03	2.36	6.02	C, NA	[55]
VII27 ^f	3ZCU	T68	68,000	4.17	4.17	0.00	5.06	−0.89	0.79	4.13	C	[55]
VII28	3ZCP	F58	–		4.87		5.12		1.71	4.43	C	[55]
VII29 ^e	3ZCV	N85	4000	5.40	6.14	−0.74	5.30	0.10	1.81	5.37	C, NA	[55]

^aCalculated using Eq. 6^bCalculated using Eq. 8a^cAuto-loaded utilizing the C-QSAR program^dBinding Site: C = catalytic site, NA = new allosteric (indole) binding site^eNot used in deriving Eq. 6^fNot used in deriving Eq. 8a**Fig. 8** Binding modes of **VII16**, **VII17**, **VII26**, and **VII29** at the catalytic site (right) and the new allosteric binding site (left)

$$\text{pIC}_{50} = -9.54(\pm 8.55) \text{CMR} + 62.05(\pm 52.49)$$

$$n = 3, r^2 = 1.00, q^2 = 0.92, s = 0.023 \quad (5c)$$

outlier : **VII8**

$$\text{CPI} = 8.22(\pm 4.13) \text{CMR} - 48.27(\pm 25.34)$$

$$n = 3, r^2 = 1.00, q^2 = 0.97, s = 0.011 \quad (5d)$$

VIII–VIII15 are allosteric inhibitors that bind at two binding sites of GP, and the allosteric QSARs expressed as Eq. 5a is a normal parabola correlation.

Alexacou et al. [54] mentioned the binding of these inhibitors at the new allosteric site slightly shifted in the vicinity residues, indicating a conformational change.

N-substituted-N-β-D-Glucopyranosyl)urea derivatives

Chrysina et al. [55] and others [56, 57] reported N-substituted-N-β-D-glucopyranosyl)ureas listed in Table 8. These compounds are structurally not so much different from those in Table 7. Unlike those compounds in Table 7, four of the 14 compounds (**VII16–VII17**, **VII26**, and **VII29**) bound at the two separate binding sites, and 10 compounds bound only at the catalytic site.

Figure 8 shows the two binding modes (the catalytic site and the new allosteric binding site) of **VII16–VII17**, **VII26**, and **VII29**.

From the compounds listed in Table 8, Eq. 6 was developed. The correlation was a normal parabolic one, not an inverted one. Three compounds were outliers in Eq. 6

(QSAR1). (When pKi values were plotted against CPI or CMR, an indication of a parabolic relationship with CPI or CMR was observed with the three outliers in QSAR2. However, because not enough data points were available, further investigation was not considered.)

$$\begin{aligned} \text{pKi} &= 8.61(\pm 2.79) \text{CMR} - 0.74(\pm 0.25) \text{CMR}^2 - 18.74(\pm 7.52) \\ n &= 10, r^2 = 0.92, q^2 = 0.85, s = 0.193 \quad (6) \\ \text{outlier} &: \text{VIII16, VII20, VII29} \\ \text{optimum CMR} &= 5.82(\pm 0.17). \end{aligned}$$

The information about the crystal structures of most compounds in Table 8 was obtained from the RCSB PDB protein data bank since no paper has been published. An exception was Oikonomakos et al.'s paper. Oikonomakos et al. [56] reported VIII16 bound tightly at the catalytic site and induced substantial conformational changes in the loop containing residues 282–287 of 280 s loop. They showed VIII6 equally bound at the new allosteric site, about 33 Å from the catalytic site. Three other compounds (VIII17, VII26, and VII29) showed that they were equally bound at both the catalytic site and the new allosteric binding site as VIII6.

As the normal parabola correlation of Eq. 5a which was discussed with the compounds in Table 7, Eq. 6 is a normal parabola correlation. This correlation provides another example of normal parabola allosteric QSAR, which involves conformational changes in the protein-inhibitor allosteric interactions.

The binding modes of the final three outliers (VIII16, VII20, VII29) from Eq. 6 are not identical: VIII16 and VII29 bound at the two different binding sites (C, NA), whereas VII20 bound only at the catalytic site (C). Regarding such outliers, please see the further discussion below under the titles of ‘Different binding modes at the allosteric secondary binding sites of glycogen phosphorylase’ and ‘Dual inhibitions and their QSARs.’

N-(β-D-glucopyranosyl)-N'-oxamide derivatives

Czifrak et al. [58] and Hadjiloi et al. [59] studied the binding modes of several N-(β-D-glucopyranosyl)-N'-oxamide analogs (Table 9). They are competitive inhibitors of rabbit muscle GPb with respect to α-D-glucose-1-phosphate. The ligand-bound crystal structures revealed the inhibitors were accommodated at the catalytic site at approximately the same position as α-D-glucose and stabilized the T-state conformation of the 280 s loop. Examination of the crystal structures revealed that only one of the eight compounds listed in Table 9 bound at two separate binding sites.

Figure 9 shows VII33 bound at two separate binding sites. VII33 and VII26 bound at the same two binding sites. However, comparison of VII33 with VII26 and VII29 revealed

substantial differences in the binding mode of VII33 from the other two compounds at the new allosteric binding site. The binding modes of VII26 and VII29 were essentially identical at both binding sites, but a large difference in the binding mode of VII33 could be observed. (Regarding the effects of such a difference in the binding mode on the outlier in QSAR, please see the further discussion below and Ref. [1].)

Equation 7a was developed from the compounds listed in Table 9. One compound (VII35) became an outlier. The inhibitory potencies of these compounds were correlated with CMR with a reasonable s value. Because of the narrow range of the pKi values involved, the squared correlation coefficient (r^2) was not as high as one would hope to see. Anyhow, Eq. 7a indicates the importance of CMR as in Eq. 6.

$$\begin{aligned} \text{pKi} &= 0.19(\pm 0.17) \text{CMR} + 2.93(\pm 0.70) \\ n &= 7, r^2 = 0.63, q^2 = 0.27, s = 0.291 \quad (7a) \\ \text{outlier} &: \text{VII35}. \end{aligned}$$

Hadjiloi et al. [59] discussed comparisons of the compounds in this series with the lead compound N-acetyl-β-D-glucopyranosylamine presented previously. They described that the hydrogen bonding interaction of the amide nitrogen with the main-chain carbonyl oxygen of His377 is missing in these complexes. As they suggested, the differences in the Ki values of these compounds could be partially due to the subtle conformational changes of the protein residues [59].

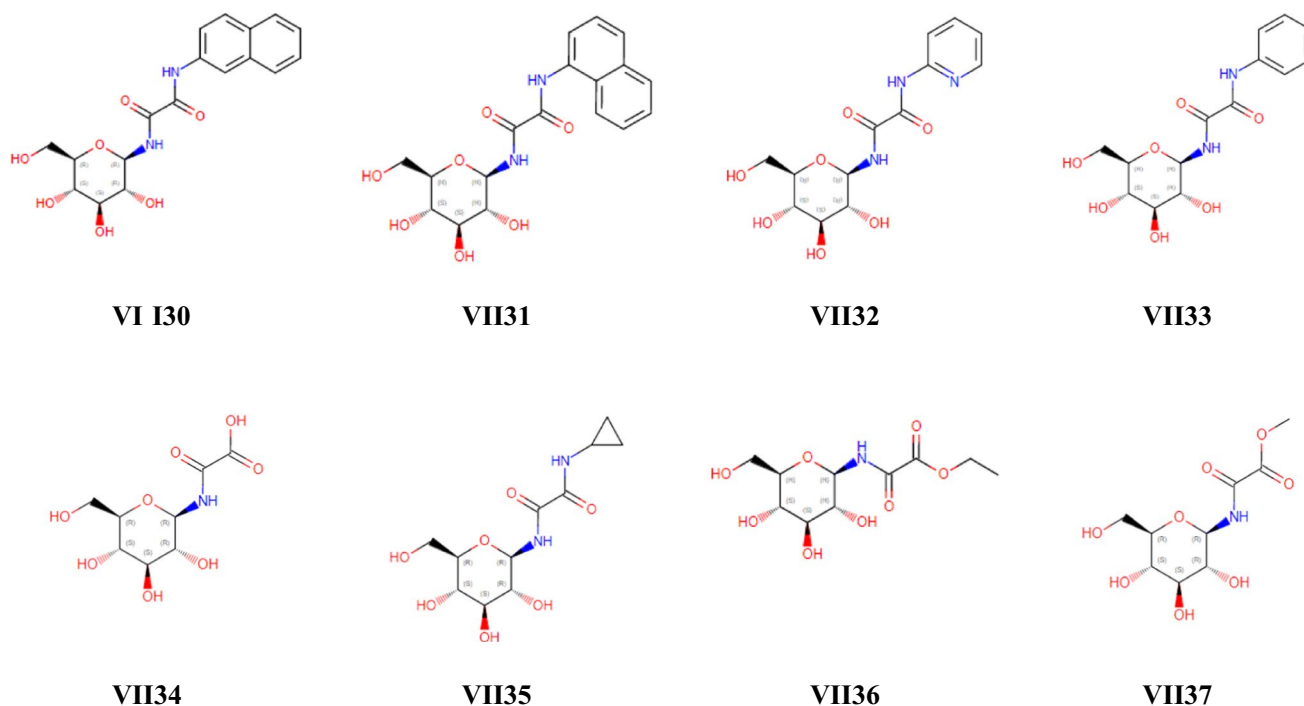
Even though Eq. 7a (or 7b–7d in Supplemental Material 2) is not a parabola/bilinear QSAR, the QSAR describes their allosteric effects.

Acyl urea derivatives

Oikonomakos et al. [60], Anderka et al. [52], and Klabunde et al. [61] reported a ‘novel’ class of GP inhibitors listed in Table 10. They are structurally similar to the side chains of the β-D-glucopyranosyl analogs listed in Table 8 but different in the core structure. They are benzoylaminocarbon-ylaminophenyl analogs.

The X-ray crystal structures of VIII1–VIII4 were done with rmGPb, whereas VIII5–VIII7 were done with hGPb. The first set of four compounds (VIII1–VIII4) bound at the allosteric activator (AMP) binding site. These authors reported the acyl urea analogs inhibited GP by direct inhibition of AMP binding and by indirect inhibition of the substrate-binding through stabilization of the T'-state.

The second set of three compounds (VIII5–VIII7) was equally bound at the allosteric (AMP) binding site, competing with the physiological activator AMP and acting

Table 9 N-(β-D-glucopyranosyl)-N'-oxamide analogs as inhibitors of GP reported by Czifrak et al. [58] and Hadjiloi et al. [59] and their X-ray crystal structure information

Compd	PDB	Ligand	Ki (nM)	pKi			Cal ^b	Dev	CPI ^c	CMR ^c	I ^d	Bind Site ^e	Ref.
				Obs	Cal ^a	Dev							
VII30	3CUT	179	56,000	4.25	4.08	0.17	4.01	0.24	2.41	6.02	1	C	[58]
VII31	3CUU	376	144,000	3.84	4.08	-0.24	4.01	-0.17	2.41	6.02	1	C	[58]
VII32	3CUV	475	230,000	3.64	3.72	-0.08	3.64	0.00	0.19	4.13	1	C	[58]
VII33	3CUW	445	100,000	4.00	3.76	0.24	3.68	0.32	1.23	4.34	1	C, NA	[58]
VII34	2F3P	4GP	710,000	3.15	3.24	-0.09	3.13	0.02	-1.05	1.61	1	C	[59]
VII35 ^f	2F3U	8GP	1,410,000	2.85	3.52	-0.67	3.43	-0.58	-0.37	3.08	1	C	[59]
VII36	2F3S	7GP	920,000	3.04	3.41	-0.37	3.32	-0.28	-1.53	2.54	1	C	[59]
VII37	2F3Q	6GP	210,000	3.68	3.32	0.36	3.23	0.45	-2.05	2.07	1	C	[59]

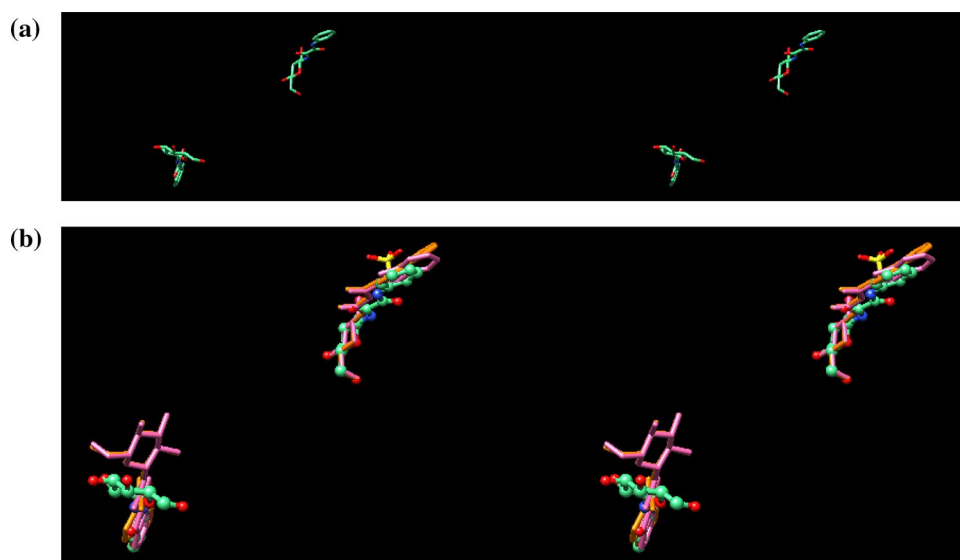
^aCalculated using Eq. 7a^bCalculated using Eq. 7c^cThe parameter values were auto-loaded utilizing the C-QSAR program^dIndicator variable I_{oxamide} was assigned the value of 1.0 for the N-(β-D-glucopyranosyl)-N'-oxamide compounds (VII30–VII37) in Table 9 and 0.0 for all others in Tables 7 and 8^eBinding site: C = catalytic site, N = new allosteric (indole) binding site^fNot used in Eq. 7c

synergistically with glucose. VIII5 occupied only the lower part of the bifurcated AMP site, whereas VIII6 exploited the full binding pocket. Anderka et al. [52] suggested the binding entropy of VIII6 was due to the extensive displacement of solvent molecules as well as to ionic interactions with the phosphate recognition site.

Equation 8 was derived from VIII1–VIII4. Because of the limited number of compounds included to develop Eq. 8,

a statistically weak QSAR was obtained. Nonetheless, it was an inverted parabola correlation. There was a visible indication that the relationship was a reverse parabola correlation when pKi was plotted against CMR. No statistically significant correlation with CPI existed for the corresponding parabola QSAR ($r^2 = 0.19$, $s = 0.20$). Equation 8 provides another example of an inverted parabola QSAR for the allosteric interaction of GP inhibitors.

Fig. 9 **a** Binding modes of **VII33** (PDB ID: 3CUW, 445) at the catalytic site (right) and the new allosteric binding site (left). **b** The difference in the binding modes of **VII33** (3CUW, green, ball-and-stick), **VII26** (orange), and **VII29** (pink) at the catalytic site (right) and the new allosteric binding site (left). While the binding modes of **VII26** and **VII29** are essentially identical at both binding sites, a large difference in the binding mode of **VII33** from the other two compounds can be seen at the new allosteric binding sites



$$\begin{aligned} \text{pKi} = & -4.72(\pm 1.83)\text{CMR} + 0.38(\pm 0.15)\text{CMR}^2 \\ & + 20.24(\pm 5.68) \end{aligned} \quad (8)$$

$$n = 4, r^2 = 1.00, q^2 = -8.00, s = 0.007$$

No QSAR was attempted with the set of **VIII5–VIII7** because of the modest range of the biological activity values and the limited number of compounds available.

β -D-glucopyranosyl triazole, pyrrole, imidazole, thiazole, tetrazole derivatives

The crystal structures of many β -D-glucopyranosyl triazole, pyrrole, imidazole, thiazole, and tetrazole analogs have been reported by Leonidas and his co-workers [26, 62–67]. They are listed in Table 11. Most compounds are bound to the catalytic site. However, **IX7**, **IX8**, and **IX26** are equally bound at the new allosteric binding site in addition to the catalytic site (Fig. 10).

From the compounds listed in Table 11, Eqs. 9a–9c was developed using the ‘split QSAR’ method [5]. For Eq. 9b, an indicator variable I_{thiazole} was assigned for the four thiazole derivatives (**IX29–IX32**). Equations 9a–9c explains 88% of the initial dataset, 94% of the second dataset, and 89% of the final dataset, respectively. No normal or inverted parabola correlation was obtained with these sets. Three QSARs indicated the critical role of the hydrophobic parameter CPI. In addition, the molar refractivity parameter played a significant part to explain the biological activity of those compounds used in Eq. 9a. The coefficient values of CPI for Eqs. 9a and 9c were essentially identical indicating their similar roles in protein–ligand interactions. However, the coefficient of CPI in Eq. 9b was different, suggesting a distinct role of these compounds in their protein–ligand interactions. Such diverse nature of correlations formulated from

the sub-datasets represented the fundamental idea of proposing the ‘split QSAR’ method. The negative coefficients of CPI for Eqs. 9a and 9c may lead to an inverse parabola QSAR if the value of CPI is extended.

The negative coefficient of I_{thiazole} in Eq. 9b indicated that the average amount of thiazole compounds were more than 1000-fold weaker than the others. This result was consistent with that of Kyriakis et al. [26]. They suggested the importance of hydrogen bond interactions between the imidazole ring and the main chain carbonyl group of His377. When replaced by a sulfur atom, such hydrogen bond interaction led to a decrease in the inhibitory activity due to geometrical constraints.

$$\begin{aligned} \text{pKi} = & -1.18(\pm 0.24)\text{CPI} + 1.23(\pm 0.24)\text{CMR} \\ & + 0.67(\pm 0.91) \end{aligned}$$

$$n = 20, r^2 = 0.88, q^2 = 0.84, s = 0.318 \quad (9a)$$

outlier: **IX6, IX8, IX11, IX12, IX16, IX23–IX26, IX29, IX30**

$$\begin{aligned} \text{pKi} = & 0.94(\pm 0.50)\text{CPI} - 3.41(\pm 1.25)I_{\text{thiazole}} + 4.03(\pm 0.96) \end{aligned}$$

$$n = 7, r^2 = 0.94, q^2 = 0.00, s = 0.319$$

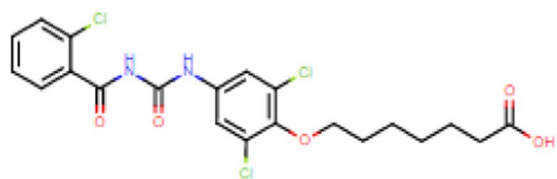
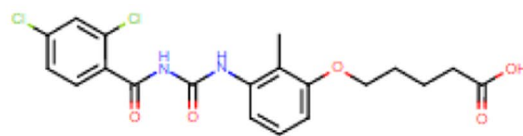
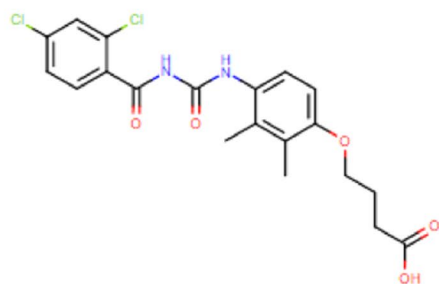
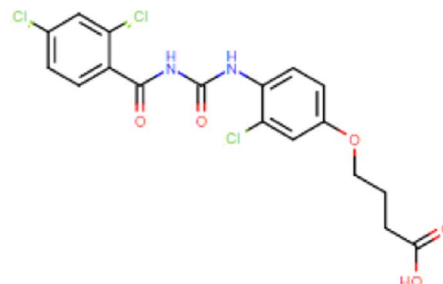
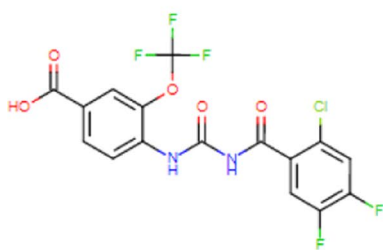
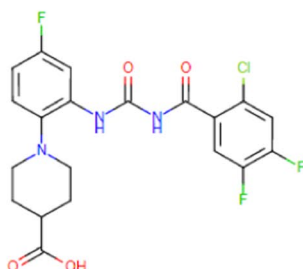
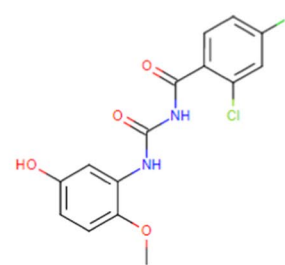
outlier: **IX11, IX23, IX24, IX29** (9b)

$$\begin{aligned} \text{pKi} = & -1.10(\pm 4.64)\text{CPI} + 8.34(\pm 11.17) \end{aligned}$$

$$n = 3, r^2 = 0.89, q^2 = -0.95, s = 0.324 \quad (9c)$$

outlier: **IX24**

Kandsami et al. [67] reported that the binding of the inhibitors **IX5–IX8** did not trigger any significant conformational change of the overall protein structure. The crystal structures of **IX7** and **IX8** showed the inhibitors were

Table 10 Acyl urea analogs as inhibitors of GP reported by Oikonomakos et al. [60], Anderka et al. [52], and Klabunde et al. [61] and their X-ray crystal structure information**VIII1** (based on PDB)**VIII2****VIII3****VIII4****VIII5** (AVE, AVE5688)**VIII6** (AVF, AVE2865)**VIII7**

Compd	PDB	Ligand	IC ₅₀ (nM) GP ^b	pIC ₅₀			IC ₅₀ (nM) GP ^a	pIC ₅₀ GP ^a	CPI ^b	CMR ^b	Bind Site ^c	Ref.
				Obs	Cal ^a	Dev						
VIII1 ^d	1WUT	BN2	1900	5.72	5.72	0.00	2,000	5.70	5.03	7.17	A	[60]
VIII2	1WV1	BN5	2800	5.55	5.54	0.01	1,300	5.89	2.72	5.72	A	[60]
VIII3	1WV0	BN4	2900	5.54	5.54	0.00	2,480	5.61	2.97	5.72	A	[60]
VIII4	1WUY	BN3	1600	5.80	5.80	0.00	650	6.19	2.91	5.29	A	[60]
VIII5	3CEH	AVE	430	6.37			915	6.04			A	[52]
VIII6	3CEJ	AVF	14	7.85			24	7.62			A	[52]
VIII7	2ATI	1HU	–	–			23	7.64			A	[61]

^aCalculated using Eq. 8^bAuto-loaded utilizing the program C-QSAR^cBinding Site: A = allosteric (AMP) binding site^dThe original paper reported this compound as 7-[2,6-dichloro-4-[(2-chlorophenyl)carbonylcarbamoylamino]phenoxy]heptanoic acid with one methylene less than the structure described in PDB ID: 1WUT

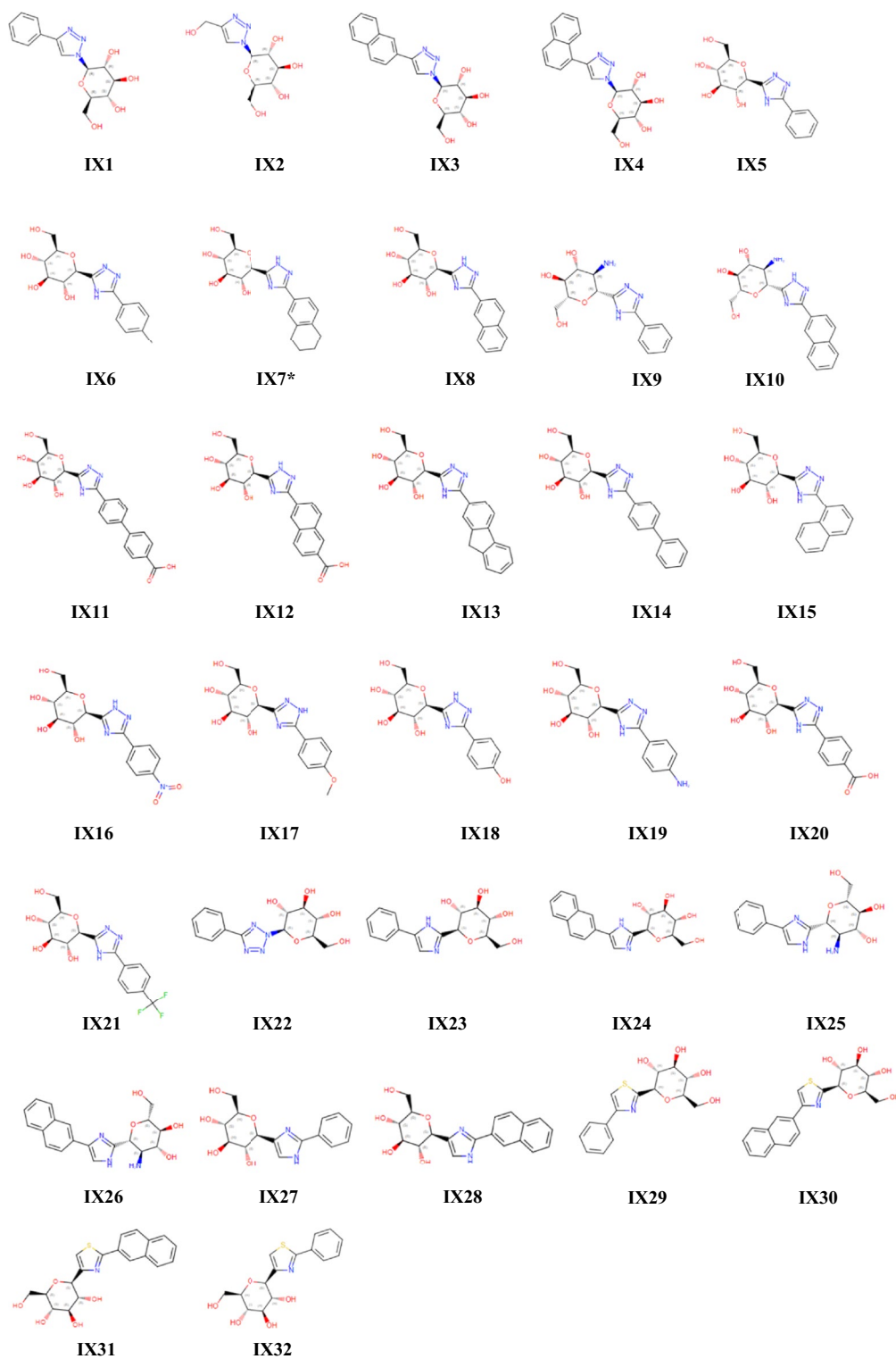
Table 11 The inhibitory potencies and the crystal structure information of β -D-glucopyranosyl triazole, pyrrole, imidazole, thiazole, tetrazole analogs, and their X-ray crystal structure information

Table 11 (continued)

Compound	PDB	Ligand	Ki (nM)	pKi			CPI ^b	CMR ^b	I _{thia} ^b	Bind Site ^c	Ref.
				Obs	Cal ^a	Dev					
IX1	3G2H	KOT	151,300	3.82	3.69	0.13	1.71	4.12	0	C	[62]
IX2	3G2I	RUG	13,700	4.86	5.08	−0.22	−1.43	2.22	0	C	[62]
IX3	3G2K	SKY	16,100	4.79	4.37	0.42	2.88	5.81	0	C	[62]
IX4	3G2L	LEW	135,900	3.87	4.37	−0.50	2.88	5.81	0	C	[62]
IX5	5LRC	73E	7000	5.16	4.71	0.45	0.84	4.12	0	C	[63]
IX6 ^d	5LRD	KS2	1700	5.77	5.29	0.48	1.34	4.58	0	C	[63]
IX7	5LRE	KS382	1600	5.80	5.39	0.41	2.02	5.81	0	C, NA	[67]
IX8 ^d	5LRF	KS3	410	6.39	6.29	0.10	2.42	5.80	0	C, NA	[63]
IX9	5O54	9LB	35,200	4.45	4.71	−0.26	0.84	4.12	0	C	[64]
IX10	5O56	9L8	4800	5.32	5.39	−0.07	2.02	5.81	0	C	[64]
IX11 ^e	6F3J	CKQ	4420	5.36	5.62	−0.26	2.48	7.28	0	C	[65]
IX12 ^d	6F3L	CJW	5050	5.30	5.68	−0.38	1.76	6.46	0	C	[65]
IX13	6F3R	CKZ	1190	5.92	5.95	−0.03	2.78	6.99	0	C	[65]
IX14	6F3S	CKW	2380	5.62	5.56	0.06	2.73	6.63	0	C	[65]
IX15	6F3U	CNK	11,500	4.94	5.39	−0.45	2.02	5.81	0	C	[65]
IX16 ^d	5OX0	B1H	33,500	4.48	4.61	−0.13	0.62	4.73	0	C	[66]
IX17	5OX1	B1K	1950	5.71	5.40	0.31	0.91	4.74	0	C	[66]
IX18	5OX3	B1N	2900	5.54	5.35	0.19	0.46	4.27	0	C	[66]
IX19	5OX4	B1W	670	6.17	6.30	−0.13	−0.11	4.49	0	C	[66]
IX20	5OWY	B0W	–				0.70	4.77	0	C	[66]
IX21	5OWZ	B0Z	111,000	3.96	4.27	−0.31	1.75	4.63	0	C	[66]
IX22	6S52	KVN	162,300	3.79	3.48	0.31	1.67	3.91	0	C	[26]
IX23 ^e	5JTT	6MY	280	6.55	6.46	0.09	1.71	4.33	0	C	[67]
IX24 ^f	5JTU	6NE	31	7.51	6.72	0.79	2.89	6.02	0	C	[67]
IX25 ^d	5O50	9L2	2670	5.57	5.63	−0.06	1.71	4.33	0	C	[64]
IX26 ^d	5O52	9LE	191	6.72	6.73	−0.01	2.89	6.02	0	C, NA	[64]
IX27	6S4H	KUQ	68,600	4.16	3.95	0.21	1.71	4.33	0	C	[26]
IX28 ^e	6S4O	KV5	4580	5.34	5.17	0.17	2.89	6.02	0	C	[26]
IX29	6S4K	KVW	310,000	3.51	3.80	−0.29	2.23	4.71	1	C	[26]
IX30 ^d	6S4P	KVE	158,000	3.80	3.80	0.00	3.40	6.40	1	C	[26]
IX31	6S4R	KVH	26,200	4.58	4.48	0.10	3.40	6.40	1	C	[26]
IX32	6S51	KVQ	326,000	3.49	3.80	−0.31	2.23	4.71	1	C	[26]

^apKi values were calculated using Eq. 9a for all except those used in Eq. 9b and 9c

^bCPI and CMR values were auto-loaded utilizing the C-QSAR program. I_{thia} is an indicator variable assigned the value of 1.0 for the four thiazole derivatives (IX29–IX32) and 0.0 for all other compounds

^cBinding site: C = catalytic site, NA = new allosteric (indole) binding site

^dpKi values were calculated using Eq. 9b

^epKi values were calculated using Eq. 9c

^fNot used in deriving Eq. 9c

*The structure of IX7 is given incorrectly as KS3 in the small molecule structure drawing of the PDB database. The correct structure was acquired from the ligand coordinates of 5LRE as well as the original paper by Kantsadi et al. [67]

equally bound at the new allosteric site and the catalytic site. Kandsami et al. [67] suggested that the primary binding site was the catalytic site. They also indicated that the binding to the new allosteric binding site might be a result of the experimental concentration (10 mM) of the inhibitor solution used for soaking the crystals. However, under the same inhibitor

concentration used for IX7 and IX8, other compounds such as IX5 and IX6 did not bind at the new allosteric site. An insightful observation that Kandsami et al. made was that the binding of IX7 and IX8 at the new allosteric site triggered a significant conformational change of this site.

Fig. 10 Binding modes of **IX1–IX32** in ligand-GPb complexes. **IX7**, **IX8**, and **IX26** (PDB ID: 5LRE, 5LRF, and 5O52) at the two separate binding sites: catalytic site (C, right) and new allosteric (indole) binding site (NA, left)

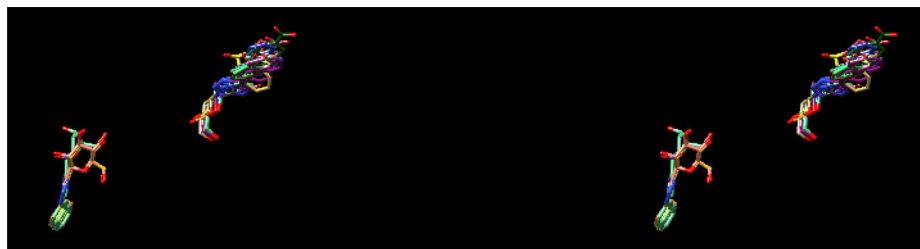


Fig. 11 a Binding modes of **X1** (three structures, orange), **X2** (five structures, green), and **X3** (blue). **b** Binding modes of **X2** (five structures, green), and **X3** (blue), and **X4** (magenta). (Supplementary Material 3)



The critical roles of CPI and CMR indicated in Eqs. 9a–9c were consistent with the explanation of Kandsami et al. When the prime binding site was the same as the other analogs, no other parameters were required in QSAR to account for the effects due to the binding at the secondary site.

There are other series of GP inhibitors whose binding modes were reported based on their ligand-bound GP X-ray crystal structures: 5-chloroindolyl derivatives (Table S4 in Supplementary Material 3) and phthalic acid derivatives and anthranilimide derivatives (Table S5 in Supplementary Material 3). The binding site of 5-chloroindolyl derivatives is the new allosteric (indole) binding site (NA) (Fig. 11) and the binding site of phthalic acid and anthranilimide derivatives is the allosteric activator (AMP) binding site (A). Interestingly, none of these inhibitors bound at the catalytic binding site. Because of their structural diversity and a limited number of compounds involved, no SAR/QSARs have been discussed, but Tables S4 and S5 were included for comparison. Further study would be possible when the binding modes of additional compounds become available.

Multiple binding sites of glycogen phosphorylase inhibitors

Among the crystal structures that we examined, 36 GP inhibitors were bound at two separate binding sites. They are summarized in Table 12. (Five single-site binders are also included for the purpose of discussion.) Except for **IV23**, **IV24**, **IV25**, **V2**, **VI5**, **VI7**, the primary binding site of these compounds is the catalytic site. Their secondary binding sites include the novel allosteric

binding site (N), the allosteric (AMP) binding site (A), the new allosteric (indole) binding site (NA), the inhibitor (purine) binding site (I), and the quercetin binding site (Q) (Fig. 12). There are only two indirubin derivatives (**V1** and **V2**) with reported ligand-bound GP crystal structures. Both compounds are bound at the inhibitor (purine) binding site (I), but **V2** is equally bound at the allosteric (AMP) binding site (A). Unlike the other five flavonoid analogs in Table 6, **VI5** and **VI7** are bound at the quercetin binding site (Q) and the allosteric (AMP) binding site (A), respectively. However, they are not bound at the catalytic binding site (C). **III15**, **VIII1–VIII17**, **VII26**, **VII29**, **VII33**, **IX7**, **IX8**, and **IX26** are equally bound at the catalytic site (C) and the new allosteric (indole) binding site (NA).

There is no apparent ligand's structural reason why some compounds are bound at more than one site. As discussed above, the effects of inhibitor binding at the secondary binding site on the correlations were not significant when the inhibitor bound at both the primary and the secondary binding sites. On the contrary, the effects were noticeable when the inhibitors bound at the secondary binding site without binding at the primary one. The latter group of compounds would ultimately end up as outliers in SAR/QSAR (for example, Eq. 4g).

Different binding modes at the allosteric binding sites of glycogen phosphorylase

Twenty-three of the 36 compounds in Table 12 are bound at the catalytic site (C) and the new allosteric (indole) binding site (NA). The binding modes of these

Table 12 Summary of GP inhibitors that are bound to more than one site discussed in this paper

No	Compd	PDB	Ligand	Primary binding site ^a	Secondary binding site ^a
1	I1	3NP7	Z15	C	N
2	II2	3GPB	GIP	C	A
3	II3	4GPB	GFP	C	A
4	II4	5GPB	GPM	C	A
5	III15	6QA6	HT8	C	NA
6	IV21	3BD7	CKB	C	I
7	IV22	3BDA	C4B	C	I
8	IV23	3BCR	AZZ	(C) ^b	I
9	IV24	3BCU	THM	(C) ^b	I
10	IV25	3BD6	RDD	(C) ^b	A
11	V2	1Z62	IAA	I	A ^c
12	VI5	4MRA	QUE	(I) ^d	Q
13	VI7	3CEM	AVD	(I) ^d	A
14	VII1	3MT9	18O	C	NA
15	VII2	3MT8	17 T	C	NA
16	VII3	3MT7	16O	C	NA
17	VII4	3MS7	22S	C	NA
18	VII5	3MS4	21 N	C	NA
19	VII6	3MS2	18S	C	NA
20	VII7	3MQF	20X	C	NA
21	VII8	3MRX	17S	C	NA
22	VII9	3MTA	22O	C	NA
23	VII10	3MTB	23 V	C	NA
24	VII11	3NC4	26O	C	NA
25	VII12	3MRT	12E	C	NA
26	VII13	3MSC	24S	C	NA
27	VII14	3MRV	16F	C	NA
28	VII15	3MTD	25E	C	NA
29	VII16	1K06 1K08 2QNB	BZD	C	NA
30	VII17	2QN8	NBY	C	NA
31	VII26	3ZCT	VMP	C	NA
32	VII29	3ZCV	N85	C	NA
33	VII33	3CUW	445	C	NA
34	IX7	5LRE	KS382	C	NA
35	IX8	5LRF	KS3	C	NA
36	IX26	5O52	9LE	C	NA

^aBinding site: C=catalytic site, NA=new allosteric (indole) binding site, A=allosteric (AMP) binding site, I=inhibitor (purine) binding site, Q=quercetin binding site, N= novel allosteric binding site. The quercetin binding site and the novel allosteric binding site are significantly overlapping

^bBinding at this site was not observed, but this site is assigned as the primary binding site because other analogs are bound at this site

^cV2 is unique in that two V2 molecules are bound at the secondary binding site (the allosteric AMP binding site)

^dBinding at this site was not observed, but this site is temporarily assigned as the primary binding site based on the binding site of their structural analogs

compounds at the catalytic site are similar and not appreciably different from the customary binding modes of structural analogs. However, the binding modes at the new allosteric (indole) binding site are relatively diverse as shown in Fig. 12. There are three distinctive binding modes (Fig. 12a–12c). These diverse binding modes are most likely due to the location of the binding site, less buried than the catalytic site.

Despite such distinctive binding modes, no other parameter was required in various allosteric QSARs examined above. The results indicated that the effects of binding at the secondary binding site were minimal when the inhibitor was equally bound at the primary binding site. On the other hand, the effects of binding at the secondary binding site were significant if the inhibitor was only bound at the secondary binding site. Such effects could be explained with the allosteric mechanism. When the inhibitor binds at the catalytic site, the access of the substrate glycogen to the catalytic site is restricted by the 280 s loop. In this manner, the binding of an inhibitor at the catalytic site stabilizes the T-state conformation of the enzyme and blocks the enzyme activity. Since the enzyme function is already reduced at this point, additional binding of the inhibitor at the allosteric site would not affect the enzyme activity further. On the other hand, when the inhibitor binds only at the allosteric binding site, the binding causes conformational changes of the enzyme by different mechanisms of action [68, 69]. Therefore, the binding at the secondary site influences the enzyme activity. This would eventually yield the outliers in SAR/QSAR as in Eq. 4g.

Table 13 summarizes the number of inhibitors bound at numerous binding sites of glycogen phosphorylase. Most of the inhibitors were bound at the catalytic site and an allosteric binding site. However, thirty-one out of 167 inhibitors (indirubin derivatives (Table 5), flavonoids (Table 6), acyl urea derivatives (Table 8), 5-chloroindolyl derivatives (Table S4), and phthalic acid and anthranilimide derivatives (Table S5)) did not bind to the catalytic site. Glucopyranosyl nucleoside derivatives (Table 4) were bound at three different binding sites including the catalytic site and two allosteric binding sites. Flavonoids (Table 6) were bound at three separate allosteric binding sites. Phthalic acid and anthranilimide derivatives (Table S5) were bound at single allosteric binding site.

Among the various allosteric binding sites, most dual-binding inhibitors preferred to bind at the new allosteric (indole) binding site, the allosteric (AMP) binding site, and the inhibitor (purine) binding site of GP in that order. Only one inhibitor each bound at the quercetin binding site and the novel allosteric binding.

Allosteric enzymes refer to the enzymes which have another site other than the active site. Allosteric enzymes

Fig. 12 The three distinctive binding modes of the structural analogs at the new allosteric (indole) binding site and location of different binding sites. **a** Binding modes of the group I compounds at the new allosteric (indole) binding site (bottom) and the catalytic site (top). Compounds included are **VIII–VIII12**, **VIII14**, **VIII15**, **IX7**, **IX8**, **IX26**. **b** Binding modes of the group II compounds at the catalytic site (top) and the new allosteric (indole) binding site (bottom). Compounds included are **III15**, **VIII13**, **VIII16** (1K06, 1K08, 2QNB), **VIII17**, **VII33**. **c** Binding modes of the group III compounds at the new allosteric (indole) binding site (top) and the catalytic site (bottom). Compounds included are **VII26** and **VII29**

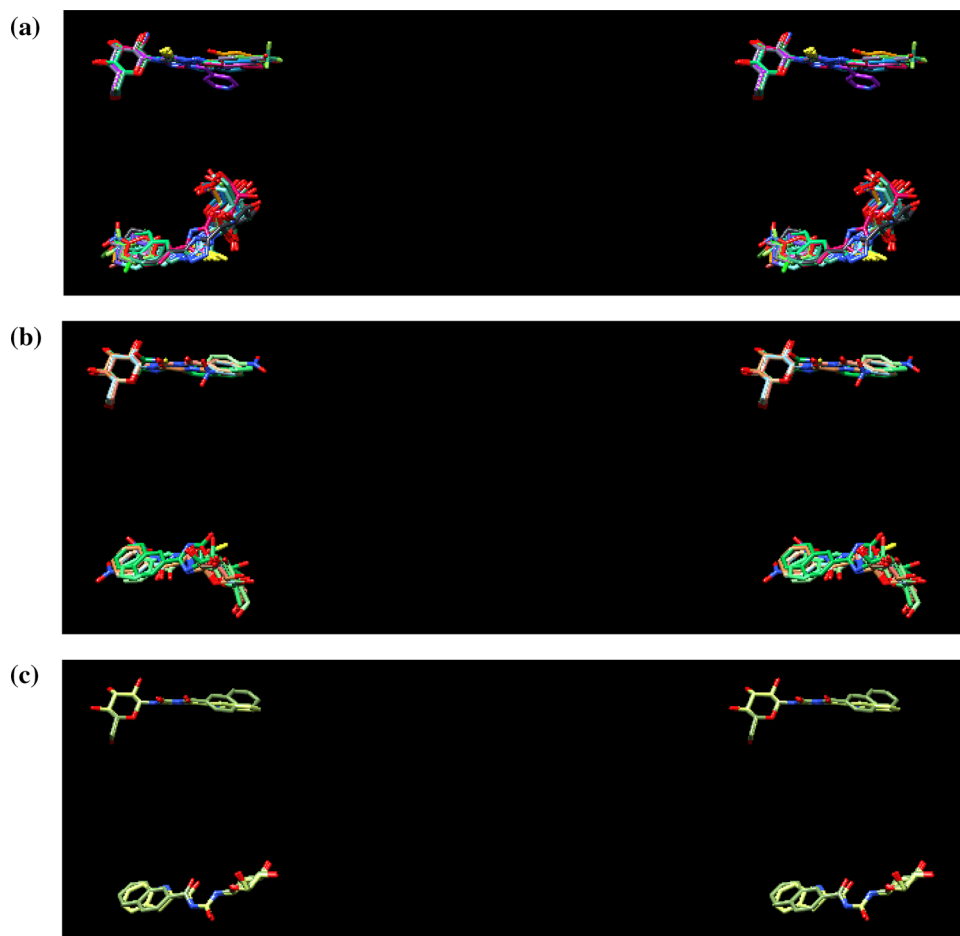
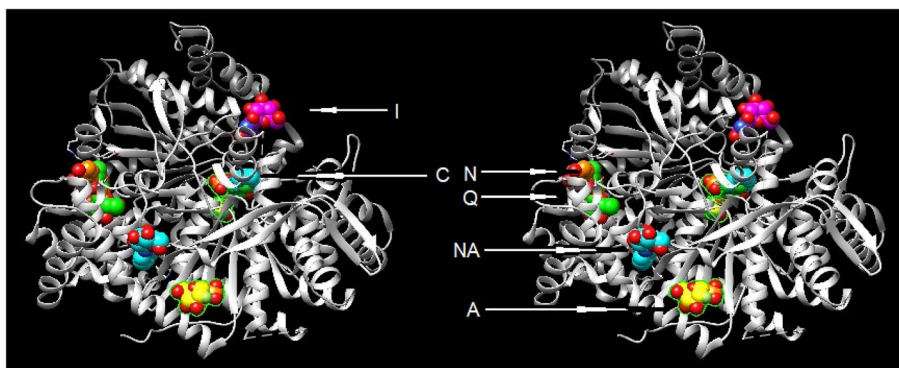


Fig. 13 Location of the five allosteric binding sites in GP: the allosteric (AMP) binding site, (A; **II3**, yellow, 4GPB), the new allosteric (indole) binding site (NA; **III15**, cyan, 6QA6), the inhibitor (purine) binding site (I; **IV21**, magenta, 3BD7), the quercetin binding site (Q; **V15**, orange, 4MRA), the novel allosteric binding site (N; green, **II**, 3NP7: Z15), and the catalytic site (C; center, between orange and magenta)



can have more than one allosteric site. Allosteric sites are different from the active site and the substrate-binding site [70]. An allosteric inhibitor is a molecule that binds to the enzyme at an allosteric site, and allosteric inhibition is a form of noncompetitive inhibition. A noncompetitive inhibitor is not directly competing with the substrate at the active site. Instead, it is indirectly altering the structure of the

enzyme. After changing the structure, the enzyme becomes inactive and does not bind with its corresponding substrate. The result is slowing down the formation of subsequent products [71].

Figure 13 shows five allosteric binding sites in GP. Table 13 shows that the inhibitors of five (Tables 5, 6, 8, S4, and S5) out of 11 series are only bound to an allosteric

Table 13 Number of inhibitors bound at different binding sites of glycogen phosphorylase examined in this study

Table	Total number of compounds	Binding site ^a					
		C	NA	A	I	Q	N
1	5	5					(1) ^b
2	6	6		3			
3	17	17	1				
4	25	22		1	4		
5	2			1	2		
6	8			1	5	1	
7a	15	15	15				
7b	14	14	4				
7c	8	8	1				
7d	12	12					
7e	5	5					
8	7			7			
9	32	32	3				
S4	4		4				
S5	7			7			
Total	167	136	28	20	11	1	(1)

^aBinding site: C=catalytic site, NA=new allosteric (indole) binding site, A=allosteric (AMP) binding site, I=inhibitor (purine) binding site, Q=quercetin binding site, N= novel allosteric binding site

^bObserved for the mixture of **11** and **12**

site: these inhibitors are single allosteric site inhibitors. Because of their structural diversity, QSAR analyses were performed only for the dataset in Tables 6 and 8, and those in Tables 5, S4, and S5 were not done. QSARs from the data in Tables 6 and 8 are both an inverse parabola correlation (Eqs. 4g and 8). Except for two flavonoid analogs (**VI6** and **VI7**) in Table 6 and **VIII5**, **VIII6**, and **VIII7** in Table 8, which were structurally diverse, all the remaining compounds whose binding modes were identical were included in the QSAR analysis and correctly identified as a single site allosteric binder. One compound (**IV5** from Table 6), which is a different allosteric site binder, was identified as an outlier in the corresponding QSAR (Eq. 4g). The outcome was not surprising because unlike all the other inhibitors, **IV5** bound at a separate binding site (quercetin binding site). The correlations of inverted parabola QSARs for the allosteric interactions were consistent with the suggestion and allosteric QSAR results of Verma and Hansch [5–10].

Dual inhibitions and their QSARs

The inhibitors of five (Tables 2, 3, 4, 7 and 9) out of 11 series described in Table 11 are bound to an allosteric site as well as the catalytic site (orthosteric site). These inhibitors are dual binders. The QSARs (Table 12) for these series of compounds include linear and normal as well as inverted parabola correlations. Equations 2a (or 2b), 3a, 3b, 5a, and 6 are normal parabola QSARs. Equations 2d and 8 are inverted

parabola correlations, suggesting this is allosteric QSAR. All other equations are linear correlations (Table 14).

It is worth noting that both normal and inverse parabola correlations are included in describing these dual inhibitions. Since linear and/or normal parabola QSARs are reported from various correlation studies, and inverted parabola QSARs are seen from allosteric inhibition studies, QSARs of all such forms are deemed natural to describe dual allosteric inhibitions. Ultimately, these QSARs can contain outliers observed in many QSARs [1, 2, 4] that are due to various possible reasons (Ref. [2] and the references cited therein).

Sharma and Gupta [72, 73] reported normal and inverse parabolic relationships with CMR in several inhibitor series, suggesting a dual allosteric binding mode in glycine/NMDA antagonism. They proposed that some molecules may be altering the shape of the active site residues, leading to normal and inverted allosteric correlations. They supported their suggestions based on a molecular docking simulation study [73], unlike this study based on the inhibitor-bound enzyme X-ray structures. Several recent studies indicated an effective dual-targeting therapeutic mechanism involving allosteric and orthosteric binding sites [74–80].

Normal parabolic or bilinear correlations for allosteric interactions

Verma and Hansch [5] reported two allosteric inverted parabolic QSARs (Eqs. 11a and 12b) with GP inhibitors as

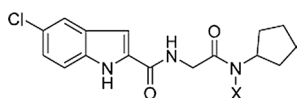
Table 14 Summary of QSAR of GP inhibitors described in this study

Table	QSAR	<i>n</i>	<i>r</i> ²	<i>s</i>	Eq.
1	pKi = 0.99(± 0.47) CPI + 2.65(± 0.60)	5	0.94	0.149	1
3	pKi = 4.15(± 3.55) MolVol - 0.85(± 0.82) MolVol ² + 0.01(± 3.76)	9	0.77	0.143	2a
	pKi = 0.79(± 0.68) CMR - 0.04(± 0.04) CMR ² + 1.51(± 2.58)	9	0.75	0.148	2b
	pKi = -7.80(± 6.47) CMR + 0.56(± 0.45) CMR ² + 30.30(± 22.35)	6	0.88	0.537	2d
4	pKi = -0.07(± 0.06) CPI ² - 2.83(± 0.38) I + 5.23(± 0.22)	14	0.96	0.287	3a
	pKi = -0.58(± 0.65) CPI ² + 5.30(± 1.41)	5	0.73	0.835	3b
6	pKi _(rmGPb) = 0.28(± 0.15) CMR ² + 3.32(± 1.14)	6	0.87	0.167	4g
7a	pIC ₅₀ = 2.18(± 1.30) CPI - 0.72(± 0.46) CPI ² - 1.30(± 0.50) Iz + 3.17(± 0.82)	11	0.87	0.259	5a
	pIC ₅₀ = -9.54(± 8.55) CMR + 62.05(± 52.49)	3	1.00	0.023	5c
7b	pKi = 8.61(± 2.79) CMR - 0.74(± 0.25) CMR ² - 18.74(± 7.52)	10	0.92	0.193	6
7c	pKi = 0.19(± 0.17) CMR + 2.93(± 0.70)	7	0.63	0.291	7a
7d	pKi = 0.21(± 0.14) CMR - 1.67(± 0.41) I _{oxamide} + 4.44(± 0.79)	20	0.91	0.355	7b
	pKi = 0.20(± 0.08) CMR - 1.42(± 0.32) I _{oxamide} - 1.53(± 0.40) I _{flex} + 4.23(± 0.44)	31	0.87	0.346	7c
	pKi = 0.22(± 0.10) CPI - 1.31(± 0.36) I _{oxamide} - 1.55(± 0.42) I _{flex} + 4.83(± 0.26)	31	0.86	0.366	7d
8	pKi = -4.72(± 1.83) CMR + 0.38(± 0.15) CMR ² + 20.24(± 5.68)	4	1.00	0.007	8
9	pKi = -1.18(± 0.24) CPI + 1.23(± 0.24) CMR + 0.67(± 0.91)	20	0.88	0.318	9a
	pKi = 0.94(± 0.50) CPI - 3.41(± 1.25) I _{thiazole} + 4.03(± 0.96)	7	0.94	0.319	9b
	pKi = -1.10(± 4.64) CPI + 8.34(± 11.17)	3	0.89	0.324	9c

shown in datasets 1 and 5 of Table S2. The compounds they studied were 5-chloroindolyl derivatives which belong to Table S4. Four compounds (**X1–X4**) are listed in Table S4 and are all bound at the new allosteric (indole) binding (NA) site. It was and is still not known whether any of the structural analogs that Verma and Hansch reported concerning the allosteric QSARs bound at more than one binding site or at a separate binding site.

They included all the compounds of the corresponding structures reported by Wright et al. [81] without considering their binding sites. Interestingly, the authors reported an additional normal parabolic QSAR (Eqs. 11b and 12a). In each case, even though these equations were not mentioned as allosteric QSARs, a single parameter Eq. 11c was additionally included. There was one final outlier in each example after the ‘splitting QSAR’ development.

Inhibition of glycogen phosphorylase A (GPA, EC 2.4.1.1)



by 5-chloroindolyl derivatives I

$$\log 1/C = -4.96(\pm 2.61) \text{CMR} + 0.20(\pm 0.12) \text{CMR}^2 + 36.54(\pm 14.44)$$

$$n = 21, r^2 = 0.855, q^2 = 0.819, s = 0.193$$

$$\text{inversion point for CMR} = 12.38(11.85-4.30)$$
(11a)

$$\log 1/C = 28.01(\pm 6.42) \text{C log P} - 4.19(\pm 0.95) \text{C log P}^2 - 39.39(\pm 10.74)$$

$$n = 10, r^2 = 0.940, q^2 = 0.866, s = 0.249$$

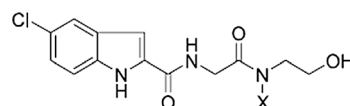
$$\text{optimum C log P} = 3.34(3.28-3.40)$$
(11b)

$$\log 1/C = -0.81(\pm 0.41) \text{C log P} - 3.01(\pm 1.55)$$

$$n = 5, r^2 = 0.931, q^2 = 0.831, s = 0.215$$

$$\text{outlier} = \text{CH}_2\text{CH}_2\text{CH}_2\text{N}(\text{CH}_3)_2$$
(11c)

Inhibition of glycogen phosphorylase A (GPA, EC 2.4.1.1) by 5-chloroindolyl derivatives II



$$\log 1/C = 0.62(\pm 0.13) \text{C log P} + 4.10(\pm 1.10) \text{CMR} - 0.23(\pm 0.06) \text{CMR}^2 - 13.03(\pm 5.60)$$

$$n = 20, r^2 = 0.881, q^2 = 0.819, s = 0.172$$

$$\text{optimum CMR} = 8.92(8.48-9.25)$$
(12a)

$$\log 1/C = -1.52(\pm 0.39) \text{C log P} + 0.38(\pm 0.10) \text{C log P}^2 + 7.74(\pm 0.32)$$

$$n = 7, r^2 = 0.967, q^2 = 0.926, s = 0.131$$

$$\text{inversion point for C log P} = 2.00(1.85-2.18)$$

$$\text{outlier} = 3\text{-Tetrahydrofuryl}$$
(12b)

Hansch's group reported inverted parabola or bilinear QSAR correlations to indicate the allosteric interactions and change of conformations involved. In our recent search of the C-QSAR database for all the reported QSAR equations using three particular search queries (carbonic anhydrase, elastase inhibitor, and rhinovirus inhibitor) [2], 270 equations were retrieved. Among the 270 equations, 19 equations were inverted parabolic or bilinear correlations (Table S1) and 43 equations were normal parabolic or bilinear correlations. Supuran [82] reported one of the carbonic anhydrase inhibition mechanisms represented an allosteric interaction with conformational change suggested based on the crystal structure. Our results presented here also show that in addition to the inverted parabola QSARs, the normal parabola QSARs and the linear parameter QSARs can equally describe dual allosteric interactions.

In the study of the allosteric site of muscarinic acetylcholine M₂-receptors, Bender et al. [83] reported a QSAR with a significant correlation between the volume of the substituents and the allosteric potency. One significant point to note about their allosteric correlation is that their QSAR is a normal parabola, not an inverted one. The allosteric potencies of the compounds they studied cover more than two orders of magnitude, and the dataset was suitable to establish a QSAR.

In another report, Sharma and Gupta [72] examined several sets of compounds as selective glycine/NMDA (N-Methyl-D-aspartic acid) site antagonists and reported ten QSARs suggesting dual allosteric binding interactions. Interestingly, two alternative forms of allosteric QSARs were reported: normal and inverted parabola. Among the ten QSARs, three were normal parabolic correlations (with CMR) and three were inverted parabolic correlations (one with ClogP and two with CMR). Additionally, there were other linear parameter equations with CMR or CPI. Unlike the suggestions made by Verma and Hansch, they described the normal parabola relationship for the allosteric interactions as well.

It was assumed that at the inversion point the structure of a receptor is forced to change into a new shape. This would result in an altogether different type of interaction. Another possibility would be that there is more than one binding site. In such a case the ligand should not bind with the same parameters defined in the first half of the equation [8].

Hansch et al. [6] reported 60 examples of the normal parabolic or bilinear QSARs with CMR and 27 examples with MgVol from their QSAR database. Even so, they were not certain if any of those results involved allosteric effects. They suggested other researchers should check such cases for the possibility of allostery and the role of QSARs in rationalizing such results [18].

Based on the various reports described above, as well as our results presented here, it is clear that normal parabola/bilinear QSARs, linear parameters, and inverted QSARs

can all describe allosteric interactions, especially in the case of dual allosteric interactions. Since normal parabola/bilinear QSARs are frequently observed in QSARs for various biological activities, the key QSARs that can be used to uncover allosteric interactions are the inverted parabola/bilinear QSARs, as suggested by Hansch and his co-workers. Even though this study confirmed Hansch et al.'s attribution of inverted parabolic/bilinear QSAR to the allosteric ligand-binding mechanism, additional studies with other allosteric binders and proteins (including dual binders) with experimental binding information would further confirm and can firmly generalize this point.

Conclusion

We examined over 200 X-ray crystal structures of the ligand-bound allosteric enzyme glycogen phosphorylase. The QSAR analyses of the inhibitors resulted in the inverted parabola correlations in several cases. In addition, we obtained the normal parabola as well as linear correlations. These results indicated that linear, normal parabola/bilinear and inverted parabola/bilinear correlations could all describe the allosteric interactions, particularly dual allosteric interactions. In many cases, the binding of various allosteric inhibitors accompanied the conformational change. This study supported Hansch and his co-workers' proposal that inverted parabola/bilinear QSARs describe the allosteric interactions and such QSARs could be used to uncover such allosteric interactions.

The crystal structures revealed many ligands bound at more than one binding site of the enzyme. Some compounds were bound at the secondary binding site only and not at the primary binding site where most other structural analogs were bound. It was not apparent at present why these compounds bound more than one binding site, unlike their close structural analogs. We initially expected that compounds bound at an uncommon secondary binding site would be outliers in QSAR. On the contrary, the results revealed that the effects of binding at the secondary binding site on many SAR/QSARs were not significant when the inhibitor was equally bound at the primary binding site. However, the effects were noticeable when the inhibitors bound at the secondary binding site without binding at the primary binding site. We proposed such a phenomenon could be explained with the allosteric mechanism. We also suggested compounds belonging to the latter group would eventually end up as outliers in SAR/QSAR of that series. As in Eq. 4g, QSAR analysis may be able to identify such compounds as outliers.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11030-021-10365-6>.

Acknowledgements The author expresses sincere gratitude to Dr. Albert Leo and Mr. Michael Medlin for their generous permission to use the C-QSAR and Bio-Loom programs. He dedicates this paper to Professor Gary L. Grunewald, who was the greatest graduate research advisor at the Department of Medicinal Chemistry, University of Kansas. His leadership and friendship along with his genuine concern for the author's research career as well as his personal life are deeply appreciated. This paper was written honoring Prof. Grunewald, a respected teacher and highly regarded medicinal chemist. (Special issue in honor of Professor Gary L. Grunewald: *Med. Chem. Res.* 30(7), 2021.)

References

- Kim KH (2007) Outliers in SAR and QSAR: is unusual binding mode a possible source of outliers? *J Comput Aided Mol Des* 21:63–86
- Kim KH (2021) Outliers in SAR and QSAR: 3. Importance of considering the role of water molecules in protein-ligand interactions and quantitative structure-activity relationship study. *J Comput-Aided Mol Des* 35:371–396
- Kurup A (2003) C-QSAR: a database of 18,000 QSARs and associated biological and physical data. *J Comput Aided Mol Design* 17:187–196
- Kim KH (2007) Outliers in SAR and QSAR: 2. Is a flexible binding site a possible source of outliers? *J Comput Aided Mol Des* 21:421–435
- Verma RP, Hansch C (2005) An approach toward the problem of outliers in QSAR. *Bioorg Med Chem* 13:4597–4621
- Hansch C, Garg R, Kurup A, Mekapati SB (2001) Searching for allosteric effects via QSARs. *Bioorg Med Chem* 9(2):283–289
- Hansch C, Garg R, Kurup A, Mekapati SB (2003) Allosteric interactions and QSAR: on the role of ligand hydrophobicity. *Bioorg Med Chem* 11(9):2075–2084
- Garg R, Kurup A, Mekapati SB, Hansch C (2003) Searching for allosteric effects via QSAR. Part II *Bioorg Med Chem* 11(4):621–628
- Garg R, Bhatarai B (2008) Possible allosteric interactions of monoindazole-substituted P2 cyclic urea analogues with wild-type and mutant HIV-1 protease. *J Comput-Aided Mol Des* 22:737–745
- Hansch C, Hoekman D, Leo A, Weininger D, Selassie CD (2002) Chem-Bioinformatics: comparative QSAR at the Interface between chemistry and biology. *Chem Rev* 102:783–812
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. *Nucleic Acids Res* 28:235–242
- Burley SK, Berman HM, Bhikadiya C, Bi C, Chen L, Costanzo LD, Christie C, Dalenberg K, Duarte J, Dutta S, Feng Z, Ghosh S, Goodsell DS, Green RK, Guranovic V, Guzenko D, Hudson B, Kalro T, Liang Y, Lowe R, Namkoong H, Peisach E, Periskova P, Prlic A, Randle C, Rose A, Rose P, Sala R, Sekharan M, Shao C, Tan L, Tao Y-P, Valasatava Y, Voigt M, Westbrook J, Woo J, Yang H, Young J, Zhuravleva Z, Zardecki C (2019) RCSB protein data bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res* 47:D464–D474
- Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, Lopez R (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* 47:W636–W641
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—A visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612
- Leo A, Michael LM. BioByte: 201 W. 4th St., #204, Claremont, CA 91711–4707. clogp@biobyte.com. 909–624–5992
- Deng Q, Lu Z, Bohn J, Ellsworth KP, Myers RW, Geissler WM, Harris G, Willoughby CA, Chapman K, McKeever B, Mosley R (2005) Modeling aided design of potent glycogen phosphorylase inhibitors. *J Mol Graph Model* 23:457–464
- Liu X, Lu S, Song K, Shen Q, Ni D, Li Q, He X, Zhang H, Wang Q, Chen Y, Li X, Wu J, Sheng C, Chen G, Liu Y, Lu X, Zhang J (2020) Unraveling allosteric landscapes of allosterome with ASD. *Nucleic Acids Res* 48:D394–D401
- Garg R, Kurup A, Hansch C (2001) Possible allosteric effects in anticancer compounds. *Bioorg Med Chem* 9(12):3161–3164
- Kurup A, Garg R, Hansch C (2001) Comparative QSAR study of tyrosine kinase inhibitors. *Chem Rev* 101(8):2573–2600
- Mekapati SB, Hansch C (2001) Comparative QSAR studies on bibenzimidazoles and terbenzimidazoles inhibiting topoisomerase I. *Bioorg Med Chem* 9(11):2885–2893
- Mekapati SB, Kurup A, Verma RP, Hansch C (2005) The role of hydrophobic properties of chemicals in promoting allosteric reactions. *Bioorg Med Chem* 13(11):3737–3762
- Mathieu C, Dupret JM, Lima FR (2017) The structure of brain glycogen phosphorylase—from allosteric regulation mechanisms to clinical perspectives. *FEBS J* 284:546–554
- Oikonomakos NG, Tiraidis C, Leonidas DD, Zographos SE, Kristiansen M, Jessen CU, Norskov-Lauritsen L, 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
- Newgard CB, Hwang PK, Fletterick RJ (1989) The family of glycogen phosphorylases: structure and function. *Critical Rev Biochem Molecular Biol* 24:69–99
- Barford D, Johnson LN (1992) The molecular mechanism for the tetrameric association of glycogen phosphorylase promoted by protein phosphorylation. *Protein Sci* 1:472–491
- Kyriakis E, Karra AG, Papaioannou O, Solovou T, Skamnaki VT, Liggri PGV, Zographos SE, Szennyes E, Bokor E, Kun S, Psarra A-MG, Somsak L, Leonidas DD (2020) The architecture of hydrogen and sulfur s-hole interactions explain differences in the inhibitory potency of C- β -D-glucopyranosyl thiazoles, imidazoles and an N- β -D-glucopyranosyl tetrazole for human liver glycogen-D phosphorylase and offer new insights to structure-based design. *Bioorg Med Chem* 28:115196–115205
- Christopoulos A (2002) Allosteric binding sites on cell-surface receptors: novel targets for drug discovery. *Nat Rev Drug Disc* 1:198–210
- Chetter BA, Kyriakis E, Barr D, Karra AG, Katsidou E, Koulas SM, Skamnaki VT, Snape TJ, Psarra AG, Leonidas DD, Hayes JM (2020) Synthetic flavonoid derivatives targeting the glycogen phosphorylase inhibitor site: QM/MM-PBSA motivated synthesis of substituted 5,7-dihydroxyflavones, crystallography, in vitro kinetics and ex-vivo cellular experiments reveal novel potent inhibitors. *Bioorg Chem* 102:104003–104003
- Szabo KE, Kyriakis E, Psarra AG, Karra AG, Sipos A, Docsa T, Stravodimos GA, Katsidou E, Skamnaki VT, Liggri PGV, Zographos SE, Mandi A, Kiraly SB, Kurtan T, Leonidas DD, Somsak L (2019) Glucopyranosylidene-spiro-imidazolinones, a new ring system: synthesis and evaluation as glycogen phosphorylase inhibitors by enzyme kinetics and x-ray crystallography. *J Med Chem* 62:6116–6136
- Alexacou KM, Zhang YZ, Praly JP, Zographos SE, Chrysina ED, Oikonomakos NG, Leonidas DD (2011) Halogen-substituted (C- β -D-glucopyranosyl)-hydroquinone regioisomers: Synthesis, enzymatic evaluation and their binding to glycogen phosphorylase. *Bioorg Med Chem* 19:5125–5136
- He L, Zhang YZ, Tanoh M, Chen GR, Praly JP, Chrysina ED, Tiraidis C, Kosmopoulou M, Leonidas DD, Oikonomakos NG

- (2007) In the Search of glycogen phosphorylase inhibitors: synthesis of C-D-glycopyranosylbenzo(hydro)quinones. Inhibition of and binding to glycogen phosphorylase in the crystal. *Eur J Org Chem* 2007:596–606
32. Martin JL, Johnson L, Withers SG (1990) Comparison of the binding of glucose and glucose 1-phosphate derivatives to T-state glycogen phosphorylase b. *Biochem* 29(48):10745–10757
 33. Johnson LN, Acharya KR, Jordan MD, McLaughlin PJ (1990) Refined crystal structure of the phosphorylase-heptulose 2-phosphate-oligosaccharide-AMP complex. *J Mol Biol* 211:645–661
 34. Chrysina ED, Kosmopoulou MN, Kardakaris R, Bischler N, Leonidas DD, Kannan T, Loganathan D, Oikonomakos NG (2005) Binding of beta-d-glucopyranosyl bismethoxyphosphoramidate to glycogen phosphorylase b: kinetic and crystallographic studies. *Bioorg Med Chem* 13:765–772
 35. Bentifa M, Hayes JM, Vidal S, Gueyrard D, Goekjian PG, Praly JP, Kizilis G, Tiraidis C, Alexacon KM, Chrysina ED, Zographos SE, Leonidas DD, Archontis G, Oikonomakos NG (2009) Glucose-based spiro-isoxazolines: a new family of potent glycogen phosphorylase inhibitors. *Bioorg Med Chem* 17:7368–7380
 36. Watson KA, Chrysina ED, Tsitsanou KE, Zographos SE, Archontis G, Fleet GWJ, Oikonomakos NG (2005) Kinetic and crystallographic studies of glucopyranose spirohydantoin and glucopyranosylamine analogs inhibitors of glycogen phosphorylase. *PROTEINS Struct Funct Bioinform* 61(4):966–983
 37. Czifrak K, Deak S, Pahi A, Kover KE, Docsa T, Gergely P, Alexacou KM, Papakonstantinou M, Leonidas DD, Zographos SE, Chrysina ED, Somsak L (2014) Glucopyranosylidene-Spiro-Iminothiazolidinone, a new bicyclic ring system: synthesis, derivatization, and evaluation for inhibition of glycogen phosphorylase by enzyme kinetic and crystallographic methods. *Bioorg Med Chem* 22(15):4028–4041
 38. Oikonomakos NG, Shamnaki VT, Osz E, Szilagyi L, Somsak L, Docsa T, Toth B, Gergely P (2002) Kinetic and crystallographic studies of glucopyranosylidene spirothiohydantoin binding to glycogen phosphorylase b. *Bioorg Med Chem* 10:261–268
 39. Gregoriou M, Noble MEM, Watson KA, Garman EF, Johnson LN, Krulle TM, De La Fuente C, Fleet GWJ, Oikonomakos NG (1998) The structure of a glycogen phosphorylase glucopyranose spirohydantoin complex at 1.8 Å resolution and 100 K: the role of the water structure and its contribution to binding. *Protein Sci* 7:915–927
 40. Abboud JLM, Mo O, de Paz JLG, Yanez M, Esseffar M, Bouab W, El-Mouhtadi M, Mokhlisse R, Ballesteros E, Herreros M, Homan H, Lopez-Mardomingo C, Notario R (1993) Thiocarbonyl versus carbonyl compounds: a comparison of intrinsic reactivities. *J Am Chem Soc* 115:12468–12476
 41. Tsirkone VG, Tsoukala E, Lamprakos C, Manta S, Hayes JM, Skamnaki VT, Drakou C, Zographos SE, Komiotis D, Leonidas DD, (2010) 1-(3-Deoxy-3-fluoro-beta-d-glucopyranosyl) pyrimidine derivatives as inhibitors of glycogen phosphorylase b: Kinetic, crystallographic and modelling studies. *Bioorg Med Chem* 18:3413–3425
 42. Kantsadi AL, Hayes JM, Manta S, Skamnaki VT, Kiritsis C, Psarra AMG, Koutsogiannis Z, Kontou M, Papadopoulos G, Zographos SE, Komiotis D, Leonidas DD (2012) The σ -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(4):722–732
 43. Manta S, Xipnitou A, Kiritsis C, Kantsadi AL, Hayes JM, Skamnaki VT, Lamprakos C, Kontou M, Zoumpoulakis P, Zographos SE, Leonidas DD, Komiotis D (2012) 3'-Axial CH(2) OH substitution on glucopyranose does not increase glycogen phosphorylase inhibitory potency. *QM/MM-PBSA calculations suggest why*. *Chem Biol Drug Des* 79:663–673
 44. Kantsadi AL, Manta S, Psarra AM, Dimopoulou A, Kiritsis C, Parmenopoulou V, Skamnaki VT, Zoumpoulakis P, Zographos SE, Leonidas DD, Komiotis D (2012) The binding of C5-alkynyl and alkylfurano[2,3-d]pyrimidine glucopyranonucleosides to glycogen phosphorylase b: synthesis, biochemical and biological assessment. *Eur J Med Chem* 54:740–749
 45. Mamais M, Esposti AD, Kouloumoundra V, Gustavsson T, Monti F, Venturini A, Chrysina ED, Markovitsi D, Gimisis T (2017) A new potent inhibitor of glycogen phosphorylase reveals the basicity of the catalytic site. *Chem Eur J* 23:8800–8805
 46. Cisma C, Sovantzis DA, Hadjiloi T, Stathis D, Gimisis T, Hayes JM, Zographos SE, Leonidas DD, Chrysina ED, Oikonomakos NG. D-Glucopyranosyl pyrimidine nucleoside binding to muscle glycogen phosphorylase b. <https://www.rcsb.org/structure/3BCS>
 47. Kosmopoulou MN, Leonidas DD, Chrysina ED, Bischler N, Eisenbrand G, Sakarellos CE, Pauptit R, Oikonomakos NG (2004) Binding of the potential antitumour agent indirubin-5-sulphonate at the inhibitor site of rabbit muscle glycogen phosphorylase b. Comparison with ligand binding to pCDK2-cyclin A complex. *Eur J Biochem* 271:2280–2290
 48. Kosmopoulou MN, Leonidas DD, Chrysina ED, Eisenbrand G, Oikonomakos NG (2005) Indirubin-3'-Aminoxy-Acetate inhibits glycogen phosphorylase by binding at the inhibitor and the allosteric site. Broad specificities of the two sites. *Lett Drug Design Discov* 2:377–390
 49. Tsitsanou KE, Hayes JM, Keramioti M, Mamais M, Oikonomakos NG, Kato A, Leonidas DD (2013) Zographos SE. Sourcing the affinity of flavonoids for the glycogen phosphorylase inhibitor site via crystallography, kinetics and QM/MM-PBSA binding studies: comparison of chrysin and flavopiridol. *Food Chem Toxicol* 61:14–27
 50. Oikonomakos NG, Schnier JB, Zographos SE, Skamnaki VT, Tsitsanou KE, Johnson LN (2000) Flavopiridol inhibits glycogen phosphorylase by binding at the inhibitor site. *J Biol Chem* 275:34566–34573
 51. Kantsadi AL, Apostolou A, Theofanous S, Stravodimos GA, Kyriakis E, Gorgogietas VA, Chatzileontiadou DS, Pegiou K, Skamnaki VT, Stagos D, Kouretas D, Psarra AM, Haroutounian SA, Leonidas DD (2014) Biochemical and biological assessment of the inhibitory potency of extracts from vinification byproducts of *Vitis vinifera* extracts against glycogen phosphorylase. *Food Chem Toxicol* 67:35–43
 52. Anderka O, Loenze P, Klabunde T, Dreyer MK, Defossa E, Wendt KU, Schmoll D (2008) Thermodynamic characterization of allosteric glycogen phosphorylase inhibitors. *Biochem* 47(16):4683–4691
 53. Kato A, Nasu N, Takebayashi K, Adachi I, Minami Y, Sanae F, Asano N, Watson AA, Nash RJ (2008) Structure-activity relationships of flavonoids as potential inhibitors of glycogen phosphorylase. *J Agric Food Chem* 56:4469–4473
 54. Alexacou KM, Tenchiu AC, Chrysina ED, Charavgi MD, Kostas ID, Zographos SE, Oikonomakis NG, Leonidas DD (2010) The binding of β -d-glucopyranosyl-thiosemicarbazone derivatives to glycogen phosphorylase: a new class of inhibitors. *Bioorg Med Chem* 18(22):7911–7922
 55. Chrysina ED, Nagy V, Felföldi N, Konya B, Telepo K, Praly JP, Docsa T, Gergely P, Alexacou KM, Hayes JM, Konstantakaki M, Kardakaris R, Leonidas DD, Zographos SE, Oikonomakos NG, Somsak L. Synthesis, Kinetic, Computational and Crystallographic Evaluation of N-Acyl-N-Beta-D- Glucopyranosyl)Ureas, Nanomolar Glucose Analogue Inhibitors of Glycogen Phosphorylase, Potential Antidiabetic Agents. <https://www.rcsb.org/structure/3ZCT>
 56. Oikonomakos NG, Kosmopoulou M, Zographos SE, Leonidas DD, Chrysina ED, Somsak L, Nagy V, Praly JP, Docsa T, Toth B, Gergely P (2002) Binding of N-acetyl-N'-beta-D-glucopyranosyl

- urea and N-benzoyl-N'-beta-D-glucopyranosyl urea to glycogen phosphorylase b: kinetic and crystallographic studies. *Eur J Biochem* 269:1684–1696
57. Nagy V, Felfoldi N, Praly JP, Docsa T, Gergerly P, Chrysina ED, Tiraidis C, Kosmopoulou MN, Alexacou K-M, Konstantakaki M. N-(4-substituted-benzoyl)-N'-(beta-D-glucopyranosyl)ureas, inhibitors of glycogen phosphorylase: synthesis, kinetic and crystallographic evaluation. <https://www.rcsb.org/structure/3BCS>
 58. Czifrak K, Kyriaki C, Felfoldi N, Docsa T, Gergely P, Chrysina ED, Siafaka-Kapadai A, Leonidas DD, Zographos SE, Oikonomakos NG. Tracking structure activity relationships of glycogen phosphorylase inhibitors: synthesis, kinetic and crystallographic evaluation of analogues of N-(D-glucopyranosyl)-N'-oxamides. <https://www.rcsb.org/structure/2QN3>
 59. Hadjiloi T, Tiraidis C, Chrysina ED, Leonidas DD, Oikonomakos NG, Tsipos P, Gimisis T (2006) Binding of oxalyl derivatives of beta-d-glucopyranosylamine to muscle glycogen phosphorylase b. *Bioorg Med Chem* 14:3872–3882
 60. Oikonomakos NG, Kosmopoulou MN, Chrysina ED, Leonidas DD, Kostas ID, Wendt KU, Klabunde T, Defossa E (2005) Crystallographic studies on acyl ureas, a new class of glycogen phosphorylase inhibitors, as potential antidiabetic drugs. *Protein Sci* 14:1760–1771
 61. Klabunde T, Wendt KU, Kadereit D, Brachvogel V, Burger H-J, Herling AW, Oikonomakos NG, Kosmopoulou MN, Schmoll D, Sarubbi E, von Roedem E, Schnafinger K, Defossa E (2005) Acyl ureas as human liver glycogen phosphorylase inhibitors for the treatment of type 2 diabetes. *J Med Chem* 48(20):6178–6193
 62. Chrysina ED, Bokor E, Alexacou K-M, Charavgi M-D, Oikonomakos GN, Zographos SE, Leonidas DD, Oikonomakos NG, Somsak L (2009) Amide-1,2,3-triazole bioisosterism: the glycogen phosphorylase case. *Tetrahedron Asymmetry* 20:733–740
 63. Kantsadi AL, Stravodimos GA, Kyriakis E, Chatzileontiadou DSM, Solovou TGA, Kun S, Bokor E, Somsak L, Leonidas DD (2017) van der Waals interactions govern C-β-d-glucopyranosyl triazoles' nM inhibitory potency in human liver glycogen phosphorylase. *J Struct Biol* 199(1):57–67
 64. Bokor E, Kyriakis E, Solovou TGA, Koppany C, Kantsadi AL, Szabo KE, Szakacs A, Stravodimos GA, Docsa T, Skamnaki VT, Zographos SE, Gergely P, Leonidas DD, Somsak L (2017) Nanomolar inhibitors of glycogen phosphorylase based on beta-d-glucosaminyl heterocycles: a combined synthetic, enzyme kinetic, and protein crystallography study. *J Med Chem* 60:9251–9262
 65. Kun S, Begum J, Kyriakis E, Stamati ECV, Barkas TA, Szenyies E, Bokor E, Szabo KE, Stravodimos GA, Sipos A, Docsa T, Gergely P, Moffatt C, Patraskaki MS, Kokolaki MC, Gkerdi A, Skamnaki VT, Leonidas DD, Somsak L, Hayes JM (2018) A multidisciplinary study of 3-(beta-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
 66. Kyriakis E, Solovou TGA, Kun S, Czifrak K, Szocs B, Juhasz L, Bokor E, Stravodimos GA, Kantsadi AL, Chatzileontiadou DSM, Skamnaki VT, Somsak L, Leonidas DD (2018) Probing the β-pocket of the active site of human liver glycogen phosphorylase with 3-(C-β-d-glucopyranosyl)-5-(4-substituted-phenyl)-1, 2, 4-triazole inhibitors. *Bioorg Chem* 77:485–493
 67. Kantsadi AL, Bokor E, Kun S, Stravodimos GA, Chatzileontiadou DSM, Leonidas DD, Juhasz-Toth E, Szakacs A, Batta G, Docsa T, Gergely P, Somsak L (2016) Synthetic, enzyme kinetic, and protein crystallographic studies of C-β-d-glucopyranosyl pyrroles and imidazoles reveal and explain low nanomolar inhibition of human liver glycogen phosphorylase. *Eur J Med Chem* 123:737–745
 68. Nussinov R, Tsai C-J (2012) The different ways through which specificity works in orthosteric and allosteric drugs. *Curr Pharm Des* 18(9):1311–1316
 69. Diez MDAA, Leanzi MC, Iglesias AA, Ballicora MA (2014) A novel dual allosteric activation mechanism of escherichia coli ADP-glucose pyrophosphorylase: the role of pyruvate. *PLoS One* 9(8):e103888
 70. A Level Chemistry, Allosteric Enzymes. <https://alevelchemistry.co.uk/definition/allosteric-enzymes/>
 71. Thienthanh (2020) Allosteric Inhibition, in *Biology Dictionary*. <https://biologydictionary.net/allosteric-inhibition/>
 72. Sharma M, Gupta VB (2010) Dual allosteric effect in glycine/NMDA receptor antagonism: a comparative QSAR approach. *Pharmaceuticals* 3(10):3167–3185
 73. Sharma M, Gupta VB (2012) Dual allosteric effect in glycine/NMDA receptor antagonism: a molecular docking simulation approach. *Intern J Drug Design Disc* 3:718–730
 74. Qiu Y, Yin X, Li X, Wang Y, Fu Q, Huang R, Lu S (2021) Untangling dual-targeting therapeutic mechanism of epidermal growth factor receptor (EGFR) based on reversed allosteric communication. *Pharmaceutics* 13:747–774
 75. Huang J, Yuan Y, Zhao N, Pu D, Tang Q, Zhang S, Luo S, Yang X, Wang N, Xiao Y, Zhang T, Liu Z, Sakata-Kato T, Jiang X, Kato N, Yan N, Yin H (2021) Orthosteric-allosteric dual inhibitors of PfHT1 as selective antimalarial agents. *PNAS* 118(3):e2011749118
 76. Yadav Sm Mathur P (2020) Orthosteric and allosteric modulation of human kinases: a mechanistic view. *Frontiers Biosci Landmark* 25:1462–1487
 77. Zhang M, Jang H, Nussinov R (2020) PI3K inhibitors: review and new strategies. *Chem Sci* 11:5855–5865
 78. Meijer FA, Oerlemans GJM, Brunsved L (2021) Orthosteric and allosteric dual targeting of the nuclear receptor RORγt with a bitopic ligand. *ACS Chem Biol* 16(3):510–519
 79. Fodor M, Price E, Wang P, Lu H, Argintaru A, Chen Z, Glick M, Hao H-X, Kato M, Koenig R, LaRochelle JR, Liu G, McNeill E, Majumdar D, Nishiguchi GA, Perez LB, Paris G, Quinn CM, Ramsey T, Sendzik M, Shultz MD, Williams SL, Stams T, Blacklow SC, Acker MG, LaMarche MJ (2018) Dual allosteric inhibition of SHP2 phosphatase. *ACS Chem Biol* 13(3):647–656
 80. Brown JA, Thorpe IF (2015) Dual allosteric inhibitors jointly modulate protein structure and dynamics in the hepatitis C virus polymerase. *Biochem* 54(26):4131–4141
 81. Wright SW, Rath VL, Genereux PE, Hageman DL, Levy CB, McClure LD, McCoid SC, McPherson RK, Schelhorn TM, Wilder DE, Zavadoski WJ, Gibbs EM, Treadway JL (2005) 5-Chloroindoloyl glycine amide inhibitors of glycogen phosphorylase: synthesis, in vitro, in vivo, and X-ray crystallographic characterization. *Bioorg Med Chem Lett* 15:459–463
 82. Supuran CT (2016) How many carbonic anhydrase inhibition mechanisms exist? *J Enzyme Inhibit Med Chem* 31(3):345–360
 83. Bender W, Staudt M, Trankle C, Mohr K, Holzgrabe U (2000) Probing the size of a hydrophobic binding pocket within the allosteric site of muscarinic acetylcholine M2-receptors. *Life Sci* 66(18):1675–1682

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.