Jaida Begum, Vassiliki T. Skamnaki, Colin Moffatt, Nicolas Bischler, Josephine Sarrou, Alexios-Leandros Skaltsounis, Demetres D. Leonidas, Nikos G. Oikonomakos, Joseph M. Hayes. *An evaluation of indirubin analogues as phosphorylase kinase inhibitors*. Journal of Molecular Graphics & Modeling, 61, 231-242, 2015. DOI: 10.1016/j.jmgm.2015.07.010

An Evaluation of Indirubin Analogues as Phosphorylase Kinase Inhibitors

Jaida Begum,[a,b] Vassiliki T. Skamnaki,[c] Colin Moffatt,[b] Nicolas Bischler,[d] Josephine Sarrou,[d] Alexios-Leandros Skaltsounis,[e] Demetres D. Leonidas,[c] Nikos G. Oikonomakos,[d†] Joseph M. Hayes[a,b]*

[a] *School of Physical Sciences & Computing, Division of Chemistry, University of Central Lancashire, Preston PR1 2HE, United Kingdom*

[b]*School of Forensic & Investigative Sciences, University of Central Lancashire, Preston PR1 2HE, United Kingdom*

[c] *Department of Biochemistry and Biotechnology, University of Thessaly, 26 Ploutonos Str. 41221 Larissa, Greece*

[d] *Institute of Organic and Pharmaceutical Chemistry, National Hellenic Research Foundation, 48 Vas. Constantinou Ave., 116 35 Athens, Greece*

[e]*Division of Pharmacognosy, Department of Pharmacy, University of Athens, Panepistimiopolis-Zografou, Athens 15771, Greece*

† Deceased on Aug 31, 2008.

* Corresponding author: tel: +441772894334 (JMH), fax: +441772894981; email: jhayes@uclan.ac.uk. Address: Centre for Materials Science, Division of Chemistry, University of Central Lancashire, Preston PR1 2HE, United Kingdom.

Abstract

Phosphorylase kinase (PhK) has been linked with a number of conditions such as glycogen storage diseases, psoriasis, type 2 diabetes and more recently, cancer (Camus S. et al., Oncogene 2012, 31, 4333). However, with few reported structural studies on PhK inhibitors, this hinders a structure based drug design approach. In this study, the inhibitory potential of 38 indirubin analogues have been investigated. 11 of these ligands had IC_{50} values in the range 0.170 – 0.360 μM, with indirubin-3'-acetoxime (**1c**) the most potent. 7-bromoindirubin-3'-oxime (**13b**), an antitumor compound which induces caspase-independent cell-death (Ribas J. et al., Oncogene, **2006**, *25*, 6304) is revealed as a specific inhibitor of PhK (IC_{50} = 1.8 μM). Binding assay experiments performed using both PhK-holo and PhK-γtrnc confirmed the inhibitory effects to arise from binding at the kinase domain (γ subunit). High level computations using QM/MM-PBSA binding free energy calculations were in good agreement with experimental binding data, as determined using statistical analysis, and support binding at the ATP-binding site. The value of a QM description for the binding of halogenated ligands exhibiting σ-hole effects is highlighted. A new statistical metric, the 'sum of the modified logarithm of ranks' (SMLR), has been defined which measures performance of a model for both the "early recognition" (ranking earlier/higher) of active compounds and their relative ordering by potency. Through a detailed structure activity relationship analysis considering other kinases (CDK2, CDK5 and GSK-3α/β), 6'(Z) and 7(L) indirubin substitutions have been identified to achieve selective PhK inhibition. The key PhK binding site residues involved can also be targeted using other ligand scaffolds in future work.

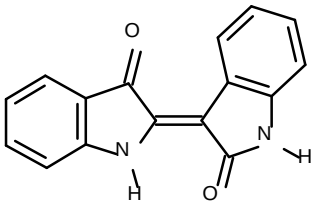
Keywords: QM/MM-PBSA, indirubins, sigma-hole, type 2 diabetes, kinase inhibitors, glycogen phosphorylase

Introduction

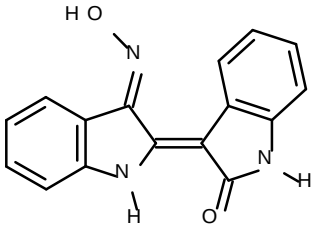
Many physiological processes in cells are controlled via reversible phosphorylation of target proteins through the activity of kinases and phosphatases. Dysfunction affects signal transduction and leads to numerous diseases such as cancer, neurodegenerative diseases and diabetes [1]. Among the hundreds of kinases/phosphatases identified to date, phosphorylase kinase (PhK) plays a key role in the glycogen metabolism, catalyzing the phosphorylation of a single residue (Ser14) on glycogen phosphorylase b (GPb), creating its more active GPa state [2]. PhK is a large 1.3 MDa hexadecameric complex composed of four subunits (α, β, γ, δ). The γ subunit is the kinase domain (45 kDa) carrying the active site, the δ subunit is an endogeneous calmodulin conferring the calcium sensitivity, while the α and β subunits possess phosphorylation sites which regulate the activity of PhK. Genetic mutation on the genes coding for PhK induce severe disorders like glycogen storage diseases (i.e. VIII, IX and X) whereas physiological hyperactivity of the enzyme was shown to be linked with psoriasis [3, 4]. The highly specific activity of PhK also makes it a potential target for drugs involved in the control of the glucose metabolism, such as type 2 diabetes (T2D) [2, 5]. Recently, PhK has also been identified as a new potential target for development of anti-angiogenic therapies [6]. The structure of the PhK heterotetramer $(\alpha\beta\gamma\delta)_4$ has been determined to 9.9 Å resolution using cryo-electron microscopy single particle reconstruction [7]. While the kinase domain of PhK (PhK-γtrnc) has been expressed and crystallized together with nucleotides (ATP and ATP analogue, AMPPMP) and substrate analogues [2, 8-10], until very recently this had not been achieved with inhibitors. In this regard, the crystal structure of the γ subunit of human PhK (hg2-PhK) in complex with sunitinib (pdb: 2Y7J) has been solved. This work has yet to be published, but this is the first successful crystallization of the γ-PhK catalytic site without the

presence of ATP. Recently, also, we have refined using molecular dynamics the structure of PhK-γtrnc in complex with two indirubin and two staurosporine derivatives [5, 11].

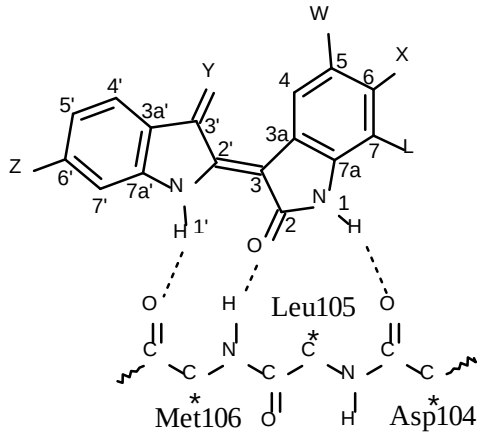
Although kinase inhibitors have been discovered from various sources [12-15], the major obstacle to kinase-directed drugs is the cross-inhibition linked to their highly homologous ATP-binding sites [12]. Staurosporine, although too toxic for clinical use, is a potent inhibitor of PhK (K_i = 0.37 nM) [5]. However, its kinase activity is not specific to PhK [16]. Its analogue, KT5720, is a more specific inhibitor of PhK (K_i = 18.4 nM) and interactions leading to its specificity we have previously studied [5]. Meanwhile, competitive inhibitors from the indirubin family acting on CDKs and GSK-3 have been quite extensively studied [17-20] but information on their activity towards PhK has been limited and has in part motivated the current study [5]. Indirubin (**1a**, Scheme 1) is a natural bis-indol molecule found in plants and mollusks [21]. It has demonstrated antileukemic properties, although the exact mechanism of this remains unclear [22]. Modifications of indirubin have lead to the discovery of other potential anticancer compounds [20, 22-24]. The indirubin family of compounds target several kinases, some of which have been co-crystallised with indirubin analogues: GSK3, CDK2, CDK5, PfPK5 [18, 25-30]. Despite this apparent promiscuity towards kinases, through modification of the indirubin scaffold, kinase specific indirubin inhibitors can be designed.[18, 19] For example, an indirubin analogue over 4000 times more potent for GSK-3α/β compared to CDK1/cyclin B and CDK5/p25 has been reported [18]. In terms of PhK inhibition, we recently reported indirubin-3'-oxime (**1b**, Scheme 1) as a moderately potent inhibitor of PhK (IC_{50} = 144 nM) [5]. However, apart from this study, structure based design studies on PhK ATP-binding site inhibitor are limited. In light of recent results highlighting PhK inhibition from a therapeutic perspective [6], there is an immediate need for new rational drug design efforts.



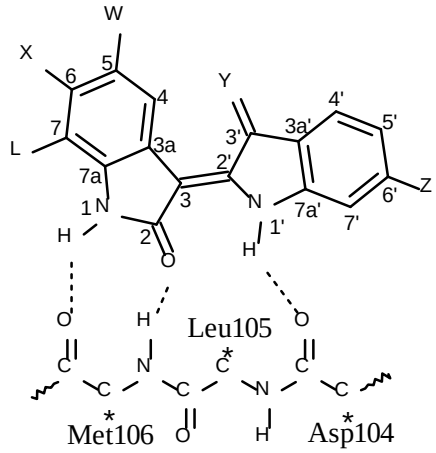
1a Indirubin
IC_{50} = >50 µM



1b Indirubin-3oxime
IC_{50} = 0.210 µM



Expected binding mode



Alternative inverse binding mode

Scheme 1. Indirubin (**1a**) and indirubin-3'oxime (**1b**) together with their IC_{50}'s for PhK-holo inhibition and whose binding at PhK was studied in detail in previous work [5]. Shown also are the hinge region hydrogen bonds as observed in MD studies [5], which is the "expected binding mode" consistent with the binding of indirubins to other kinases [18, 25, 27, 29]. However, a recent study for the binding of indirubins at DYRK kinases, revealed that an "inverse binding mode" is also possible [30]. L, W, X, Y and Z represent the substitution positions of the 38 indirubin analogues studied in this work (Table 1), while substitution of the hydrogen on N1 with methyl was also explored.

In the current study, the binding potential (IC_{50}'s) of 38 indirubin derivatives (shown in Table 1) against PhK has been studied. Experiments using both the truncated catalytic site (PhK-γtrnc) and holoenzyme (PhK-holo) for 5 of these analogues suggest that the inhibitors target the γ-

subunit. Accordingly, exploiting our previously refined MD structure of PhK-γtrnc in complex with indirubin-3'-oxime [5], we have performed extensive modelling calculations in the form of Glide molecular docking followed by more advanced Quantum Mechanics/Molecular Mechanics (QM/MM) and QM/MM-PBSA (Poisson Boltzmann Surface Area) computations [31] to better understand the binding mechanisms at the PhK ATP-binding site. The performance of the computational models are compared by statistical analysis, for which we defined a new metric, the 'sum of the modified logarithm of ranks' (SMLR). Finally, the activity and potential selectivity of indirubin analogues towards PhK inhibition is thoroughly analysed via structure activity relationship (SAR) analysis. The predicted thermodynamics of binding from the QM/MM-PBSA calculations and binding data (structural and activity) for homologous kinases is considered for this purpose.

MATERIALS & METHODS

Protein production: The holoenzyme of PhK from fast-twitch skeletal muscle of female New Zealand White rabbits was purified by the modified method of Cohen [32] as described in Venien-Bryan *et al.*[33]. PhK-γtrnc was expressed as an N-terminal Glutathione-S-Transferase (GST) fusion [5]. Rabbit muscle GPb was purified according to Fischer and Krebs [34]. The concentrations of GPb and PhK were determined from absorbance measurements at 280 nm using extinction coefficients $\varepsilon^{1\%}_{1cm}$ =13.2 [35] and $\varepsilon^{1\%}_{1cm}$ =12.4 [36], respectively. PhK-γtrnc concentration was determined according to Bradford [37].

S***ynthesis of indirubins:*** The 38 indirubin analogues were prepared and characterized according to methods we have described previously [18, 38].

IC_{50} measurements: The enzymatic activities of PhK-holo and PhK-γtrnc were measured by monitoring the conversion of glycogen phosphorylase b (GPb) to GPa by assaying phosphorylase activity in the presence of 10 μM AMP and 0.5 mM caffeine in the direction of glycogen synthesis [2, 39]. In the presence of inhibitors, activities were assayed at constant concentrations of GPb (1 mg/ml) and ATP (50 μM). Blank values were subtracted and activities were calculated after 20 min of incubation at 30°C. The concentration of the various indirubin analogues tested varied from 25 nM to 50 μM during the assay. Activities were expressed as percentages of the maximal activity (i.e. in the absence of inhibitors). Most of the inhibitors showed adequate solubility up to 1 mM in DMSO and the presence of the solvent did not affect the activities of both enzyme forms if its final concentration in the assay was kept below 5% (v/v). None of the compounds tested showed significant inhibition towards GPb. Data was analysed using the non-linear regression program GraFit [40].

Protein and ligand preparation for calculations*:* A MD model of PhK-γtrnc in complex with indirubin-3'-oxime (**1b**, Scheme 1) from our previous work [5] was used as the starting structure for docking of the 38 indirubins, with all H_2Os and counter-ions deleted. This complex corresponds to the minimized representative receptor-ligand conformation which gave the lowest binding free energy (BFE) from MM-GBSA calculations [11] and which exhibited all the key protein-ligand binding interactions from the MD simulations. The 38 indirubins were prepared for calculations using Schrödinger's Maestro and the LigPrep 2.5 program [31].

Docking calculations: Flexible-ligand docking calculations were initially performed using Glide 5.8 in both standard- (SP) and extra-precision (XP) modes, as well as with quantum mechanics – polarized ligand docking (QM-PLD) [31]. Using the PhK-γtrnc/indirubin-3'oxime (**1b**) MD

model, the shape and properties of the catalytic site were mapped onto grids with dimensions ~ 23.7 Å × 23.7 Å × 23.7 Å, centred on the **1b** ligand. Standard parameters were applied including van der Waals (vdW) scaling of non-polar atoms (by 0.8) to include modest induced-fit effects, with up to 5 poses per ligand saved in the Glide-SP and -XP calculations. As per our previous study [5], protein "hinge-region" hydrogen bonds (*c.f.* Scheme 1) were defined for Asp104(O), Met106(H) and Met106(O), with accepted ligand docking poses required to make at least one of these contacts. For the QM-PLD calculations, the output docking poses from Glide SP docking with OPLS-AA atomic partial charges were used to obtain more "accurate" electrostatic potential (ESP) fit ligand atomic charges using the program QSite 5.8 and the QM/MM method in the 'field' of the receptor. Single point energy calculations using the default B3LYP [41-43] density functional theory (DFT) method and the LACVP* basis set on the ligands (QM region) and the OPLS-AA forcefield for the receptor (MM region) were used for this purpose. The ligands with the reparametrized atomic partial charges were then redocked into the catalytic site using Glide-XP.

QM/MM and QM/MM-PBSA calculations: Two types of post-docking calculations were employed. QM/MM interaction energies ($\Delta E_{QM/MM}$) were calculated directly for each saved docking pose (up to 5 per ligand), considering the bound and unbound states of the predicted protein-ligand complexes:

$$\Delta E_{QM/MM} = E_{complex} - E_{rec} - E_{lig} \qquad Eq.(1)$$

Ligand binding free energies were also calculated using QM/MM-PBSA with Eq. (2), the foundations of which we have described before [44].

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

For these calculations, therefore, re-ranking of protein-ligand binding conformations from docking was performed and a more theoretically rigorous estimate of relative activities obtained. All calculations employed DFT with the M06-2X functional [45] and the all electron MIDIX basis set [46, 47] was used for the QM region (the ligands), previously shown to be effective for halogen substituted ligands in studies of this type [48]. The PhKγ-trnc protein was described using MM with the OPLS-AA(2005) forcefield [49]. No cut-off for non-bonded interactions was employed. Effectively, the MM region polarizes the QM region, with electrostatic interactions between the MM point charges and the QM wavefunction, and van der Waals interactions between QM and MM atoms accounted for [50]. Bulk solvation effects were included using PBSA [51], allowing ΔG_{solv} to be calculated. The default solute (internal) dielectric constant of 1.0 was used. An estimate for the loss of solute entropy (ΔS_{MM}) on binding was calculated using MM with the OPLS-AA(2005) forcefield [49] and the Rigid Rotor Harmonic Oscillator (RRHO) approximation. Using this method, the change in vibrational, rotational and translational entropy of the ligands on binding was considered. All QM/MM and QM/MM-PBSA calculations were performed using QSite 5.8 [31]. RRHO calculations were performed using MacroModel 9.9 [31].

Statistical Analysis of Binding Predictions*:* Predicted ranking of the indirubin analogue potencies was based on GlideScore for all docking calculations, $\Delta E_{QM/MM}$ (Eq. (1)) for the QM/MM calculations and ΔG_{bind} (Eq. (2)) for the QM/MM-PBSA calculations. Performance of the different docking and post-docking methods with respect to recovery of the actives (compounds with experimental inhibition constants below a defined threshold, e.g. IC_{50}'s $\leq$ 10 μM) was analyzed in a statistical manner [52, 53]. Several performance metrics were considered including the basic hit-rate to those that better highlight "early recognition" (the ranking of actives early/higher in an ordered list). The hit-rate used is a simple metric which represents the % of real

actives in the predicted top ranked *n* ligands, where *n* is the number of actives. A receiver operator characteristic (ROC) plot shows the relationship between the true positive rate (*TPR, or sensitivity*) and the false positive rate (*FPR, or specificity*), with the Area Under the ROC curve (AU-ROC) a common way to summarize the overall quality of a ROC plot:

$$AU - ROC = \frac{1}{n}\sum_{i=1}^{n}(1 - f_i) \quad Eq.(3)$$

where f_i is the fraction of non-actives ranked higher than the *i*th active. AU-ROC varies between 0 and 1, where 0.5 corresponds to a random ranking and 1 meaning that all actives are ranked above the inactives. The AU-ROC values, however, do not distinguish early and late performance [52]. Accordingly, we also used the "sum of logarithms of ranks" (SLR) metric [53] which accounts for "early recognition" of actives and which we defined as [54, 55]:

$$SLR = -\sum_{i=1}^{n} log_e\left(\frac{r_i}{N}\right) \quad Eq.(4)$$

where r_i is the rank of the *i*th active among *N* total compounds, and the negative logarithm emphasizes early recognition.

To further extend the "early recognition" SLR metric to reward for the correct ordering of actives by their known IC_{50}'s, assigning greater weight to the more active compounds, we have defined a new metric, the "sum of the modified log ranks" (SMLR). If r = [r_1,r_2, …., r_n] is the vector of ranks of the *n* actives and they are ordered by their inhibition constants, with r_1 the most active, the SMLR value is calculated:

$$SMLR = -\sum_{i=1}^{n} log_e \frac{i + \sqrt{r_i}}{n + \sqrt{N}} \quad Eq.(5)$$

The denominator in the expression requires the addition of the square root of the total number of compounds, to ensure the quotient always lies between 0 and 1. Thus, the negative of the log produces values of the same polarity.

In an ideal situation, all actives will be ranked in the first *n* positions for an optimal SLR value and additionally ranked according to potencies for SMLR. The worst scenario is when all the *n* actives are ranked sequentially from $N - n + 1$ to N (for SLR) and additionally in reverse order of potency (for SMLR). This allows us to calculate theoretical "minimum" and "maximum" SLR and SMLR values, and normalized SLR and SSLR values according to:

$$N(M)SLR = \frac{(S)SLR - (S)SLR_{min}}{(S)SLR_{max} - (S)SLR_{min}} \quad Eq.(6)$$

The metric values can therefore range from 0 to 1, with 1 being the ideal situation.

The null distributions of (N)SLR and (N)SMLR were obtained through a simulation procedure. Under the null hypothesis, ranks of active molecules are the result of random assignment across all possible ranks. The empirical random ranking distributions were simulated using a bespoke function by repetitively drawing ranks from a uniform distribution over time. 1 million repetitions has been advocated elsewhere [53], and was applied in this study. Simulations were performed using the statistical software program R [56]. The null distributions can be used as probability distributions, where the relative positions of the actual values of (N)SLR and (N)SMLR on these distributions converts to a probability. Since there were no tied ranks amongst the predictions, random drawing of ranks was done without replacement to imitate reality. Where there were tied values of experimental IC_{50}'s, the ranking of the ties was determined by using their relative predicted ranking.

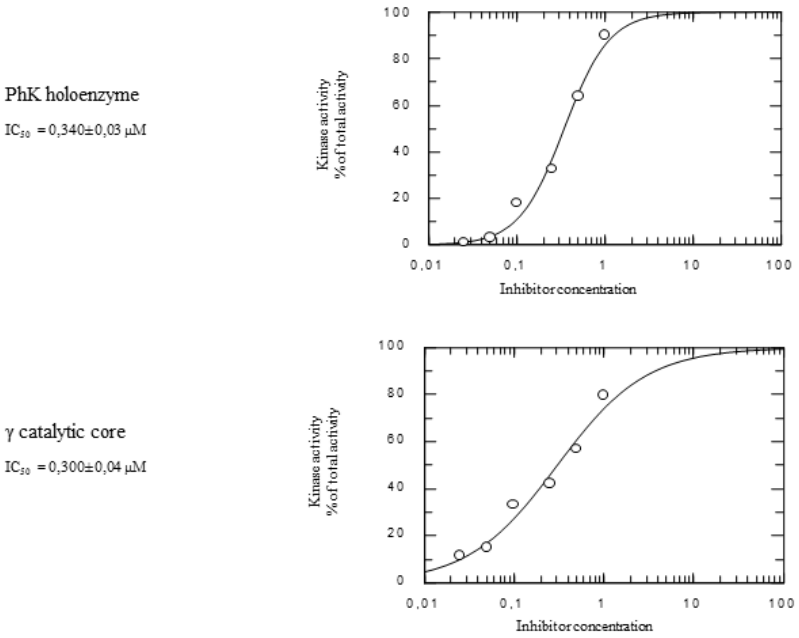


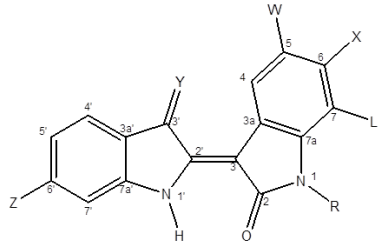
Figure 1. IC_{50} of 6-bromoindirubin-3'-oxime (**4b**) in the presence of the PhK holoenzyme compared to the catalytic core of the enzyme (PhK-γtrnc).

RESULTS AND DISCUSSION

Binding Assay Experiments: Although PhK carries its kinase catalytic activity only on its γ subunit, other nucleotide-binding sites have been reported on other subunits of the holoenzyme. The presence of several sites can be justified by two other nucleotide-dependent activities linked to the enzyme: autophosphorylation [57] and ATPase [58] although no definite relationship has yet been detected between the sites and the activities. At least one adenine nucleotide-binding site has been reported on each of the large regulatory subunits [59-61].

The full set of IC_{50} results for the 38 indirubins using PhK-holo calculated from dose-response curves (Figure 1) are shown in Table 1. For comparison, the corresponding IC_{50} values for CDK1/cyclin B, CDK5/p25 and GSK-3α/β inhibition are also shown [18, 20, 62]. 21 of the 38 inhibitors had IC_{50}'s ≤ 10 µM for PhK-holo inhibition with the best of these having IC_{50}'s ~ 0.2

Table 1. Inhibitory effects of indirubin analogues on PhK-holo where IC_{50} values were calculated from the dose-response curves. The standard error in all experiments was ≤10%. IC_{50}'s for homologous kinases are shown for comparison [18, 20, 62].



No.	Compound	X	Y	Z	W	R	L	IC_{50}(μM)			
								CDK1/ cyclin B[a]	CDK5/ p25[a]	GSK-3α/β[a]	PhK
1a	Indirubin	H	O	H	H	H	H	10.000	10.000	1.000	>50
1b	Indirubin-3′-oxime	H	NOH	H	H	H	H	0.180	0.100	0.022	0.210
1c	Indirubin-3′-acetoxime	H	NOAc	H	H	H	H	1.200	0.700	0.200	0.170
1d	Indirubin-3′-methoxime	H	$NOCH_3$	H	H	H	H	1.000	0.400	0.150	0.340
2b	6,6′-dibromoindirubin-3′-oxime	Br	NOH	Br	H	H	H	17.000	1.300	0.120	2.330
3b	6′-bromoindirubin-3′-oxime	H	NOH	Br	H	H	H	3.000	1.200	0.340	0.290
4a	6-bromoindirubin	Br	O	H	H	H	H	>100	53.0	0.045	>50
4b	6-bromoindirubin-3′-oxime	Br	NOH	H	H	H	H	0.320	0.083	0.005	0.340
4c	6-bromoindirubin-3′-acetoxime	Br	NOAc	H	H	H	H	63.000	2.400	0.010	0.330
5a	6-bromo-N-methylindirubin	Br	O	H	H	CH_3	H	>100	>100	>100	>50
5c	6-bromo-N-methylindirubin-3′-acetoxime	Br	NOAc	H	H	CH_3	H	-	-	-	>50
6a	6-chloroindirubin	Cl	O	H	H	H	H	>100	>100	0.140	>50
6b	6-chloroindirubin-3′-oxime	Cl	NOH	H	H	H	H	0.650	0.100	0.020	0.230
7b	6-iodoindirubin-3′-oxime	I	NOH	H	H	H	H	1.300	0.300	0.010	0.330
8a	6-vinylindirubin	$CH=CH_2$	O	H	H	H	H	4.200	2.400	0.240	>50
8b	6-vinylindirubin-3′-oxime	$CH=CH_2$	NOH	H	H	H	H	1.200	0.420	0.060	0.550
8c	6-vinylindirubin-3′-acetoxime	$CH=CH_2$	NOAc	H	H	H	H	1.600	0.400	0.065	0.540
9a	6-fluoroindirubin	F	O	H	H	H	H	1.500	1.000	0.650	>50
9b	6-fluoroindirubin-3′-oxime	F	NOH	H	H	H	H	0.320	0.150	0.130	0.220
10a	6-bromo-5-methylindirubin	Br	O	H	CH_3	H	H	30.000	60.000	0.025	>50
10c	6-bromo-5-methylindirubin-3′-acetoxime	Br	NOAc	H	CH_3	H	H	31.000	30.000	0.007	0.360
11a	6,5-dichloroindirubin	Cl	O	H	Cl	H	H	45.000	60.000	0.030	>50
11b	6,5-dichloroindirubin-3′-oxime	Cl	NOH	H	Cl	H	H	0.140	0.060	0.004	0.200
11c	6,5-dichloroindirubin-3′-acetoxime	Cl	NOAc	H	Cl	H	H	30.000	0.100	0.004	0.820
12a	6-bromo-5-nitroindirubin	Br	O	H	NO_2	H	H	>100	>100	0.100	>50
12b	6-bromo-5-nitroindirubin-3′-oxime	Br	NOH	H	NO_2	H	H	12.000	0.150	0.007	1.000
12c	6-bromo-5-nitroindirubin-3′-acetoxime	Br	NOAc	H	NO_2	H	H	11.000	31.000	0.006	1.200
13a	7-bromoindirubin	H	O	H	H	H	Br	>100[c]	>100[c]	>100[c]	>50
13b	7-bromoindirubin-3′-oxime	H	NOH	H	H	H	Br	22[b]	33[b]	32[b]	1.800
13c	7-bromoindirubin-3′-acetoxime	H	NOAc	H	H	H	Br	>100[c]	>100[c]	>100[c]	10.00
13d	7-bromoindirubin-3′-methoxime	H	$NOCH_3$	H	H	H	Br	>100[c]	>100[c]	>100[c]	>50
14b	6-Methoxindirubin-3′-oxime	OCH_3	NOH	H	H	H	H	-	-	-	0.700
14c	6-Methoxindirubin-3′-acetoxime	OCH_3	NOAc	H	H	H	H	-	-	-	1.130
15a	7-carbocylindirubin	H	O	H	H	H	COOH	-	-	-	>50
16a	7-bromo-N-methylindirubin	H	O	H	H	CH_3	Br	100[c]	>100[c]	>100[c]	>50
16b	7-bromo-N-methylindirubin-3′-oxime	H	NOH	H	H	CH_3	Br	>100[b]	>100[b]	>100[b]	>50
16c	7-bromo-N-methylindirubin-3′-acetoxime	H	NOAc	H	H	CH_3	Br	70[c]	>100[c]	>100[c]	>50
16d	7-bromo-N-methylindirubin-3′-methoxime	H	$NOCH_3$	H	H	CH_3	Br	>100[c]	>100[c]	>100[c]	>50

[a] Polychronopoulos et al.[18] [b] Ribas et al.[20] [c] Meijer et al.[62]

μM. The inhibition constants of some of the most potent indirubin analogues (**1b**, **1c**, **1d**, **4b**, **4c**) were also determined for PhK-γtrnc (Table 2), where the similarity of the IC_{50} values for both PhK-holo and PhK-γtrnc implies that the indirubin analogues target the γ subunit.

Table 2: Potency comparison of some indirubins analogues for both forms of PhK, PhK-holo and PhK-γtrnc. The IC_{50} values were calculated from dose-response curves, with the experimental errors all within 10%.

No.	Compound	IC_{50} PhK-holo (μM)	IC_{50} PhK-γtrnc (μM)
1b	Indirubin-3'-oxime	0.210	0.144
1c	Indirubin-3'-acetoxime	0.170	0.240
1d	Indirubin-3'-methoxime	0.340	0.280
4b	6-bromoindirubin-3'-oxime	0.340	0.300
4c	6-bromoindirubin-3'-acetoxime	0.330	0.304

Statistical Analysis of Docking, QM/MM and QM/MM-PBSA Performance: To compare the performance of the docking and post-docking calculations to identify actives and correctly rank indirubin analogue potencies in the set of 38 compounds, statistical analysis of results was performed. For our statistical evaluation, we defined two active sets, each employing a different threshold for "activity". **Active Set 1** (21 ligands) was defined as those indirubins with IC_{50}'s ≤ 10 μM (the remaining 17 ligands had IC_{50}'s > 50 μM), while **Active Set 2** (11 ligands) represents the most active indirubins with IC_{50}'s of 0.170 – 0.360 μM. The statistical performance metrics employed were hit-rates, ROC plots and the corresponding AU-ROC values (Eq. (3)), normalized SLR (NSLR) and SMLR (NSMLR) values. NSLR, as defined in Eq.'s (4) and (6), measures the "early recognition" of actives, with the use of the logarithm in Eq. (4) meaning that lower/better ranks of actives contribute proportionately more to the sum, *but* all actives are treated equally irrespective of their relative potencies. Our newly defined (N)SMLR is a metric that additionally

rewards the correct ordering of actives according to their potencies (Eq.'s (5) and (6)). It does this by reducing relative to SLR, the weight of the predicted rank through its square root and taking into account the actual rank through addition. There are many potential metrics which can be devised to account for the correct relative ordering amongst actives. SMLR was used since it is functionally similar to SLR, and it seems to behave reasonably. To illustrate this, a set of situations comparing NSLR and NSMLR values for different hypothetical predicted ranking of actives are given in the Supporting Information (Tables S1-S2).

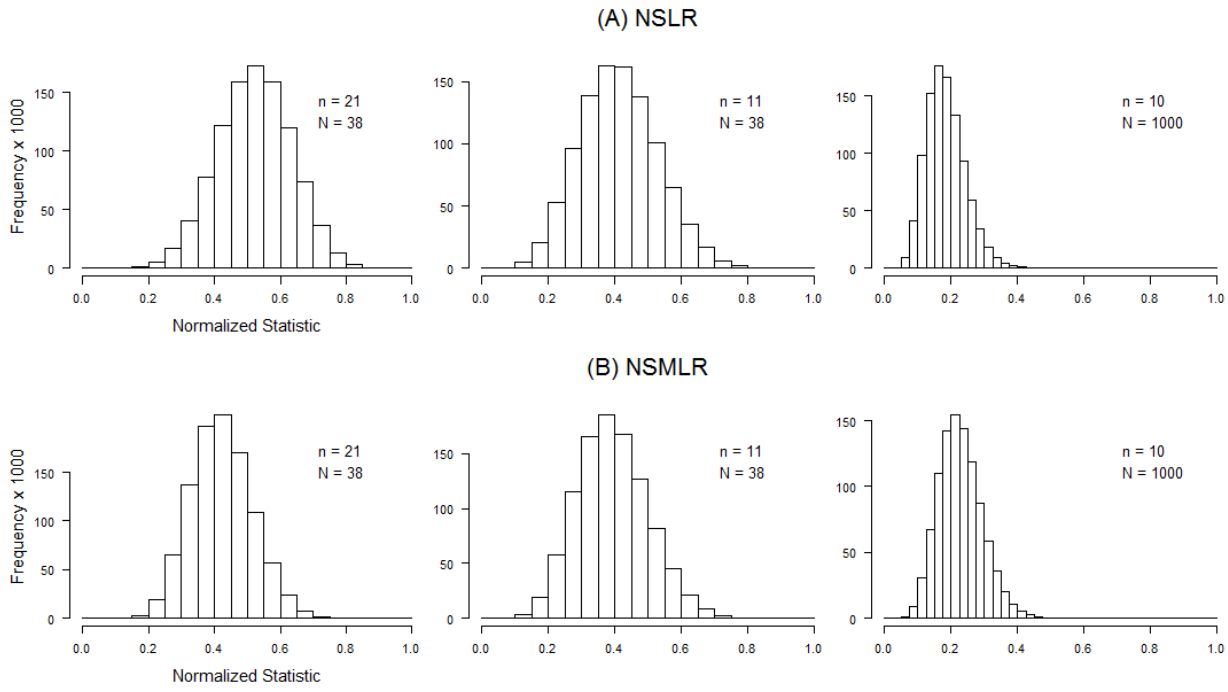


Figure 2. Null distributions for **(A)** the normalized 'sum of the log ranks' (NSLR) and **(B)** the normalized 'sum of the modified log ranks' (NSMLR) metrics. These random ranking distributions were calculated by replacement considering n = 21 actives among N = 38 compounds (**Active Set 1**), n = 11 actives among N = 38 compounds (**Active Set 2**), and an approximately standard situation in virtual screening with n = 10 actives among N = 1000 compounds.

Null distributions comparing NSLR and NSMLR statistics derived by random ranking simulations considering different number of actives *n* within a dataset of *N* compounds are shown in Figure 2. Having *n* = 10 actives within a set of *N* = 1000 compounds represents a situation typical of virtual screening studies and the application of NSLR and NSMLR. However, replicating the situation in the current study, a retrospective analysis of a smaller set of compounds with a much greater ratio of actives ($^{n}/_{N}$), the corresponding distributions for *n* = 21 and *n* = 11 within a set of *N* = 38 compounds is shown. We can see that the NSMLR distribution is similar to that of NSLR, which derives from a gamma distribution (SLR) [53]. With greater $^{n}/_{N}$ ratios such as for **Actives 1** and **Actives 2**, as expected, there is a higher probability of a prediction matching random. This is reflected in the greater spread of the NSLR and NSMLR values in the null distributions compared to the case of *n* = 10 in a set of *N* = 1000. However, the distributions allow us to produce probabilities (p-values) which reflect the robustness of the predictions and therefore entirely validate the statistical significance of the $^{n}/_{N}$ ratios used in the current study. Further comparisons of the null distributions for the different $^{n}/_{N}$ ratios are including in the Supporting Information (Figures S1 and S2). AU-ROC has been reported as independent of the ratio of actives [63], although this has been questioned and refuted in recent studies [52]. An associated problem is that when studies are performed with different ratio of actives, the values cannot be compared [52]. Our comparisons, however, are exempt from these problems, with all comparisons between values considering the same ratio of actives throughout.

The statistical performance metrics for the identification and ranking of active compounds obtained for the different docking and post-docking methods are displayed in the flowchart of Figure 3. Step-wise more theoretically rigorous methods were employed from left to right, where generally we can see the expected trend of improvement of statistics. For the most part, Glide-XP

(Figure 3, **3.2b**) out-performed Glide-SP (**3.1**).[31] With Glide-XP, the potential to identify actives with IC_{50}'s ≤ 10 μM (**Active Set 1**) within the set of 38 compounds was significant, with a hit rate of 71% and an AU-ROC of 0.77. The NSLR values (0.85) reveal that actives were recovered early, while the NSMLR value of 0.70 also revealed reasonable ordering of these actives according to their relative potencies. As could be expected, using tighter criteria for activity with **Active Set 2**, statistics were somewhat lower (hit rate 55%, AU-ROC 0.73, NSLR 0.61, NSMLR 0.62) but still significant; p-values < 0.05 for NSLR and NSMLR were obtained (0.047 and 0.020, respectively).

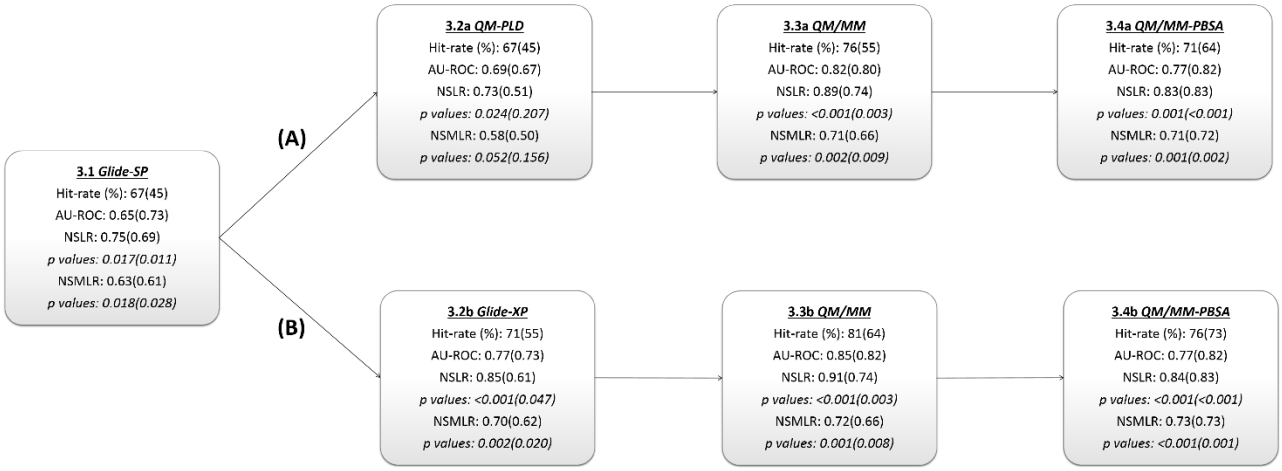


Figure 3. Statistical performance metrics for the identification and ranking of actives using different computational methods. Values are for both **Active Set 1** and **Active Set 2** (in parentheses). The flow-chart reveals the improvement in statistics that can be obtained for docking and post-docking methods by enhancing the theoretical foundations of the models (left to right). There was an exception, however, in terms of QM-PLD (**4.2a**) in pathway **(A)**, where the poorer statistics render this direction redundant for the ligands studied in this work (potential reasons as described in the text). Our main discussion in the text, therefore, follows pathway **(B)** and in particular the QM/MM-PBSA predicted complex geometries and binding free energies (Eq. (2)).

While QM-PLD (**3.2a**) docking employs reparametrized electrostatic potential (ESP) fit ligand atomic charges exploiting QM methods, this did not lead to improvement in results over Glide-XP. The statistics for **Active Set 2** were in fact the worst of all methods tested. NSLR (0.51) and NSMLR (0.50) values were considerably lower, reflected by much larger p-values (~ 0.2). Towards unravelling the source of this poor performance, we need to consider that 25 of the 38 indirubin analogues contain chlorine, bromine or iodine atoms. Through the halogen atom σ-hole effect [48, 64, 65], the unshared electrons on the halogen atom X form a region of negative electrostatic potential around the central region of the R-X bond, leaving a positive "σ-hole" on the outermost tip of the surface on X [64], as shown in Figure 4. Indirubins **9a**-**9b** contain fluorine, but fluorine cannot form halogen bonds due to its high electronegativity and the sp hybridization of its s valence electrons neutralizing the σ-hole. Atomic centred partial charges commonly applied by generic forcefields and docking programs are inadequate for modelling the σ-hole effect; in the current study, ESP fit QM-based atomic centred charges proved to be less accurate than the empirically derived OPLS-AA charges employed by Glide-XP. Recently, more accurate MM models have been proposed [66-69], such as having a δ^- charge on the halogen atom and representing the induced positive δ^+ charge with an extra point (EP) of charge at a distance r* from the halogen atom (e.g. its' atomic radius [67]).

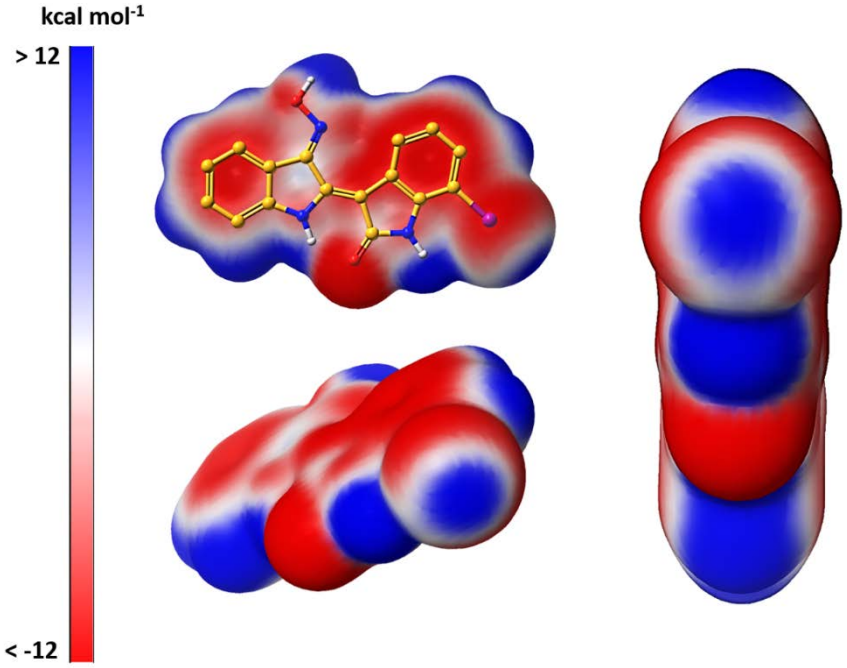


Figure 4. Different orientations of the 7-bromoindirubin-3'-oxime (**13b**) highlighting the halogen σ-hole effect for bromine, calculated using QM at the M06-2X/MIDIX level of theory. Shown is the electrostatic potential (ESP) mapped onto the electron density surface. Note that only the ESP range of -12 – 12 kcal mol^{-1} is used in the surface plot for a better visual demonstration of the σ-hole effect; actual minimum–maximum ESP values were ~-32 – 62 kcal mol^{-1}.

The elimination of parametrization issues for the halogen atoms was achieved employing QM/MM and QM/MM-PBSA calculations, with the ligands modelled entirely using QM. As could be expected based on the above argument, the Glide-XP poses rather than the QM-PLD poses in these calculations produced better performance statistics. The discussion therefore continues with respect to pathway **(B)** in Figure 3. Ranking ligand potencies using QM/MM interaction energies ($\Delta E_{QM/MM}$; Eq.(1)) achieved statistical values for **Active Set 1** of 81% for hit-rate, 0.85, 0.91 and 0.72 for AU-ROC, NSLR and NSMLR, respectively (Figure 3, **3.3b**). Similar improvement in

statistics was observed for **Active Set 2** compared to Glide-XP. Most significantly, however, QM/MM-PBSA ranking of ligands revealed consistently good statistics for both **Active Set 1** and **Active Set 2** (**3.3b**). This more theoretically advanced model (Eq. (2)) for calculating ligand binding free energies (ΔG_{bind}) includes interaction energies $\Delta E_{QM/MM}$, solvation effects modelled using PBSA (ΔG_{solv}), and an estimate of ligand vibrational, rotational and translational (VRT) entropy loss ($T\Delta S_{MM}$) using MM with the rigid rotor harmonic oscillator (RRHO) approximation. Despite the tight criteria for activity employed with **Active Set 2** (≤ 0.360 μM), a hit-rate of 73% was obtained, AU-ROC, NSLR and NSMLR values were 0.82, 0.83 and 0.73, respectively. Indeed, the p-values for NSLR and NSMLR were negligible (≤ 0.001).

Structure Activity Relationship Analysis: Breakdown of the thermodynamic contributions to binding for the 38 ligands based on the QM/MM-PBSA predictions is shown in Table 3. Although the absolute ΔG_{bind} values are too negative, a consequence of the approximations applied and inherent in the methods [44, 48, 70], it is the correct ranking of ligand potencies (relative energies) that concerns us. The top 10 predicted most potent ligands (ΔG_{bind}values < -40 kcal/mol) are highlighted in Table 3 and the predicted binding of the most potent indirubin-3'-acetoxime (**1c**) ligand shown in Figure 5. Ligands in Table 3 are ordered by experimental potencies (IC_{50}'s) and most of the highlighted ligands are close to the top of this table. The QM/MM-PBSA binding geometries and energies, excluding "outliers", forms the basis of our structure activity relationship analysis.

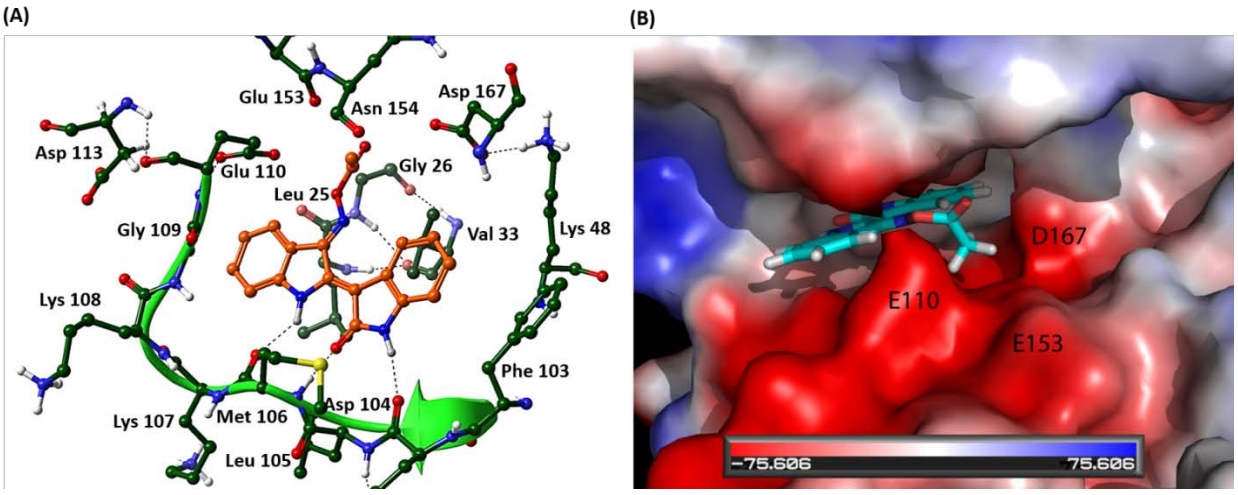


Figure 5. For the most potent (IC_{50} = 0.170 µM) indirubin-3'-acetoxime (**1c**) ligand, **(A)** the predicted binding in the PhK active site as calculated using QM/MM-PBSA and **(B)** the corresponding potential surface (kcal/mol) calculated using PyMol. This represents a "protein contact potential" which approximates the potential that would be felt by a point-charge one solvent radius from the protein surface, if we ignore solvent screening and only consider nearby atoms. The surface shown clearly highlight a triad of negative potential created the Glu110, Glu153 and Asp167 residues which is partially occupied by the acetoxime ligand substituent.

To allow us consider the significant issue of kinase selectivity, inhibition data for CDK1/cyclin B, CDK5/p25 and GSK-3α/β kinases from previous work is also shown in Table 1. Although the IC_{50} values were determined using different assays, in each case the ATP concentration used was less than the K_M for ATP (70 µM). In any case, we are comparing magnitudes of IC_{50} values through ratios (factors of the differences between potencies), rather than the actual values. Inhibition data for 34 of the 38 indirubins is included, allowing for a detailed analysis.

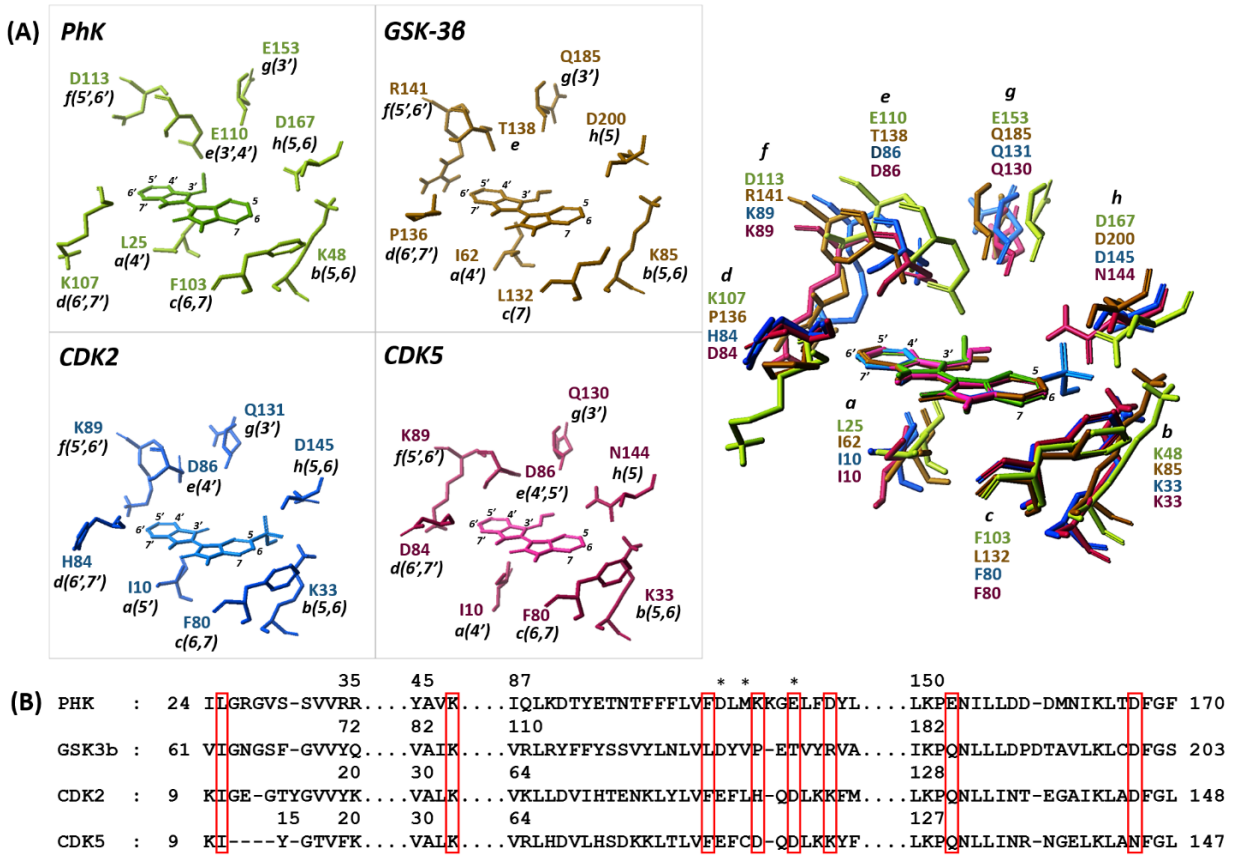
Table 3. Quantum Mechanics/Molecular Mechanics – Poisson–Boltzmann Surface Area (QM/MM-PBSA) results for estimation of binding free energies at the PhK active site. The top 10 predicted ligands are highlighted in bold/italics (ΔG_{bind} < -40 kcal/mol), with ligands ordered by their experimental potencies (IC_{50}'s).[a]

Ligand		$\Delta E_{QM/MM}$	ΔG_{solv}	$T\Delta S_{MM}$	ΔG_{bind}	Exp IC_{50} (μM)
1c	***Indirubin-3'-acetoxime***	***-88.0***	***31.6***	***-15.9***	***-40.6***	***0.17***
11b	6,5-dichloroindirubin-3'-oxime	-80.2	44.0	-17.0	-19.2	0.20
1b	***Indirubin-3'-oxime***	***-92.3***	***28.5***	***-14.9***	***-48.8***	***0.21***
9b	***6-fluoroindirubin-3'-oxime***	***-94.4***	***32.4***	***-15.0***	***-46.9***	***0.22***
6b	***6-chloroindirubin-3'-oxime***	***-102.1***	***41.0***	***-15.7***	***-45.5***	***0.23***
3b	***6'-bromoindirubin-3'-oxime***	***-95.9***	***29.5***	***-15.0***	***-51.4***	***0.29***
4c	6-bromoindirubin-3'-acetoxime	-88.7	36.2	-16.6	-35.9	0.33
7b	***6-iodoindirubin-3'-oxime***	***-99.6***	***42.7***	***-16.3***	***-40.5***	***0.33***
1d	***Indirubin-3'-methoxime***	***-78.8***	***20.7***	***-15.6***	***-42.5***	***0.34***
4b	***6-bromoindirubin-3'-oxime***	***-102.1***	***42.6***	***-15.8***	***-43.7***	***0.34***
10c	6-bromo-5-methylindirubin-3'-acetoxime	-64.0	44.0	-18.1	-1.9	0.36
8c	6-vinylindirubin-3'-acetoxime	-85.9	38.7	-17.7	-29.4	0.54
8b	6-vinylindirubin-3'-oxime	-97.0	45.2	-17.0	-34.7	0.55
14b	6-methoxindirubin-3'-oxime	-87.8	37.9	-16.9	-33.0	0.70
11c	6,5-dichloroindirubin-3'-acetoxime	-65.1	33.7	-17.6	-13.8	0.82
12b	6-bromo-5-nitroindirubin-3'-oxime	-94.0	44.4	-17.0	-32.6	1.00
14c	6-methoxindirubin-3'-acetoxime	-85.8	39.1	-17.7	-29.0	1.13
12c	6-bromo-5-nitroindirubin-3'-acetoxime	-59.5	36.7	-18.9	-3.9	1.20
13b	7-bromoindirubin-3'-oxime	-79.8	37.4	-16.3	-26.1	1.80
2b	***6,6'-dibromoindirubin-3'-oxime***	***-104.7***	***40.7***	***-15.8***	***-48.1***	***2.33***
13c	7-bromoindirubin-3'-acetoxime	-62.2	15.0	-17.2	-30.1	10.00
1a	Indirubin	-66.5	14.9	-14.4	-37.2	> 50
4a	6-bromoindirubin	-73.4	21.3	-15.1	-37.0	> 50
5a	6-bromo-N-methylindirubin	-42.5	21.8	-17.6	-3.1	> 50
5c	6-bromo-N-methylindirubin-3'-acetoxime	-44.9	39.8	-19.3	14.3	> 50
6a	6-chloroindirubin	-74.0	20.5	-15.0	-38.5	> 50
8a	6-vinylindirubin	-63.3	19.9	-16.4	-26.9	> 50
9a	***6-fluoroindirubin***	***-70.3***	***15.3***	***-14.5***	***-40.5***	***> 50***
10a	6-bromo-5-methylindirubin	-54.8	24.5	-16.6	-13.6	> 50
11a	6,5-dichloroindirubin	-59.9	26.9	-16.4	-16.6	> 50
12a	6-bromo-5-nitroindirubin			*no pose*		> 50
13a	7-bromoindirubin	-70.6	35.4	-15.1	-20.1	> 50
13d	7-bromoindirubin-3'-methoxime	-62.8	21.2	-16.2	-25.4	> 50
15a	7-carbocylindirubin	95.0	-61.7	-15.3	48.5	> 50
16a	7-bromo-N-methylindirubin	-67.6	29.3	-15.2	-23.0	> 50
16b	7-bromo-N-methylindirubin-3'-oxime	-93.0	47.8	-16.7	-28.5	> 50
16c	7-bromo-N-methylindirubin-3'-acetoxime	-30.5	32.4	-18.3	20.2	> 50
16d	7-bromo-N-methylindirubin-3'-methoxime	-47.6	35.8	-18.5	6.6	> 50

[a] Calculated energies are in kcal mol^{-1}. Contributions to ΔG_{bind} as per Eq. 2. Values are the Glide-XP docking poses used and represent pathway **(B)** in the flowchart shown in Figure 4.

The key PhK-γtrnc catalytic site residues for indirubin binding (MD model of indirubin-3'-oxime (**1b)** bound at the ATP-binding site [5]) are shown in Figure 6**(A)**, together with the corresponding residues in the crystal structures of indirubin-3'-oxime (**1b)** bound GSK-3β [29], indirubin-3'-oxime (**1b**) bound CDK5/p25 [27], and indirubin-5-sulphonate bound CDK2/cyclin A [25]. Structural alignment of the PhK active site residues with those in CDK2, CDK5 and GSK-3β are shown in Figure 6**(B)** as calculated using the program PDBeFold [71]. We note that the binding data for indirubins in Table 1 was obtained using CDK1/cyclin B rather than CDK2 for practical purposes [18, 20, 62]. However, CDK1 and CDK2 differ by only two amino acids in the ATP binding site, and CDK1 inhibitors were found to be equally active for CDK2 [18]. For the purposes of the below analysis, therefore, CDK1 and CDK2 will be discussed in the same light.

All ligands except 6-bromo-5-nitroindirubin-3'-oxime (**12b**), 7-bromoindirubin-3'-acetoxime (**13c**) and 7-bromoindirubin-3'methoxime (**13d**) are predicted to have the "expected binding mode" shown in Scheme 1. For ligands **5a**, **5c** and **16a-d** with a methyl R substituent at position 1' instead of a hydrogen, the expected binding mode was still adopted but with one less hydrogen bond; this, however, proved detrimental to indirubin analogue potency in all cases (IC_{50}'s > 50 μM) indicating the importance of the hinge region interactions. The relatively weaker predicted ΔG_{bind}values (> approx. -30 kcal/mol) for these ligands reflect the IC_{50} results. The predicted pose for 6-bromoindirubin-3'-methoxime (**12b**) was unexpected, with no hinge region hydrogen bonds and this was likely an error due to the constraints imposed by the rigid receptor docking calculations. The "alternative inverse binding mode" recently reported for the binding of indirubins to DYRK kinases was observed for 7-bromoindirubin-3'acetoxime (**13c**) and 7-bromoindirubin-3'-methoxime (**13d**).



(B)

```
                          35    45      87                  * *   *          150
PHK    :  24 ILGRGVS-SVVRR....YAVK....IQLKDTYETNTFFFLVFDLMKKGELFDYL....LKPENILLDD-DMNIKLIDFGF 170
                          72    82      110                               182
GSK3b  :  61 VIGNGSF-GVVYQ....VAIK....VRLRYFFYSSVYLNLVLDYVP-ETVYRVA....IKPQNLLLDPDTAVLKLCDFGS 203
                          20    30      64                                128
CDK2   :  9  KIGE-GTYGVVYK....VALK....VKLLDVIHTENKLYLVFEFLH-QDLKKFM....LKPQNLLINT-EGAIKLADFGL 148
                    15    20    30      64                                127
CDK5   :  9  KI----Y-GTVFK....VALK....VRLHDVLHSDKKLTLVFEFCD-QDLKKYF....LKPQNLLINR-NGELKLANFGL 147
```

Figure 6. **(A)** The key PhK-γtrnc catalytic site residues for indirubin binding and the corresponding residues in the GSK-3β [29], CDK2/cyclin A [25] and CDK5/p25 [27] homologous kinases with bound indirubins. Indirubin positions for which substitutions have the potential to form interactions with these residue are also listed. For PhK, the MD model of indirubin-3'-oxime (**1b)** bound at the ATP-binding site was used, while for GSK-3β, CDK2/cyclin A and CDK5/p25, the solved crystal structures complex with PDB codes with 1Q41, 1E9H and 1UNH were employed, respectively. Superimpositions were based on the core scaffold of the indirubins. Hinge region residues have been undisplayed for clarity. **(B)** The sequence/structural alignment of catalytic site residues as determined using the program PDBeFold [71]. The residues shown in (A) are highlighted in red boxes.

It is apparent from our results that the Y substituent at position 3' is critical to the inhibitory potential of the indirubin analogues. Indirubins with oxygen as the Y substituent have in each case

IC_{50}'s > 50 μM. Our previous computational study attributed the poor activity of indirubin (**1a)** to a lack of interactions (hinge region excepted) in the PhK active site [5]. In agreement with this, smaller magnitudes of $\Delta E_{QM/MM}$ values (> -75 kcal/mol) for all indirubin analogues with Y = oxygen were obtained. Although an oxygen substituent is also less favoured for CDK1/cyclin B, CDK5/p25 and GSK-3α/β compared to the other Y substituents, the effect on IC_{50}'s is less dramatic, particularly for GSK-3α/β inhibition.

While a Y = oxygen substituent is ineffective, all of oxime, acetoxime and methoxime groups at position 3' are favourable for PhK inhibition. Indeed, having Y substitutions only in the indirubin scaffold revealed indirubin-3'-oxime (**1b**; IC_{50} = 0.210 μM), indirubin-3'-acetoxime (**1c**; IC_{50} = 0.170 μM) and indirubin-3'-methoxime (**1d**; IC_{50} = 0.340 μM) to be among the most potent indirubin inhibitors in the set of 38 studied. For indirubin-3'-oxime (**1b**) and indirubin-3'-acetoxime (**1c),** the predicted $\Delta E_{QM/MM}$ contributions to binding are strong (~ -90 kcal/mol), with the magnitude for indirubin-3'-methoxime (**1d)** significantly lower (-78.8 kcal/mol). However, this is balanced by a lower desolvation cost (ΔG_{solv}) for **1d** with the less polar Y substituent.

While a Y acetoxime group is less favoured by the CDKs and GSK-3α/β compared to a favoured oxime substituent, acetoxime is in fact slightly favoured over oxime for PhK inhibition. **1c** was the most potent indirubin for PhK inhibition identified in this work and its predicted binding is shown in Figure 5. The acetoxime (-NOAc) methyl group partially occupies a space of negative ESP created by Glu153 O, and the Glu110 and Asp167 sidechain carboxylates. Rational inhibitor design would indicate that this pocket can be exploited to increase ligand potency, assuming that the desolvation energy cost does not out-weigh the enthalpy gains from any extra contacts formed [44]. In terms of kinase selectivity, we note that the corresponding aligned "**e**" residue (Figure 6**(A)**) to the PhK Glu110 from the triad is Asp for CDK2 and CDK5, and Thr in the case of GSK-

3β. Comparing the superimpositions in Figure 5**(A)**, the closer contacts with the PhK Glu110 sidechain identifies this residue as a potential means to enhance selectivity.

Of particular significance, we have unravelled the indirubin 6' position (Z substitution) as a means of directing selectivity towards PhK. While the PhK inhibitory potential of indirubin-3'-oxime (**1b**; IC_{50} = 0.210 μM) is relatively unaffected by the introduction of a Br substituent at 6' with 6'-bromoindirubin-3'-oxime (**3b**, IC_{50} = 0.290 μM), the corresponding halogen substitution decreases the inhibitory potency for CDK1/cyclin B, CDK5/p25 and GSK-3α/β by a factor of 12-17. In this regard, the influence of the superimposed "**f**" residues Asp113 (PhK), Arg141 (GSK-3β), Lys89 (CDK2) and Lys89 (CDK5) is revealed as crucial. In the case of GSK-3β, CDK2 and CDK5, the residues are *positively charged* and directed towards the 6' indirubin position, so that repulsive interactions with the positive σ-hole on the 7-Br atom are the likely source of the decreased potency of **3b**. On the contrary, the corresponding residue in PhK is the *negatively charged* Asp167 sidechain which can be exploited in further studies.

Considering halogen substitutions at position 5, we can consider **11a-11c**, but for which there are also substitutions at position 6. Comparing 6,5-dichloroindirubin-3'-oxime (**11b**) with 6-chloroindirubin-3'-oxime (**6b**), a Cl substitution at position 5 is favored by CDK1/cyclin B, CDK5/p25 and GSK-3α/β by factors of 2-5, but only marginally in the case of PhK (IC_{50} = 0.200 μM versus 0.230 μM). Only one halogen (Cl) at position 5 was explored in this study. However, through halogen bonding fine-tuning [64, 65, 72, 73] at position 5, interactions with the PhK Asp167 sidechain carboxylate (Figure 6**(A)**) can potentially be optimized. The corresponding aligned "**h**" residues (Figure 6**(A)**) in the other kinases are either also Asp or Asn (CDK5), but these adopt different conformations and distances from the indirubins in the actives sites. Neither

a 5-Me (**10c**) or 5-NO_2 (**12b-c**), substitution is favorable for PhK inhibition compared to their parent 5-unsubstituted compounds.

For the 6-position (X) of indirubin, halogen substitution with Cl, Br and I atoms is particularly favored for GSK-3α/β inhibition, but not with fluorine. The Br substitution in 6-bromoindirubin-3'-oxime (**4b**) is most favoured (IC_{50} = 0.005 μM) and **4b** is a highly potent/specific inhibitor of GSK-3α/β [18, 19]. In the case of the other kinases (PhK, CDK1/cyclin B and CDK5/p25), 6-halogen substitutions lead to decreased potencies, although the effect is not dramatic. For PhK inhibition, bulkier 6-halogen substitutions of indirubin-3'-oxime does decrease potency (H (**3b**) < F (**9b**) < Cl (**6b**) < Br (**4b**) ~ I (**7b**)), but not by an appreciable amount (IC_{50}'s = 0.210 – 0.340 μM). The QM/MM-PBSA calculations (ΔG_{bind}values; Table 4) reproduce this trend exactly, highlighting the potential of this method to correctly predict favorable halogen substitutions of the indirubin scaffold. The breakdown of the energetics revealed more favorable interactions of the bulkier halogens through $\Delta E_{QM/MM}$ (~-100 kcal/mol for Cl, Br and I substitutions compared to -92.3 and -94.4 kcal/mol for H and F, respectively), but ΔG_{bind} suffers overall due to greater ΔG_{solv} desolvation costs (~40 kcal/mol for Cl, Br and I substitutions compared to ~30 kcal/mol for H and F). The strong preference of GSK-3α/β inhibition towards 6-halogen substitution, however, may lead to selectivity issues for PhK inhibition.

Finally, a significant finding came from our analysis of position 7 (L) substitutions. While a 7-Br substitution decreases the potency of PhK inhibition, the effect is only minor compared to the negative effects on CDK1/cyclin B, CDK5/p25 and GSK-3α/β inhibition. 7-bromoindirubin-3'-acetoxime (**13c**) has an IC_{50} of 10 μM for PhK inhibition, but is inactive (> 100 μM) for the other kinases. Meanwhile, 7-bromoindirubin-3'-oxime (**13b**) has been revealed as important anti-tumour agent triggering the activation of non-apoptotic cell death, although the exact mechanism

has yet to be determined [20, 62]. Interesting, therefore, **13b** exhibits an IC_{50} of 1.8 μM with PhK, a factor of 12-18 times more potent than with CDK1/cyclin B (IC_{50} = 22 μM), CDK5/p25 (IC_{50} = 33 μM) and GSK-3α/β (IC_{50} = 32 μM). Indeed, a weak and gradually decreasing inhibition of CDK1, CDK5 and GSK-3α/β has been reported with different 7-halogeno-indirubin-3'-oximes when the size of the 7-atom substitution was increased [62]. This is suggestive of steric hindrance with the "**c**" aligned residues (Figure 6**(A)**). The residue of significance is a leucine (Leu132) in the case of GSK-3α/β, with the IC_{50} of indirubin-3'-oxime (**3b**) decreasing by a factor of almost 1500 from 0.022 μM to 32 μM for 7-bromoindirubin-3'-oxime (**13b**). For PhK, CDK2 and CDK5, the corresponding residue is phenylalanine. Analysis of the distances in the CDK2 and CDK5 PDB structures used in Fig. 5**(A)** reveal that the 7-H indirubin substituent is orientated towards the sidechain $–CH_2$- and phenyl (C1) atoms at a distance of ~ 2.5 Å. In the PhK model, these distances are > 3.0 Å (3.0 – 3.2 Å). 7-substitution with a bulkier halogen, therefore, appears to be more acceptable in the case of PhK, consistent with the IC_{50} results.

CONCLUSION

38 indirubin analogues have been investigated as inhibitors of the hexadecameric PhK enzyme. Through enzymatic inhibition experiments with both PhK-holo and PhK-γtrnc, binding at the γ-subunit is supported, while our computational analysis revealed favorable interactions of the indirubins at the ATP binding site. The most active indirubins had IC_{50}'s of 0.170 – 0.360 μM. Indirubin-3'-acetoxime (**1c**) was the most potent, with the 3'-acetoxime substituent more favourable for PhK inhibition compared to the other kinases. We have identified a catalytic site cavity formed by a triad of negatively charged residues (Glu110, Glu153 and Asp167) which can potentially be targeted with further 3' (Y group) modifications. While alchemical molecular

dynamics (or Monte Carlo) free energy simulations represent the most accurate method for calculating binding free energies, they are time-consuming and also limited in terms of forcefield accuracy.[74, 75] QM/MM based free energy perturbation (FEP) methods have also recently been developed.[75] Here, a computational protocol (exploiting diverse docking poses) leading to QM/MM-PBSA binding free energy predictions in good agreement with IC_{50} binding data was derived for the set of 38 compounds. As the majority of the ligands studied were halogenated, a QM description of their σ-hole effects was necessary to achieve better agreement with experiment, where we have defined a new statistical metric ((N)MSLR) to evaluate computational performance. The predictive power of our QM/MM-PBSA model, however, can only be truly validated through independent tests, for example, in lead optimization efforts. In this respect, substitutions at indirubin 6'(Z) and 7(L) positions have been identified as means through which more selective PhK inhibition can be achieved. The 7-substituted anti-tumour agent, 7-bromoindirubin-3'-oxime (**13b**), is revealed as a specific inhibitor of PhK; 7-bromoindirubin-3'-acetoxime (**13c**) as a particularly specific inhibitor. Alternatively, 6'-indirubin substituents have the potential to exploit space close to the negatively charged Asp113 PhK residue, with the aligned residues in the other kinases all having positively charged sidechains. It should be highlighted that apart from **13b** and **13c**, the specificity for PhK inhibition over the other kinases is not present for the indirubins studied. Greater specificity is observed, in general, for GSK-3α/β inhibition. For some of the most potent PhK inhibitors, for example **1c** and **3b**, IC_{50}'s are similar for PhK-holo and GSK-3α/β inhibition. However, the potential for more selective PhK inhibition exists through novel 6'(Z), 7(L) and 3'(Y) substitutions. Exploring kinase structural differences to improve selectivity is crucial in kinase inhibitor design. Structure based inhibitor design efforts targeting PhK are to date limited [5]. Recent findings highlight even more the need for new studies on

PhK.[6] Through this work, valuable information and methodologies, not limited to indirubins, have been unraveled to direct further rational design of potent and specific PhK inhibitors.

ACKNOWLEDGEMENTS

JB and JMH would like to acknowledge funding in part for this research through the RSC Bursary scheme.

APPENDIX A. SUPPLEMENTARY DATA

Further comparisons regarding the new SMLR statistical metric are given in the Supporting Information.

ABBREVIATIONS

AU-ROC, area under receiver operator characteristic; DFT, density functional theory; MD, molecular dynamics; NSLR, normalized sum of the logarithm of ranks; NSMLR, normalized sum of the modified logarithm of ranks; PhK, phosphorylase kinase; QM/MM, quantum mechanics/molecular mechanics; QM/MM-PBSA, quantum mechanics/molecular mechanics-Poisson Boltzmann surface area; QM-PLD, quantum mechanics-polarized ligand docking; ROC, receiver operator characteristic; SAR, structure activity relationship; SLR, sum of the logarithm of ranks; SMLR, sum of the modified logarithm of ranks; T2D, type 2 diabetes.

REFERENCES

[1] Cohen, P. Protein kinase inhibitors for the treatment of disease: the promise and the problems. In: Handbook of Experimental Pharmacology, Pinna, L.C., T.W. Ed., Springer-Verlag, Berlin, 2005, Vol. 167, pp. 1-7.

[2] Skamnaki, V.T., Owen, D.J., Noble, M.E.M., Lowe, E.D., Lowe, G., Oikonomakos, N.G., et al. Catalytic mechanism of phosphorylase kinase probed by mutational studies. Biochemistry-Us. 1999, 38, 14718-30.
[3] Heng, M.C.Y., Song, M.K., Heng, M.K. Elevated Phosphorylase-Kinase Activity in Psoriatic Epidermis - Correlation with Increased Phosphorylation and Psoriatic Activity. Brit J Dermatol. 1994, 130, 298-306.
[4] Heng, M.C.Y., Song, M.K., Harker, J., Heng, M.K. Drug-induced suppression of phosphorylase kinase activity correlates with resolution of psoriasis as assessed by clinical, histological and immunohistochemical parameters. Brit J Dermatol. 2000, 143, 937-49.
[5] Hayes, J.M., Skamnaki, V.T., Archontis, G., Lamprakis, C., Sarrou, J., Bischler, N., et al. Kinetics, in silico docking, molecular dynamics, and MM-GBSA binding studies on prototype indirubins, KT5720, and staurosporine as phosphorylase kinase ATP-binding site inhibitors: The role of water molecules examined. Proteins. 2011, 79, 703-19.
[6] Camus, S., Quevedo, C., Menendez, S., Paramonov, I., Stouten, P.F.W., Janssen, R.A.J., et al. Identification of phosphorylase kinase as a novel therapeutic target through high-throughput screening for anti-angiogenesis compounds in zebrafish. Oncogene. 2012, 31, 4333-42.
[7] Venien-Bryan, C., Jonic, S., Skamnaki, V., Brown, N., Bischler, N., Oikonomakos, N.G., et al. The Structure of Phosphorylase Kinase Holoenzyme at 9.9 angstrom Resolution and Location of the Catalytic Subunit and the Substrate Glycogen Phosphorylase. Structure. 2009, 17, 117-27.
[8] Owen, D.J., Papageorgiou, A.C., Garman, E.F., Noble, M.E.M., Johnson, L.N. Expression, Purification and Crystallization of Phosphorylase-Kinase Catalytic Domain. J Mol Biol. 1995, 246, 374-81.
[9] Owen, D.J., Noble, M.E.M., Garman, E.F., Papageorgiou, A.C., Johnson, L.N. 2 Structures of the Catalytic Domain of Phosphorylase-Kinase - an Active Protein-Kinase Complexed with Substrate-Analog and Product. Structure. 1995, 3, 467-82.
[10] Lowe, E.D., Noble, M.E.M., Skamnaki, V.T., Oikonomakos, N.G., Owen, D.J., Johnson, L.N. The crystal structure of a phosphorylase kinase peptide substrate complex: kinase substrate recognition. EMBO J. 1997, 16, 6646-58.
[11] Hayes, J.M., Archontis, G. MM-GB(PB)SA calculations of protein-ligand binding free energies. In: Molecular dynamics: Studies of synthetic and biological macromolecules, Wang, L. Ed., Intech Open, 2012. ISBN 979-953-307-865-5.
[12] Bain, J., Plater, L., Elliott, M., Shpiro, N., Hastie, C.J., Mclauchlan, H., et al. The selectivity of protein kinase inhibitors: a further update. Biochem J. 2007, 408, 297-315.
[13] Noble, M.E.M., Endicott, J.A., Johnson, L.N. Protein kinase inhibitors: Insights into drug design from structure. Science. 2004, 303, 1800-5.
[14] Magiatis, P., Polychronopoulos, P., Meijer, L., Skaltsounis, A.L. Kinase Inhibitors of Marine Origin. Anticancer Res. 2004, 24, 3683-4.
[15] Vougogiannopoulou, K., Skaltsounis, A.L. From Tyrian Purple to Kinase Modulators: Naturally Halogenated Indirubins and Synthetic Analogues. Planta Med. 2012, 78, 1515-28.
[16] Elliott, L.H., Wilkinson, S.E., Sedgwick, A.D., Hill, C.H., Lawton, G., Davis, P.D., et al. K252a Is a Potent and Selective Inhibitor of Phosphorylase-Kinase. Biochem Bioph Res Co. 1990, 171, 148-54.
[17] Leclerc, S., Garnier, M., Hoessel, R., Marko, D., Bibb, J.A., Snyder, G.L., et al. Indirubins inhibit glycogen synthase kinase-3 beta and CDK5/P25, two protein kinases involved in abnormal tau phosphorylation in Alzheimer's disease - A property common to most cycline-dependent kinase inhibitors? J Biol Chem. 2001, 276, 251-60.

[18] Polychronopoulos, P., Magiatis, P., Skaltsounis, A.L., Myrianthopoulos, V., Mikros, E., Tarricone, A., et al. Structural basis for the synthesis of indirubins as potent and selective inhibitors of glycogen synthase kinase-3 and cyclin-dependent kinases. J Med Chem. 2004, 47, 935-46.
[19] Vougogiannopoulou, K., Ferandin, Y., Bettayeb, K., Myrianthopoulos, V., Lozach, O., Fan, Y., et al. Soluble 3 ',6-Substituted Indirubins with Enhanced Selectivity toward Glycogen Synthase Kinase-3 Alter Circadian Period. J Med Chem. 2008, 51, 6421-31.
[20] Ribas, J., Bettayeb, K., Ferandin, Y., Knockaert, M., Garrofe-Ochoa, X., Totzke, F., et al. 7-bromoindirubin-3 '-oxime induces caspase-independent cell death. Oncogene. 2006, 25, 6304-18.
[21] Cooksey, C.J. Tyrian purple: 6,6 '-dibromoindigo and related compounds. Molecules. 2001, 6, 736-69.
[22] Xiao, Z.J., Hao, Y.S., Liu, B.C., Qian, L.S. Indirubin and meisoindigo in the treatment of chronic myelogenous leukemia in China. Leukemia Lymphoma. 2002, 43, 1763-8.
[23] Eisenbrand, G., Hippe, F., Jakobs, S., Muehlbeyer, S. Molecular mechanisms of indirubin and its derivatives: novel anticancer molecules with their origin in traditional Chinese phytomedicine. J Cancer Res Clin. 2004, 130, 627-35.
[24] Shin, E.K., Kim, J.K. Indirubin derivative E804 inhibits angiogenesis. BMC Cancer. 2012, 12.
[25] Davies, T.G., Tunnah, P., Meijer, L., Marko, D., Eisenbrand, G., Endicott, J.A., et al. Inhibitor binding to active and inactive CDK2: The crystal structure of CDK2-cyclin A/indirubin-5-sulphonate. Structure. 2001, 9, 389-97.
[26] Hoessel, R., Leclerc, S., Endicott, J.A., Nobel, M.E.M., Lawrie, A., Tunnah, P., et al. Indirubin, the active constituent of a Chinese antileukaemia medicine, inhibits cyclin-dependent kinases. Nat Cell Biol. 1999, 1, 60-7.
[27] Mapelli, M., Massimiliano, L., Crovace, C., Seeliger, M.A., Tsai, L.H., Meijer, L., et al. Mechanism of CDK5/p25 binding by CDK inhibitors. J Med Chem. 2005, 48, 671-9.
[28] Holton, S., Merckx, A., Burgess, D., Doerig, C., Noble, M., Endicott, J. Structures of P-falciparum PfPK5 test the CDK regulation paradigm and suggest mechanisms of small molecule inhibition. Structure. 2003, 11, 1329-37.
[29] Bertrand, J.A., Thieffine, S., Vulpetti, A., Cristiani, C., Valsasina, B., Knapp, S., et al. Structural characterization of the GSK-3 beta active site using selective and non-selective ATP-mimetic inhibitors. J Mol Biol. 2003, 333, 393-407.
[30] Myrianthopoulos, V., Kritsanida, M., Gaboriaud-Kolar, N., Magiatis, P., Ferandin, Y., Durieu, E., et al. Novel Inverse Binding Mode of Indirubin Derivatives Yields Improved Selectivity for DYRK Kinases. ACS Med Chem Lett. 2013, 4, 22-6.
[31] Schrodinger. LLC, New York, NY, 2014.
[32] Cohen, P. Subunit Structure of Rabbit-Skeletal-Muscle Phosphorylase Kinase, and Molecular Basis of Its Activation Reactions. Eur J Biochem. 1973, 34, 1-14.
[33] Venien-Bryan, C., Lowe, E.M., Boisset, N., Traxler, K.W., Johnson, L.N., Carlson, G.M. Three-dimensional structure of phosphorylase kinase at 22 angstrom resolution and its complex with glycogen phosphorylase b. Structure. 2002, 10, 33-41.
[34] Fischer, E.H., Krebs, E. G. Muscle phosphorylase-B. Methods Enzymol. 1962, 5, 369-73.
[35] Kastenschmidt, L.L., Kastenschmidt, J., Helmreich, E. Subunit interactions and their relationship to the allosteric properties of rabbit skeletal muscle phosphorylase b. Biochemistry. 1968, 7, 3590-608.

[36] Chebotareva, N.A., Andreeva, I.E., Makeeva, V.F., Livanova, N.B., Kurganov, B.I. Effect of molecular crowding on self-association of phosphorylase kinase and its interaction with phosphorylase b and glycogen. Journal of Molecular Recognition. 2004, 17, 426-32.
[37] Bradford, M.M. Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding. Anal Biochem. 1976, 72, 248-54.
[38] Ferandin, Y., Bettayeb, K., Kritsanida, M., Lozach, O., Polychronopoulos, P., Magiatis, P., et al. 3 '-substituted 7-halogenoindirubins, a new class of cell death inducing agents. J Med Chem. 2006, 49, 4638-49.
[39] Uyttenhove, K., Bollen, M., Stalmans, W. An optimized assay of phosphorylase kinase in crude liver preparations. The Biochemical Journal. 1991, 278 (Pt 3), 899-901.
[40] Leatherburrow, R.J. GraFit. Erithakus Software, Staines, U.K., 1992.
[41] Becke, A.D. Density-Functional Thermochemistry .3. The Role of Exact Exchange. J Chem Phys. 1993, 98, 5648-52.
[42] Lee, C.T., Yang, W.T., Parr, R.G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron-Density. Phys Rev B. 1988, 37, 785-9.
[43] Stephens, P.J., Devlin, F.J., Chabalowski, C.F., Frisch, M.J. Ab-Initio Calculation of Vibrational Absorption and Circular-Dichroism Spectra Using Density-Functional Force-Fields. J Phys Chem-Us. 1994, 98, 11623-7.
[44] Manta, S., Xipnitou, A., Kiritsis, C., Kantsadi, A.L., Hayes, J.M., Skamnaki, V.T., et al. 3 '-Axial CH2OH Substitution on Glucopyranose does not Increase Glycogen Phosphorylase Inhibitory Potency. QM/MM-PBSA Calculations Suggest Why. Chem Biol Drug Des. 2012, 79, 663-73.
[45] Zhao, Y., Truhlar, D.G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. Theor Chem Acc. 2008, 120, 215-41.
[46] Easton, R.E., Giesen, D.J., Welch, A., Cramer, C.J., Truhlar, D.G. The MIDI! basis set for quantum mechanical calculations of molecular geometries and partial charges. Theor Chim Acta. 1996, 93, 281-301.
[47] Li, J.B., Cramer, C.J., Truhlar, D.G. MIDI! basis set for silicon, bromine, and iodine. Theor Chem Acc. 1998, 99, 192-6.
[48] Kantsadi, A.L., Hayes, J.M., Manta, S., Skamnaki, V.T., Kiritsis, C., Psarra, A.M.G., et al. The s-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. 2012, 7, 722-32.
[49] Kaminski, G.A., Friesner, R.A., Tirado-Rives, J., Jorgensen, W.L. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. J Phys Chem B. 2001, 105, 6474-87.
[50] Murphy, R.B., Philipp, D.M., Friesner, R.A. A mixed quantum mechanics/molecular mechanics (QM/MM) method for large-scale modeling of chemistry in protein environments. J Comput Chem. 2000, 21, 1442-57.
[51] Marten, B., Kim, K., Cortis, C., Friesner, R.A., Murphy, R.B., Ringnalda, M.N., et al. New model for calculation of solvation free energies: Correction of self-consistent reaction field continuum dielectric theory for short-range hydrogen-bonding effects. J Phys Chem-Us. 1996, 100, 11775-88.

[52] Truchon, J.F., Bayly, C.I. Evaluating virtual screening methods: Good and bad metrics for the "early recognition" problem. J Chem Inf Model. 2007, 47, 488-508.
[53] Zhao, W., Hevener, K.E., White, S.W., Lee, R.E., Boyett, J.M. A statistical framework to evaluate virtual screening. BMC Bioinformatics. 2009, 10.
[54] Venkatraman, V., Perez-Nueno, V.I., Mavridis, L., Ritchie, D.W. Comprehensive Comparison of Ligand-Based Virtual Screening Tools Against the DUD Data set Reveals Limitations of Current 3D Methods. J Chem Inf Model. 2010, 50, 2079-93.
[55] Parmenopoulou, V., Kantsadi, A.L., Tsirkone, V.G., Chatzileontiadou, D.S.M., Manta, S., Zographos, S.E., et al. Structure based inhibitor design targeting glycogen phosphorylase b. Virtual screening, synthesis, biochemical and biological assessment of novel N-acyl-beta-D-glucopyranosylamines. Bioorgan Med Chem. 2014, 22, 4810-25.
[56] R-Core-Team. R: A language and environment for statistical computing., R Foundation for Statistical Computing, Vienna, Austria, 2014.
[57] Brushia, R.J., Walsh, D.A. Phosphorylase kinase: the complexity of its regulation is reflected in the complexity of its structure. Frontiers in bioscience : a journal and virtual library. 1999, 4, D618-41.
[58] Paudel, H.K., Carlson, G.M. The ATPase activity of phosphorylase kinase is regulated in parallel with its protein kinase activity. The Journal of biological chemistry. 1991, 266, 16524-9.
[59] Zaman, N., Varsanyi, M., Heilmeyer, L.M., Jr., Sotiroudis, T.G., Johnson, C.M., Crabb, J.W. Reaction of fluorescein isothiocyanate with an ATP-binding site on the phosphorylase kinase alpha subunit. Eur J Biochem. 1989, 182, 577-84.
[60] Cheng, A., Carlson, G.M. Competition between nucleoside diphosphates and triphosphates at the catalytic and allosteric sites of phosphorylase kinase. The Journal of biological chemistry. 1988, 263, 5543-9.
[61] Gulyaeva, N.V., Vulfson, P.L., Severin, E.S. The role of different ATP analogs on the catalytic function of phosphorylase kinase. Biokhimiya (Moscow). 1977, 43, 373-81.
[62] Meijer, L., Bettayeb, K., Skaltsounis, A.L., Maglatis, P., Boix, J., Ribas J. 7-substituted indirubin-3'-oximes and their applications. U.S.A., 2013, Vol. WO2007099402.
[63] Triballeau, N., Acher, F., Brabet, I., Pin, J.P., Bertrand, H.O. Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4. J Med Chem. 2005, 48, 2534-47.
[64] Clark, T., Hennemann, M., Murray, J.S., Politzer, P. Halogen bonding: the sigma-hole. J Mol Model. 2007, 13, 291-6.
[65] Scholfield, M.R., Vander Zanden, C.M., Carter, M., Ho, P.S. Halogen bonding (X-bonding): A biological perspective. Protein Sci. 2013, 22, 139-52.
[66] Kolar, M., Hobza, P. On Extension of the Current Biomolecular Empirical Force Field for the Description of Halogen Bonds. J Chem Theory Comput. 2012, 8, 1325-33.
[67] Ibrahim, M.A.A. Molecular Mechanical Study of Halogen Bonding in Drug Discovery. J Comput Chem. 2011, 32, 2564-74.
[68] Ibrahim, M.A.A. Molecular mechanical perspective on halogen bonding. J Mol Model. 2012, 18, 4625-38.
[69] Jorgensen, W.L., Schyman, P. Treatment of Halogen Bonding in the OPLS-AA Force Field: Application to Potent Anti-HIV Agents. J Chem Theory Comput. 2012, 8, 3895-901.
[70] Ciancetta, A., Genheden, S., Ryde, U. A QM/MM study of the binding of RAPTA ligands to cathepsin B. J Comput Aid Mol Des. 2011, 25, 729-42.

[71] Krissinel, E., Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr D. 2004, 60, 2256-68.
[72] Riley, K.E., Murray, J.S., Fanfrlik, J., Rezac, J., Sola, R.J., Concha, M.C., et al. Halogen bond tunability II: the varying roles of electrostatic and dispersion contributions to attraction in halogen bonds. J Mol Model. 2013, 19, 4651-9.
[73] Fanfrlik, J., Kolar, M., Kamlar, M., Hurny, D., Ruiz, F.X., Cousido-Siah, A., et al. Modulation of Aldose Reductase Inhibition by Halogen Bond Tuning. ACS Chem Biol. 2013, 8, 2484-92.
[74] Aleksandrov, A., Thompson, D., Simonson, T. Alchemical free energy simulations for biological complexes: powerful but temperamental ... J Mol Recognit. 2010, 23, 117-27.
[75] Reddy, M.R., Reddy, C.R., Rathore, R.S., Erion, M.D., Aparoy, P., Reddy, R.N., et al. Free Energy Calculations to Estimate Ligand-Binding Affinities in Structure-Based Drug Design. Curr Pharm Design. 2014, 20, 3323-37.