CHEMMEDCHEM

CHEMISTRY ENABLING DRUG DISCOVERY

Accepted Article

Title: X-ray structures and feasibility assessment of CLK2 inhibitors for Phelan McDermid syndrome

Authors: Andreas Markus Lerchner, Joerg Kallen, Christian Bergsdorf, Bertrand Arnaud, Mario Bernhard, Murielle Brichet, Amanda Cobos-Correa, Azeddine Elhajouji, Felix Freuler, Ivan Galimberti, Christel Guibourdenche, Simon Haenni, Sandra Holzinger, Juerg Hunziker, Aude Izaac, Markus Kaufmann, Lukas Leder, Hans-Joerg Martus, Peter von Matt, Valery Polyakov, Patrik Roethlisberger, Guglielmo Roma, Nikolaus Stiefl, and Marianne Uteng

To be cited as: *ChemMedChem* 10.1002/cmdc.201800344

Link to VoR: http://dx.doi.org/10.1002/cmdc.201800344

A Journal of

ChemPubSoc Europe

WILEY-VCH

www.chemmedchem.org

X-ray structures and feasibility assessment of CLK2 inhibitors for Phelan McDermid syndrome

Joerg Kallen [a], Christian Bergsdorf [a], Bertrand Arnaud [c], Mario Bernhard [b], Murielle Brichet [b], Amanda Cobos-Correa [b], Azeddine Elhajouji [b], Felix Freuler [b], Ivan Galimberti [b], Christel Guibourdenche [c], Simon Haenni [b], Sandra Holzinger [b], Juerg Hunziker [b], Aude Izaac [b], Markus Kaufmann [b], Lukas Leder [b], Hans-Joerg Martus [b], Peter von Matt [c], Valery Polyakov [c], Patrik Roethlisberger [b], Guglielmo Roma [b], Nikolaus Stiefl [c], Marianne Uteng [b], and Andreas Lerchner *[c]

[a] Dr. J. Kallen, Dr. C. Bergsdorf, Novartis Institutes for BioMedical Research, Novartis Campus CH-4002 Basel Switzerland; these two first authors contributed equally to this work

[b] M. Bernhard, M. Brichet, Dr. A. Cobos-Correa, Dr. A. Elhajouji, F. Freuler, Dr. I. Galimberti, S. Haenni, S. Holzinger, A. Izaac, M. Kaufmann, Dr. L. Leder, Dr. H-J Martus, P. Roethlisberger, Dr. G. Roma, , Dr. M. Uteng, Novartis Institutes for BioMedical Research, Novartis Campus CH-4002 Basel Switzerland

[c] Dr. A. Lerchner, B. Arnaud, Dr. C. Guibourdenche, Dr. J. Hunziker, Dr. P. von Matt, Dr. V. Polyakov, Dr. N. Stiefl, Global Discovery Chemistry, Novartis Institutes for BioMedical Research, Novartis Campus CH-4002 Basel Switzerland; E-mail: andreas.lerchner@novartis.com

Abstract

CLK2 inhibition has been proposed as a potential mechanism to improve autism and neuronal functions in Phelan-McDermid syndrome (PMDS). Herein, we report the discovery of a very potent indazole CLK inhibitor series, and the CLK2 X-ray structure of its most potent analog. This new indazole series was identified via a biochemical CLK2 Caliper assay screen with 30k compounds that were selected by an *in silico* approach. Novel high resolution X-ray structures of all CLKs, including the first CLK4 X-ray structure, bound to known CLK2 inhibitor tool compounds (e.g. TG003, CX-4945), are also shown and yield insights into inhibitor selectivity in the CLK family. Efficacy of our new CLK2 inhibitors from the indazole series was demonstrated in the mouse brain slice assay, and potential safety concerns were investigated. We show genotoxicity findings in the human lymphocyte MNT assay using two structurally different CLK inhibitors, which reveals a major concern for pan-CLK inhibition in PMDS.

Introduction

Phelan-McDermid syndrome (PMDS) is a congenital disease caused by loss of the terminal segment of chromosome 22, with concomitant loss of the SHANK3 gene, a deficiency that is associated with intellectual disability and autism. In a recent work, inhibition of the serine-threonine kinase CLK2 was shown to rescue spine density in SHANK3 knockdown neurons [1]. Therefore, inhibition of CLK2 was proposed as an interesting new target that could potentially ameliorate autistic features in PMDS. Obviously, in order to assess the therapeutic concept of CLK inhibitors to treat PMDS, it was especially important to investigate some safety relevant aspects of CLK inhibition. In particular, it was already known that CLKs regulate alternative splicing by phosphorylation of serine- and arginine rich (SR) proteins, which promote the assembly of the spliceosome [2].

Several pan-CLK inhibitors with acceptable ADME properties have been described to be selective against most other kinases (apart from CLK1-4/DYRK1A family) and in particular the benzo-thiazole derivative TG003 has been shown to restore synaptic impairments in SHANK3 deficient neurons [1-2]. While it seems to be sufficient to restore spine density with the selective knockdown of CLK2 [1], a pan-CLK inhibitor might lead to higher safety concerns in comparison to the inhibition of only one CLK family member. Therefore, we investigated whether it was possible to identify CLK2 inhibitors with high selectivity over CLK1, CLK3, CLK4, and/or DYRK1A. Due to biochemical potency data reported in recent publications [3], we anticipated that this would be difficult to achieve and would require an in-depth analysis. Here we present the identification of a new CLK inhibitor scaffold, found in a biochemical CLK2 Caliper assay screen with 30k compounds that were selected by an *in silico* approach. We show seven new CLK-family X-ray structures (including the first X-ray structure for CLK4), for the complexes with TG003 [2], with a novel potent in house CLK2 inhibitor (Indazole1) and with two tool compounds. These novel CLK-family X-ray structures are the basis for compound design/optimization and address questions concerning selectivity within the CLK1-4/DYRK1A kinase family.

Results and Discussions

X-ray structure reveals the binding mode of TG003 complexed with CLK2 kinase domain

When we initiated our work, one CLK2 X-ray structure was publicly available, deposited by the SGC consortium (pdb entry 3NR9, 2.89 Å resolution). As the binding mode of TG003 was unknown, and there was the possibility that TG003 was an allosteric and/or covalent inhibitor, we embarked on determining its CLK2 X-ray co-crystal structure (see Fig. 1 for the chemical structures of all compounds described in this paper). A human CLK2 kinase domain construct containing amino acids R130–D496 was co-expressed with λ-

phosphatase in *E. coli*, affinity-purified and determined to be un-phosphorylated by LC-MS. Optimized co-crystals with TG003 diffracted to 2.6 Å (supporting information, Table S2). The electron density clearly showed that TG003 binds non-covalently in the ATP site (supporting information, Fig. S1A) and that the N- and C-terminal kinase lobes are packed tightly, contributing to the overall binding affinity (Fig. 2*A*). The benzo-thiazole core of TG003 is sandwiched between A191, V177 of the N-terminal lobe and L297, V326 of the C-terminal lobe (Fig. 2*B*).

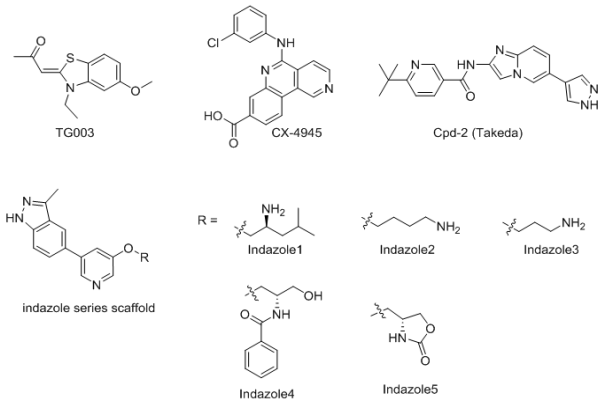


Figure 1. Chemical structures of CLK2 inhibitors described in this paper.

F243 is the gatekeeper (gk) and L246 the gk + 3 residue [4]. The methoxy oxygen accepts a hydrogen bond (3.2 Å) from NH-L246. The keto-oxygen of the propan-2-one group accepts a hydrogen bond (3.2 Å) from the catalytic lysine NZ-K193 and makes water mediated hydrogen bonds with NH-D327 and OE2-E208. Overall, the interactions of TG003 with CLK2 as determined by its X-ray structure can explain well, why such a small molecule can still reach decent potency on target (mCLK2 IC_{50} = 200 nM published by Muraki et al. [2]).

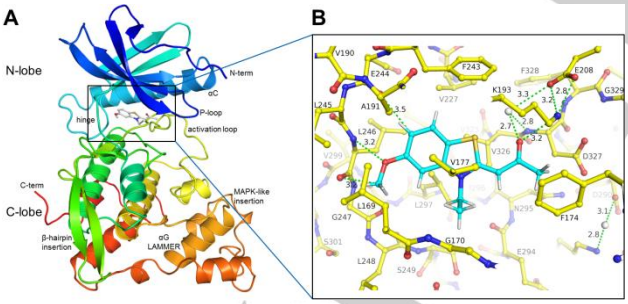


Figure 2. X-ray structure shows that TG003 binds in the ATP-binding site of CLK2. *A*, X-ray structure at 2.60 Å resolution of the CLK2 kinase domain shown as ribbon diagram (color ramping from *blue*: N-terminus to *red*: C-terminus) complexed with TG003 (*white*). *B*, details of interactions between CLK2 (carbons in *yellow*, nitrogens in *blue*, oxygens in *red*) and TG003 (carbons in *cyan*, computed hydrogen positions in *white*). Water molecules are shown as *white* spheres. Selected interactions are shown as *green* dotted lines. The coordinates for CLK2 complexed with TG003 have been deposited in the PDB databank (PDB access code = 6FYI)

Endeavor to achieve high resolution X-ray structures for CLK1, CLK2, CLK3 and CLK4 kinase domains

In order to increase our structural understanding of potential selectivity within the CLK-family and possibly versus DYRK1A, we embarked to determine X-ray structures for further selected tool compounds in complex with various CLKs (notably, no CLK4 X-ray structure was publicly available). Two publications about CX-4945 by Kim et al. describe this compound in its role as a highly potent CLK inhibitor [5] and also in its role as a highly potent DYRK1A inhibitor; DYRK1A inhibition is proposed as a potential treatment for Down syndrome [6]. The high potency of CLK inhibitor Cpd-2 is described in a publication by Araki et al. [7]. For CLK4 we designed a construct based on a sequence alignment with CLK2. All four CLK kinase domains were expressed and purified, using the constructs CLK1(H148–I484), CLK2(S136–D496), CLK3(Q275–T632) and CLK4(H146–K480). All in house X-ray structures turned out to be the un-phosphorylated forms, even though the protein batches used in crystallization for CLK1 and CLK4 were ca. 50% phosphorylated and 50% un-phosphorylated (CLK2 and CLK3 batches were 100% un-phosphorylated).

X-ray structures reveal the binding mode of CX-4945 complexed with CLK2, CLK3, and CLK4 kinase domains

We managed to obtain well-diffracting co-crystals of CX-4945 with CLK2, CLK3 and CLK4. The crystals for CLK4/CX-4945 diffracted to 2.46 Å (supporting information, Table S4). The three overall kinase domain structures turned out to be similar. The X-ray structure for CLK4/CX-4945 showed the details of the interactions in the ATP binding site (Fig. 3*A*). The benzo-naphthyridine core is sandwiched between A189, V175 of the N-terminal lobe and L295, V324 of the C-terminal lobe. It also makes an aromatic interaction with the gk residue F241. In the hinge region, the 2-N of the naphthyridine accepts a hydrogen bond (2.9 Å) from NH-L244. The carboxylate of CX-4945 makes a strong saltbridge interaction (2.8 Å) with the catalytic lysine NZ-K191. Interestingly, this carboxylate oxygen is at the same position as the water molecule which mediates the (weaker) interactions with NZ-K193 for CLK2/TG003 (Fig. 2*B*). The carboxylate of CX-4945 also makes water mediated hydrogen bonds with NH-F323, OE2-E206. Importantly, the 3-Cl of the Cl-phenyl group makes an edge interaction (3.4 Å) with F172. A comparison of the X-ray structures for CLK4/CX-4945 and CLK2/CX-4945 (1.95 Å resolution) revealed no important differences of interactions and no significant side chain differences in the ATP binding site. As a consequence we reasoned that selectivity for CLK2 versus CLK4 with an inhibitor in the ATP binding site will be very difficult to obtain. By contrast, a comparison of the X-ray structures for CLK4/CX-4945 and CLK3/CX-4945 (2.29 Å resolution) revealed a key difference (Fig. 3*B*). For CLK3, the small sidechain of residue A319 underneath provides less vdW-interactions (van der Waals interactions) with the aromatic core of the inhibitor than the larger V324 for CLK4 (V324 for CLK1, V326 for CLK2). The X-ray structures thus predict that the affinity of CX-4945 should be lowest for CLK3, and that in general selectivity favouring CLK1, CLK2, CLK4 versus CLK3 could be envisaged in compound optimization. In order to verify this prediction, CX-4945 was then submitted for testing in CLK1-4/DYRK1A radiometric protein kinase assays by ProQinase (see supporting information for assay description), which showed IC_{50} values of 12 nM (CLK1), 7 nM (CLK2), 143 nM (CLK3), 11 nM (CLK4), and 14 nM (DYRK1A). These results confirmed the prediction based on CLK X-rays to reach the highest selectivity over CLK3 within this family.

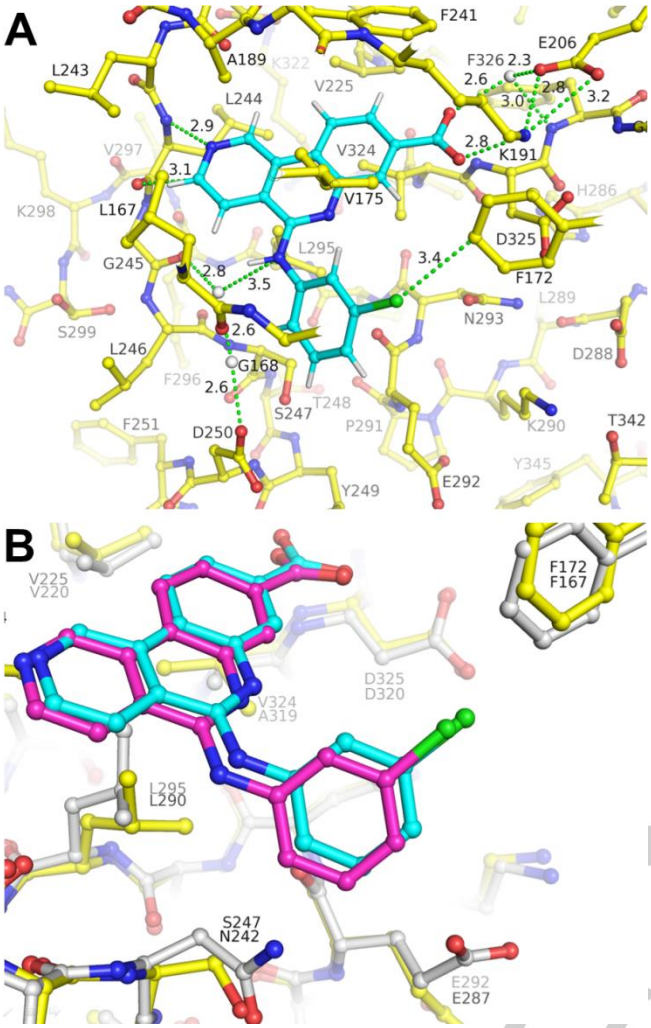


Figure 3. First X-ray structure for CLK4. Comparison of interactions made by CX-4945 in the ATP-binding sites of CLK4 and CLK3: Structural basis for selectivity against CLK3. *A*, X-ray structure at 2.46 Å resolution of the kinase domain of CLK4 (carbons in *yellow*, nitrogens in *blue*, oxygens in *red*) complexed with CX-4945 (carbons in *cyan*, computed hydrogen positions in *white*). Water molecules are shown as *white* spheres. Selected interactions are shown as *green* dotted lines. *B,* Superposed on the X-ray structure of CLK4/CX-4945 (coloring as in *A*) is the X-ray structure at 2.29 Å resolution of CLK3 (carbons in *white*) complexed with CX-4945 (carbons in *magenta*). The coordinates for CLK2, CLK3, CLK4 complexed with CX-4945 have been deposited in the PDB databank (PDB access codes = 6FYL, 6FYP, 6FYV).

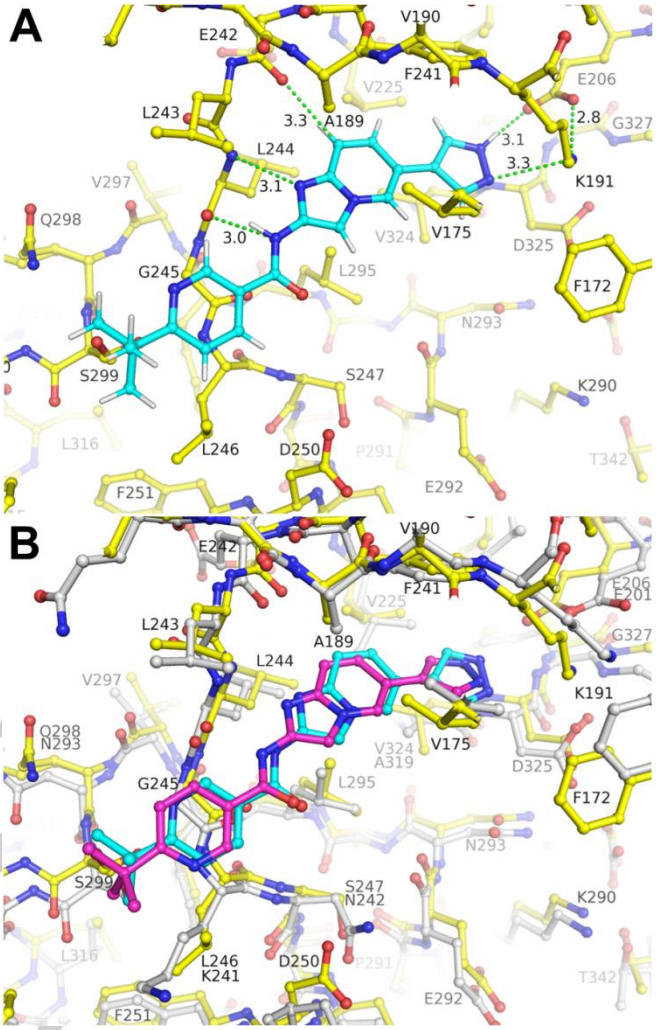


Figure 4. Comparison of interactions made by Cpd-2 in the ATP-binding sites of CLK1 and CLK3: Structural basis for selectivity against CLK3. *A*, X-ray structure at 2.32 Å resolution of the kinase domain of CLK1 (carbons in *yellow*, nitrogens in *blue*, oxygens in *red*) complexed with Cpd-2 (carbons in *cyan*, computed hydrogen positions in *white*). Water molecules are not shown. Selected interactions are shown as *green* dotted lines. *B,* Superposed on the X-ray structure of CLK1/Cpd-2 (coloring as in *A*) is the X-ray structure at 1.42 Å resolution of CLK3 (carbons in *white*) complexed with Cpd-2 (carbons in *magenta*). Upper/lower residue names for CLK1/CLK3 are indicated if there are residue differences. The coordinates for CLK1, CLK3 complexed with Cpd-2 have been deposited in the PDB databank (PDB access codes = 6FYO, 6FYR).

X-ray structures reveal the binding mode of Cpd-2 complexed with CLK1 and CLK3 kinase domains

As a further tool compound, Cpd-2 published by Araki et al. [7] was selected to determine the binding mode, and details of interactions for its pyrazole and *tert*-butyl-nicotinamide groups. The X-ray structures for Cpd-2 bound to CLK1 and CLK3 (at 2.32 Å, resp. 1.42 Å resolution) were obtained, shown in Fig. 4. Interestingly, the overall electron density for CLK1/ Cpd-2 showed no phosphorylation, in particular not at S341 or T341, and the latter loop conformation is consistent with no phosphorylation (despite the co-presence of a phosphorylated form after purification; i.e. crystallization selected the un-phosphorylated form).

The imidazopyridine core is sandwiched between A189, V175 of the N-terminal lobe and L295, V324 of the C-terminal lobe. In the hinge region, the imidazo-N accepts a hydrogen bond (3.0 Å) from NH-L244. In addition, the amide-NH donates a hydrogen bond (3.0 Å) with non-optimal geometry to CO-L244. The pyrazole makes an aromatic interaction with the gk residue F241 and also important vdW-interactions with V324. In addition, for CLK1/Cpd-2 the pyrazole makes strong direct hydrogen bond interactions with OE2-E206 and the catalytic lysine NZ-K193. Interestingly, for CLK3/Cpd-2 the pyrazole makes only indirect water mediated interactions with the corresponding E201 and K186 (Fig. 4*B* and supporting information Fig. S2). Correlating with this finding, for CLK3/Cpd-2 the pyrazole has a slightly different orientation than for CLK1, possibly caused by the

difference of V324 vs. A319 for CLK1 vs. CLK3. The X-ray structures thus predict that Cpd-2 should be slightly less potent for CLK3, compared to CLK1, CLK2, CLK4. Importantly, the *tert*-butyl group attached to the nicotinamide of Cpd-2 reaches towards a region (Fig. 10), where DYRK1A showed a significant difference to CLKs (Y243 for DYRK1A vs. L246 for CLK1, L248 for CLK2, K241 for CLK3, L246 for CLK4). In addition, derivatives in the *tert*-butyl region could be envisaged with further interactions, e.g. a hydrogen bond with OG-S299 of CLK1 or with OE1-E254 of CLK1. Interestingly, a piperazinylone derivative of Cpd-2 has been reported by the Takeda group to have high potency for CLK1 [8].

Taken together, our X-ray structures for the tool compounds TG003, CX-4945 and Cpd-2 showed the structural basis for possibilities to optimize selectivity and affinity in the ATP binding pocket. In parallel with the successful efforts to obtain CLK X-ray structures of various tool compounds from the literature, we had already initiated efforts to find novel highly potent CLK inhibitors that would be more selective over other kinase families than existing tool compounds.

Identification of a novel CLK2 inhibitor class using an *in silico* driven focused screen

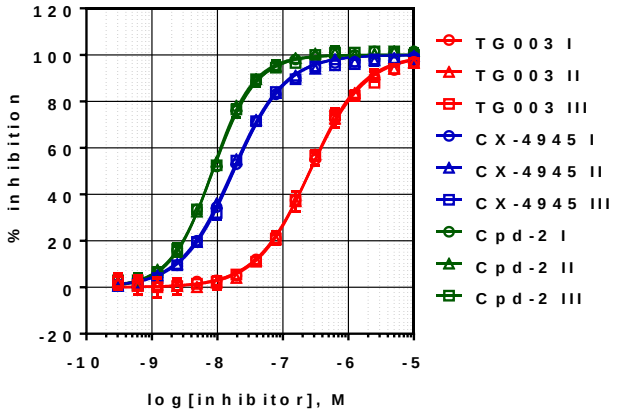
All four CLK isoforms were successfully expressed in *E. coli* and purified by a standard 2-step method. The crucial factor was the co-expression of λ-phosphatase which resulted in quite homogeneous protein preparations with no, or just a low degree of, λ-phosphorylation (supporting information Table S1). Without co-expression of the λ-phosphatase, the recombinant proteins were extensively phosphorylated (6 to > 10 phosphate residues) causing inhomogeneous elution during purification and a tendency to form high molecular weight aggregates. In addition, the closely related CLK2 and CLK3 were completely dephosphorylated in the presence of λ-phosphatase during expression, whereas CLK1 and 4 still showed some phosphorylation. More in-depth analysis performed by peptide mapping with tryptic digests of the purified proteins revealed that at least one of the two phosphate groups is attached to either S341 and/or T342 as shown in supporting information Fig. S6. Similarly CLK4 was phosphorylated at either S340 or T341. These two adjacent residues are located within the activation loop of the kinase. Further details are shown in supporting information.

To create an initial set for compound screening, two different virtual screening approaches were applied. (a) Structural information was leveraged by docking the internally available compound set into the binding site of CLK2. Since at that time point no internal structures were yet available, the structure of CLK2 (access code 3NR9 [9]) was downloaded from the PDB [10], and chain A prepared for docking experiments using the protein preparation workflow in MAESTRO. The obtained structure was used as input to two different docking engines – ICM [11] and GLIDE [12]. To obtain more diverse results from both methods, the settings for both were varied. Detailed information can be found in the supporting information. (b) Even though no activity information was available in house for CLK2, a ligand-based approach to hit list generation was followed. An internally developed machine learning technique (pQSAR [13]) based on random forests and partial-least squares was employed to build a virtual screening model. To train the model, activity data of roughly three thousand compounds against a closely related kinase was used as a surrogate for CLK2, since it was known from few compounds that selectivity between the kinases was low. Model quality for the data set was reasonable for virtual screening with an R^2_{Fit} of 0.75, a Q^2 of 0.75 and an external R^2 of 0.36. For this model, the hit list was filtered by a minimum predicted pIC50 of 5.

To reduce the size of the initial hit lists and remove obviously uninteresting compounds, duplicates were removed and the output of all methods was filtered by some simple properties (for details see supporting information) resulting in a final list of roughly 30K compounds.

Assay development of CLK2 mobility shift assay

Kinases are classical disease targets being addressed by a broad spectrum of assay technologies [14] enabling hit finding campaigns. In the case of CLK2, Caliper microfluidic technique was chosen due to its (i) high reliability, (ii) data quality, (iii) low compound interference effects and (iv) medium throughput capability. The assay principle is based on the different electrophoretic mobility of a fluorescent-label peptide after phosphorylation by CLK2 to follow the conversion of an un-phosphorylated substrate to phosphorylated product in real time. During the initial assay development phase, CLK2 kinase activity was characterized in regard to enzyme kinetic parameters as ATP K_m, enzyme time kinetic and DMSO sensitivity (experimental details shown in supporting information). CLK2 enzyme concentrations of 4 to 7 nM converted 15 to 25% of substrate after 60 minutes reaction time. The enzymatic reaction displayed a high linearity between 4 to 90 min reaction time (R^2=0.99). A substrate conversion rate of 15% ensured a reliable assay window being required for a robust screening assay with high data quality. The apparent ATP K_m values were determined in a time interval of 4 to 90 min generating ATP K_m values between 34 and 54 µM. The screening assay was performed at ATP K_m (50 µM) to guarantee a high sensitivity for CLK2 inhibitor identification. Since chemical screening libraries are dissolved in DMSO, its influence on CLK2 activity was analyzed. CLK2 activity was not significantly affected by DMSO concentrations up to 1%. Higher DMSO concentrations slightly decreased CLK2 activity, whereas the linearity of enzymatic reaction was not changed. In a last step, the CLK2 kinase assay was validated by analyzing the three CLK kinase inhibitors Cpd-2, CX-4945 and TG003 (Fig. 5). The assay generated robust data quality by showing highly reliable IC_{50} fitting curves. The determined IC_{50} values were in good agreement with the published IC_{50} data, confirming the validity of the CLK2 kinase activity assay (Fig. 5).



	TG003	CX-4945	Cpd-2
in house IC_{50}	260 +/- 2 nM	17.5 +/- 0.5 nM	8.5 +/- 0.1 nM
published IC_{50}	200 nM	3.8 nM	2.4 nM

Figure 5: CLK2 assay validation. IC_{50} fit curves of three reference inhibitors. Inhibitors tested at concentrations between 300 pM and 10 µM in three dilution series with duplicates. Green: Cpd-2, blue: CX-4945, red: TG003; circle: first dilution series, triangle: second dilution series, and square: third dilution series

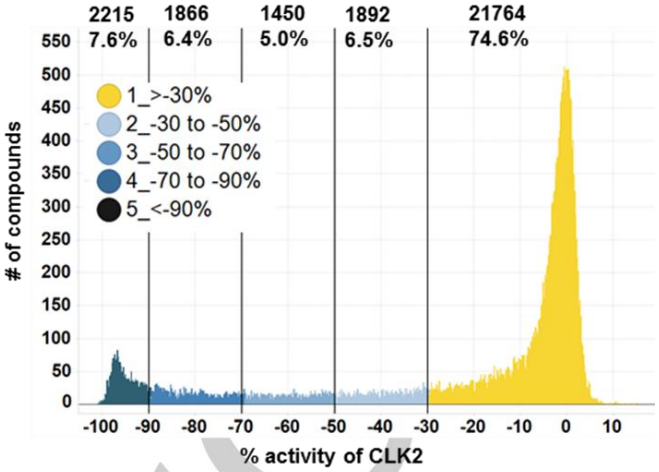


Figure 6. Enrichment of CLK2 inhibitors in primary screen. Activity distribution of hits in 30K screen. The different inhibitor potency ranges are colored from light blue (-30 to -40% activity) till dark blue (full inhibition -90 to 100%). On top of the diagram, the number of hits for the different potency ranges are displayed. Inactive hits are marked in yellow.

CLK2 30K screen and hit validation

A focused library of 30K compounds was assembled from the Novartis compound archive using the aforementioned *in silico* methods. The library was screened at 10 µM compound concentration. The screen displayed an excellent assay quality mirrored by (i) Z' > 0.5 for all 86 tested plates (average Z': 0.8) and (ii) a high reproducibility of IC_{50} values of 3 published CLK inhibitors, which were included in the beginning and the end of each screening run (assay quality data included in supporting information). The successful *in silico* compound selection of the focused screening set was underlined by (i) a high hit rate of 25.5% for compounds reducing CLK2 kinase activity by more than 30% and (ii) the enrichment of a potent inhibitor population around 100% inhibition (Fig. 6).

Initial hits were subjected to different filters e.g. chemical clustering, comparison with in house kinase selectivity data, and physicochemical properties to validate the most promising compounds (filter criteria, supporting information). Finally, 773 potential CLK2 inhibitors were selected for validation in 8-point concentration series. The predictive strength of the CLK2 assay was underlined by (i) the high confirmation rate of 81% and (ii) the good correlation between the potency (%inhibition) in the primary screen and the corresponding IC_{50} values obtained in validation. Almost half of all tested compounds displayed IC_{50} values < 1 µM. A deeper analysis of the validated hits and a SAR by archive led to the identification of several new CLK2 inhibitor classes. The indazole cluster was selected for detailed characterization by X-ray, biophysical, biochemical and cellular approaches to verify the binding mode and specificity of inhibitors.

High resolution X-ray structure reveals binding mode of new CLK2 inhibitor class exemplified by Indazole1

In order to elucidate the binding mode and interactions of the new indazole inhibitor class (as a basis for possible structure guided optimization) we embarked on an X-ray structure determination for the most potent inhibitor Indazole1. We could obtain co-crystals of CLK2(S136–D496)/Indazole1 which diffracted to 2.39 Å. The refined structure (supporting information Table S2) revealed the details of the interactions (Fig. 7*A*). The indazole core is sandwiched between A191 of the N-terminal lobe and L297, V326 of the C-terminal lobe. The benzyl part of the indazole also makes an aromatic interaction with the gk residue F243. In the hinge region, the 1-NH of the indazole donates a hydrogen bond (2.8 Å) to CO-E244 and the 2-N accepts a hydrogen bond (3.0 Å) from NH-L246. The 3-methyl extension on the indazole makes vdW-interactions with L169. Importantly, the pyridine-N of Indazole1 accepts a hydrogen bond (3.0 Å) from the catalytic lysine NZ-K193 (the position of K193 is further stabilized by a saltbridge with E208 and a water mediated hydrogen bond with D327). Interestingly, the 3-oxy ether oxygen, which is attached to the pyridine, accepts an intramolecular hydrogen bond (2.8 Å) from the amine NH3+, thus stabilizing the conformation of the 3-extension. Importantly, in this 3-extension the amine NH3+ donates a strong hydrogen bond (2.6 Å) to OD1-N295.

Indazole1 was submitted for testing in CLK1-4/DYRK1A radiometric protein kinase assays by ProQinase (see supporting information), with IC_{50} values of 12 nM (CLK1), 10 nM (CLK2), 2250 nM (CLK3), 12 nM (CLK4), and 73 nM (DYRK1A), showing the highest selectivity in inhibition of CLK2 vs CLK3 (225 fold) so far reported in literature. In order to explain the high selectivity over CLK3, it is important to consider that the small sidechain of residue A319 underneath provides less vdW-interactions with the aromatic core of the inhibitor than the larger V324 for CLK4 (V324 for CLK1, V326 for CLK2). Therefore, the interactions of

Indazole1 with NH-L246 from the hinge area and lysine NZ-K193 are most ideal in the case of CLK2 (and in analogy, CLK1 and CLK4), but not in case of CLK3. In the internal caliper kinase assay panel, Indazole1 had a CLK2 IC_{50} of 2.7 nM and showed high selectivity of at least 600-fold when tested against 34 kinases, apart from protein kinase A, where selectivity was only 70-fold (supporting information, Table S5).

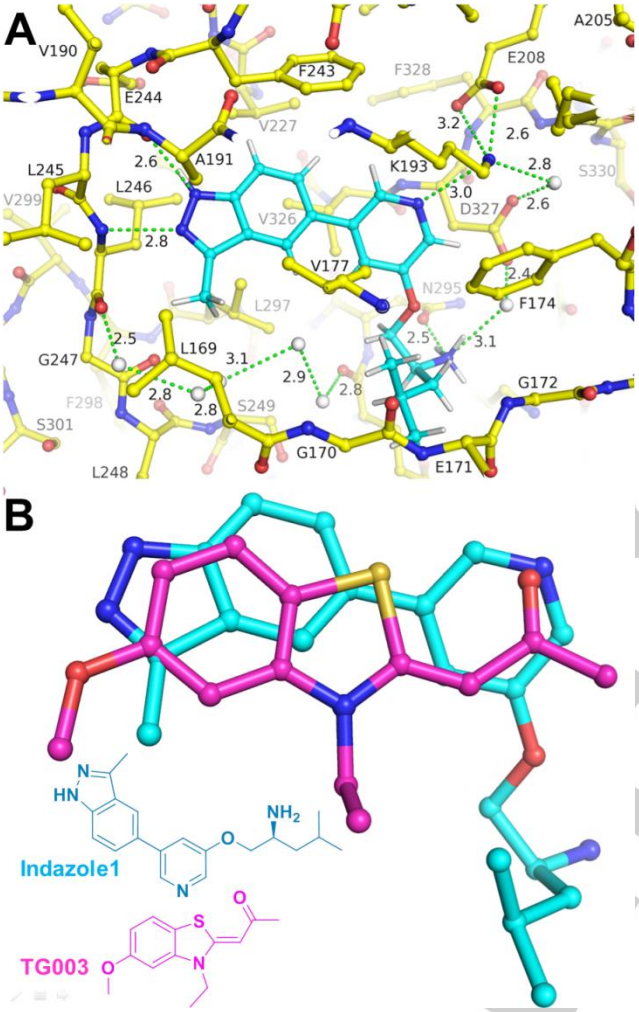


Figure 7. X-ray structure for CLK2/Indazole1 reveals binding mode and interactions, which explain the high potency of Indazole1 for CLK2 and why it is 96x more potent than TG003 in the CLK2 Caliper assay. *A*, X-ray structure at 2.39 Å resolution of the kinase domain of CLK2 (carbons in *yellow*, nitrogens in *blue*, oxygens in *red*) complexed with Indazole1 (carbons in *cyan*, computed hydrogen positions in *white*). Water molecules are shown as *white* spheres. Selected interactions are shown as *green* dotted lines. *B*, Superposed X-ray structures of TG003 (carbons in magenta) and Indazole1 (carbons in cyan), when complexed with CLK2. The coordinates for CLK2 complexed with TG003, Indazole1 have been deposited in the PDB databank (PDB access codes = 6FYI, 6FYK).

Indazole analogs of Indazole1 and initial SAR

All 4 analogs of Indazole1 from the indazole series identified via SAR by archive also showed a CLK2 IC_{50} in the Caliper assay below 50 nM, and stabilized CLK2 > 4 °C in a thermal shift assay (see table in Fig. 8A). All syntheses of Indazole2-5 can be performed in analogy to the synthesis of Indazole1, which is described in the supporting information. As demonstrated by modelling (Fig. 8B), various indazole functional side chains can form interactions with asparagine N295 or glutamic acid E171, and the potencies of Indazole1-5 were found to be in a similar range.

A

ID	Indazole1	Indazole2	Indazole3	Indazole4	Indazole5
CLK2 IC_{50} [nM]	3	25	35	6	15
CLK2 ΔT_m [°C]	7.3	5.2	4.1	7.7	5.6

B

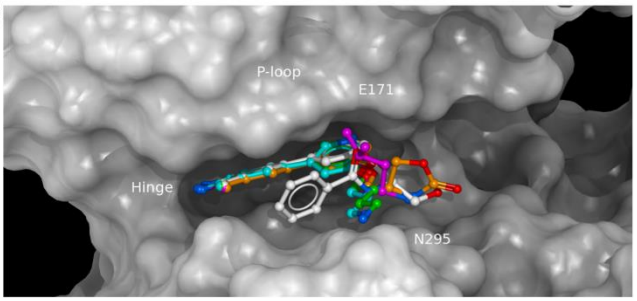


Figure 8. SAR for the CLK2 inhibitors Indazole 1-5. *A*, Table with IC_{50} values in Caliper assay and thermal shift assay data. *B*, Docking model of compounds Indazole1-5. Depending on the compound, interactions in different regions below the p-loop can be identified.

CLK2 inhibitors Indazole1 and Indazole2 rescue spine density in mouse brain slices

The tool compound TG003 was shown to rescue spine density in Shank3 deficient neurons at 10 µM concentration [1]. We used the same assay to evaluate whether Indazole1 and Indazole2 were able to rescue the loss of spine density at lower concentrations. Remarkably, Indazole2 significantly improved spine density at 1 µM concentration, whereas the most potent CLK2 inhibitor Indazole1 already showed efficacy at a concentration of 300 nM (Fig. 9). These results indicate a 10-fold and 30-fold increase of potency for Indazole2 and Indazole1, respectively, compared to TG003.

We next focused on safety related questions regarding CLK inhibition, in particular due to a publication which has linked CLK1, CLK2, and CLK4 inhibition with genotoxicity via a mechanism of action on chromosomal abscission [15].

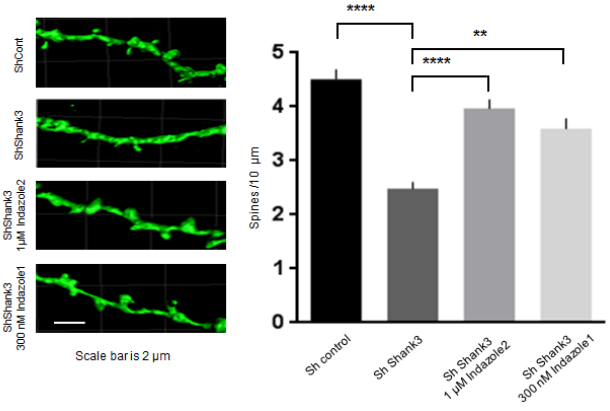


Figure 9. Indazole1 and Indazole2 rescue spine density in Shank3 knockdown neurons. *Left,* representative images of CA1 apical secondary

dendrites treated on DIV 14 for 24 hrs with Indazole1 or Indazole2 prior to immunostaining for GFP. *Right*, quantitative analysis with plots indicating group means ± SEM (one-way ANOVA, p<0.0001, Dunn's multiple comparisons test).

MNT measurements of TG003 and Indazole1 in primary human lymphocytes to assess *in vitro* genotoxicity of CLK inhibitors

To assess potential clastogenic and/or aneugenic effects of the CLK inhibition, two potent CLK inhibitors from two different chemical series were tested in the cytokinesis-blocked micronucleus (MN) assay in primary human lymphocytes. The treatment duration was 28 hours continuous incubation in the absence of S9 metabolic activation. Both compounds were tested up to clearly cytotoxic concentrations. In the experiment with Indazole1, a significant concentration-dependent increase in the frequencies of micronucleated binucleates (MNBN) was seen at all analyzed concentrations ranging from moderate to high cytotoxicity (supporting information, MNBN Table S6). TG003 also induced a significant concentration-dependent increase in the frequencies of MNBN (supporting information, MNBN Table S7). The increases reached statistically significant levels at the two highest analyzed concentrations that showed moderate to high cytotoxicity. Our data show that both compounds are genotoxic in the MN test. These results are in alignment with data already published by Petsalaki et al.[15], where TG003 or depletion of CLK1, CLK2 or CLK4 increased the frequency of micronucleated cells. Indazole1, a more potent inhibitor of CLKs than TG003, was also more potent in inducing micronuclei in reaching higher frequencies of MNBN at relatively low concentrations, correlating with IC_{50} values for CLKs. Therefore our data confirm the hypothesis that inhibition of CLKs 1, 2 and 4 confers a genotoxic potential.

Conclusions

One of our initial questions was whether CLK inhibition would be a feasible approach to treat PMDS. By leveraging a combination of orthogonal *in silico* screening methods based on structural information as well as machine learning, we were able to quickly identify a large set of different target scaffolds to answer this question. The high screening hit rate substantiates that this approach is not only feasible but can replace full deck screening. The resulting CLK inhibitors Indazole1 and Indazole2 showed efficacy on spine density restoration in Shank3 deficient neurons. However, the data on the induction of micronuclei in human lymphocytes, obtained with two compounds from different chemical series, TG003 and Indazole1, confirm the hypothesis that CLK 1, 2 or 4 inhibition *per se* will induce a genotoxic effect, which is very plausible given the mechanistic involvement of the pathway in chromosome integrity [15]. Inhibition of CLK1, CLK2, and CLK4 is therefore concluded not to be a safe strategy to treat PMDS.

Concerning the potential to reach selectivity, a comparison of the various in house X-ray structures for CLK-family members (Fig. 10) showed us that L248, S249, S301, V326 in the ATP binding site of CLK2 are possible candidates for compound interactions to obtain a certain degree of selectivity (e.g. vs. CLK3 or DYRK1A). CLK1, CLK2 and CLK4 have a valine at the V326 position of CLK2, whereas CLK3 has the smaller alanine A319 at the corresponding position. On the other hand, we concluded that selectivity for CLK2 vs. CLK1 or CLK4 would be challenging to obtain. In addition to hinge binding interactions, interactions with the CLK2 residues K193 (e.g. carboxylate, pyrazole), E208 (e.g. carboxylate, pyrazole), F243 (aromatic interaction), F174 (e.g. Cl edge interactions), L248 (vdW-interactions) are candidates for obtaining high affinity.

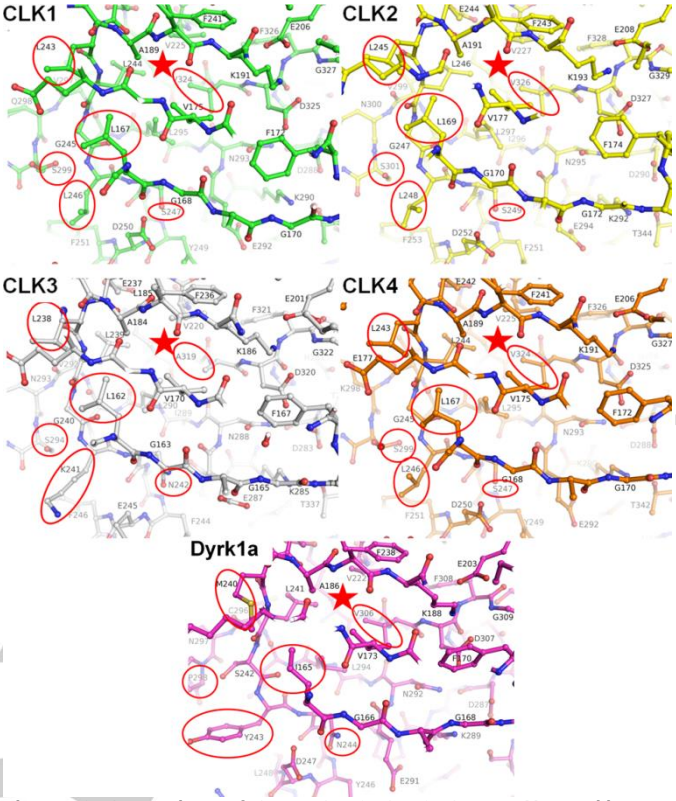


Figure 10. Comparison of CLK1, CLK2, CLK3, CLK4, DYRK1a kinase domain X-ray structures in ATP-sites. *A*, CLK1 (from complex with Cpd-2). *B*, CLK2 (from complex with CX-4945). *C*, CLK3 (from complex with CX-4945). *D*, CLK4 (from complex with CX-4945). *E*, Dyrk1a, published by Falke et al. [16]. The *red* star indicates the position where CLK1, CLK2 and CLK4 have a valine (V326 for CLK2), whereas CLK3 has the smaller alanine A319. The coordinates for CLK1/Cpd-2, CLK2/CX-4945, CLK3/CX-4945, CLK4/CX-4945 have been deposited in the PDB databank (PDB access codes = 6FYO, 6FYL, 6FYP, 6FYV).

Besides our own focus on CLK2 as a molecular target for the treatment of autistic features in PMDS patients with SHANK3 deletions [1], various groups have suggested functional modulation of the closely related kinase DYRK1A as a potential therapeutic modality in addressing the cognitive and behavioral deficits in Down syndrome [6]. We hope this work may prove helpful in the development of ATP-competitive DYRK1A inhibitors with the likely required selectivity over CLK1, CLK2, and CLK4, which are associated with adverse genotoxic potential.

Possible applications of CLK2 inhibitors in the treatment of certain cancer indications (e.g. breast cancer) are also still of high interest. Our new X-ray structures should prove to be useful for the design and/or optimization of further CLK-family inhibitor series for cancer indications, with respect to affinity, selectivity and detailed molecular mechanism.

Experimental Section

Expression and purification of human CLK1-4.

Cloning: The cDNAs of the four CLK kinase domains were codon-optimized in-house and synthesized by GeneArt AG (Regensburg, Germany). The resulting products (Invitrogen™ GeneArt™ Strings™ DNA Fragments) were cloned into a

proprietary E. coli/ T7 expression vector featuring an amino terminal hexahistidine tag (His6-tagged) followed by human rhinovirus 3C protease cleavage site (LEVLFQ'GP). The boundaries of the kinase domains were set as follows: CLK1 (Uniprot P49759-1) H148-I484, CLK2 (Uniprot P49760-1) R130-D496 or S136-D496 respectively, CLK3 (Uniprot P49761-4), Q275-T632, and CLK4 (Uniprot (Q9HAZ1-1) H146-K480.
The gene encoding full-length λ-phosphatase from lambda phage was cloned either into a pET derived expression vector or into an auxiliary plasmid for co-expression. The latter is based on the plasmid pACY184 featuring an alternative origin of replication, a different resistance marker (chloramphenicol instead of kanamycin) and the arabinose controlled BAD promoter.
Expression: BL21(DE3) Tuner *E. coli* cells (Novagen) were transformed with the individual plasmids coding for CLK isoform 1 to 4 and additionally the pACY184 vector containing the gene for λ-phosphatase in order to achieve an in vivo de-phosphorylation of the expressed kinases. The bacterial cells were cultivated in Terrific Broth (TB) culture media complemented with 0.1 M morpholino-ethanesulfonic acid (MOPS), 30 mg/l kanamycin, 30 mg/ml chloramphenicol and 2 mM arabinose (for inducing the expression of the for λ-phosphatase). Cells were grown at 37^{o} C until an optical density of 0.8 and induced with 0.1 mM isopropyl-thiogalactoside (IPTG) for induction for expression; the temperature was lowered to 20^{o} C and cells further incubated overnight.
Purification: *E. coli* cells were harvested by centrifugation and lysed with a high pressure homogenizer device. The His6-tagged CLK isoforms were all purified in a first step with immobilized metal affinity chromatography (IMAC) by using a 5 ml HisTrap column (GE Healthcare). The column was washed with loading buffer (50 mM Na-phosphate pH 8.0, 300 mM NaCl, 20 mM imidazole, 10 % glycerol and 0.5 mM TCEP) for several column volumes, once the sample was loaded. The His6-tagged kinases were eluted in an imidazole gradient ranging to 300 mM within 5 column volumes. For cleavage of the N-terminal His6-tag fractions containing sufficiently pure protein were pooled, and in-house produced HRV-3C protease (PreScission protease) was added in molar ratio of 1:100. The mixture was dialyzed against size exclusion chromatography (SEC) buffer (50 mM HEPES pH 7.4, 300 mM NaCl, 10 % glycerol and 0.5 mM TCEP) overnight at 4^{o} C. For separation of the cleaved fusion tag and as a final purification step the volume of the dialysis mixture was concentrated to approximately 5 ml and loaded onto a Superdex 75 (16/60 or 26/60) size exclusion column using the SEC buffer. Typically, the cleaved kinases eluted in a symmetric major peak around 0.5 column volumes, preceded by a smaller peak containing aggregated protein. Fractions from the large peak were pooled and deep frozen at -80°C in aliquots until further use.

CLK2 kinase activity assay

The *in vitro* CLK2 kinase assay was established using Caliper mobility shift assay technology (Perkin-Elmer, Waltham, MA). In house recombinant CLK2 kinase domain (130-496) was used to establish Caliper kinase assay. The kinase assay buffer mix contained 50 mM HEPES pH 7.5, 1 mM DTT, 0.02% Tween 20, 0.02% BSA, 10 mM beta-glycerophosphate, 10 µM sodium orthovanadate, 6 nM CLK2, 3 mM $MgCl_2$, 2 µM peptide (FL-26, 5-FAM-ARKRERTYSFGHHA-COOH, PerkinElmer, Waltham, MA). The assay buffer mix was added with a Thermo Multidrop Combi dispenser to 384well microtiter well assay plates, containing 40 nl 5 mM pre-spotted compound solution. The assay was run at 10 µM compound screening concentration (total number of compounds: 30129). CLK2 was pre-incubated with the test compounds for 45 minutes and then the enzymatic reaction was started by adding ATP (final concentration 50 µM). After 60 minutes, the kinase reaction was stopped by EDTA addition. Then, the plates were analyzed on a Caliper system.

Data analysis and visualization

K_m and IC_{50} values were determined using Prism software 7.03 (Graphpad Software Inc. San Diego, CA). Larger primary screening and IC_{50} validation raw data files were analyzed using an in house software. The data was visualized using Spotfire Decision Site 8.1 (Spotfire). The data of each screening plate was normalized by plate specific controls, (i) active control (AC, n=16), 100% inhibition and (ii) neutral control (NC, n=16), 0% inhibition using following equation: %inhibition = -100x(1-(x-AC)/(NC-AC)). Furthermore, the plate controls were used to evaluate the assay quality by calculating the Z'-factor: Z'= 1-$(3xSD_{NC} + 3xSD_{AC})/(mean_{NC} – mean_{AC})$

Differential scanning fluorimetry (DSF) assays to determine thermal stability of CLK proteins

Conformational stability reflects a common property of proteins which is influenced by the interaction with low molecular weight ligands. Thermal shift assays formats like differential scanning fluorimetry (DSF) aka ThermoFluor™ analyzing thermal stability of proteins enable the identification and characterization of ligand binding events. 384well microtiter well plate DSF assays for CLK proteins were developed using a CFX384™ real-time PCR instrument (Bio-Rad, Hercules, CA). Heat induced unfolding of CLK proteins was followed by measuring the fluorescence increase caused by a solvatochromic dye SYPRO®Orange (ThermoFisher Scientific™, Waltham, MA). The standard assay conditions were as follows: 50 mM HEPES pH 7.4, 150 mM NaCl, 0.5 mM TCEP4, 1% DMSO, 1.9 µM CLK1, CLK2, CLK3 or CLK4 and 5 µM SYPRO®Orange. All DSF assays were performed in 10 µl volume using sealed PCR plates. The CLK samples were heated from 20 to 60^{o}C with a heating rate of 1^{o}C/min and every 0.5^{o}C the fluorescence was detected by using the FRET channel (excitation: 450 - 490 nm; emission: 560 - 580 nm). To determine the stabilization of CLK proteins by ligands, control wells containing the proteins without ligand and 1% DMSO were included in each experiment. The melting temperature (T_m) was determined by plotting the fluorescence values against temperature and fitting the data to the Boltzmann equation using an in-house software program. The inflection point of the first derivative of the melting curve represented the T_m. Melting temperature shifts (ΔT_m), a measure to quantify the protein stabilization, was calculated by $\Delta T_m = T_m(CLK_{ligand}) – T_m(CLK_{DMSO})$. A ΔT_m > 3-fold standard deviation of T_m (CLK_{DMSO}) was classified as a significant stabilization. The standard deviation of CLK T_m was in the range of 0.1 to 0.2^{o}C for all experiments enabling the identification of hits causing a ΔT_m > 0.5^{o}C. For the compound tests, 40 nl of 2 mM and 10 mM compound solution was spotted into PCR plates (final concentrations of 8 and 40 µM) and mixed with 10 µl of CLK/dye solution. The assay reaction was pre-incubated for 30 min at room temperature before analysis in the PCR instrument.

Crystallization

Co-crystals of CLK1(148-484) (numbering according to P49759) in complex with Cpd-2 were obtained by adding a 3-fold molar excess of compound (all complexes were incubated ca. 2-3 hours at 4°C and centrifuged before crystallization). The reservoir solution consisted of 200 mM AS, 100 mM Bis-Tris pH5.5, 25% (w/v) PEG3350. Co-crystals of CLK2(130-496) (numbering according to P49760) in complex with TG003 were obtained by adding a 3-fold molar excess of compound. The reservoir solution consisted of 100 mM HEPES pH7.5,

25% (w/v) PEG3350. With the aim to increase the resolution for CLK2, we redesigned the CLK2 construct based on our CLK2/TG003 crystal structure by removing a flexible part at the NH2-terminus to possibly obtain a different crystal form. Co-crystals of CLK2(136-496) (numbering according to P49760) in complex with compounds Indazole1, CX-4945 were obtained by adding a 3-fold molar excess of compound. For Indazole1, the reservoir solution consisted of 200 mM trimethylamine n-oxide, 100 mM Tris pH8.5, 20% (w/v) PEG2000 MME. For CX-4945, the reservoir solution consisted of 100 mM HEPES pH7.0, 30% (v/v) Jeffamine ED-2001. Co-crystals of CLK3(275-632) (numbering according to P49761) in complex with compounds CX-4945, Cpd-2 were obtained by adding a 3-fold molar excess of compound. For CX-4945, the reservoir solution consisted of 200 mM AS, 100 mM Bis-Tris pH5.5, 25% (w/v) PEG3350. For Cpd-2, the reservoir solution consisted of 100 mM Hepes pH7.5, 25% (w/v) PEG4000. Co-crystals of CLK4(146-480) (numbering according to Q9HAZ1) in complex with compound CX-4945 were obtained by adding a 3-fold molar excess of compound. The reservoir solution consisted of 50 mM AS, 50 mM Bis-Tris pH6.5, 30%(v/v) Pentaerythritol ethoxylate (15/4 EO/OH). Co-crystals were obtained at 20 °C and by sitting drop vapor diffusion. The drops were made up of 200 nL of protein solution and 200 nL of well solution.

Data collection, structure determination and refinement

All crystals were cryo-protected in well solution supplemented with 20% ethylene glycol and flash frozen in liquid nitrogen. Data sets were collected at the Swiss Light Source Facility (SLS, Villigen, Switzerland) for all the crystals. The data were processed with XDS [17]. The X-ray structures of CLK complexes were determined by molecular replacement (PHASER [18]) using various published and in house CLK X-ray structures as the search model. Programs REFMAC [18] and COOT [19] were used for refinement and model (re)building. The final refined structures have R (Rfree) values ranging from 0.176 to 0.228 (0.189 to 0.267) and showed excellent geometry in the Ramachandran plot. All residues were in the allowed regions for all X-ray structures, except for ca. 1% of the residues (for which the electron density confirms the local structure). As an example, for the complex CLK2(136-496) and Indazole1 excellent electron density around R162, T289, D327, S343, S359 unequivocally confirmed the local structure. See supporting information Tables S2, S3, S4 for details of the data collection and structure refinement. The crystallographic data for the seven X-ray structures have been deposited at the RSCB Protein Data Bank (PDB, www.pdb.org) with the respective access codes 6FYI, 6FYK, 6FYL, 6FYO, 6FYP, 6FYR, 6FYV. All of the figures showing crystal structures were prepared using PyMOL (Schrödinger, LLC (2015) *The PyMOL Molecular Graphics System*, Version 1.8, New York).

Genotoxicity assay in primary human lymphocytes:

Test system: The study was conducted to determine the clastogenic and/or aneugenic potential of Indazole1 and TG003 in the micronucleus test *in vitro* with cultured human peripheral blood lymphocytes without S9-liver metabolic activation.

The *in vitro* cytokinesis-block micronucleus assay in human lymphocytes has the advantage of detecting in interphase both acentric chromosome fragments due to DNA breakage (clastogenicity) and chromosome loss (aneuploidy) resulting from chromatid/chromosome lagging in anaphase. In combination with the cytokinesis-block method using cytochalasin B [20], the assay allows the assessment of cell toxicity/cell cycle delay. Indeed cytochalasin B, an inhibitor of actin polymerization, blocks cytokinesis without interfering with nuclear division [21] producing two daughter nuclei in the same cytoplasm. Therefore, in the presence of the inhibitor, cells that have completed one cell cycle after treatment will appear as a binucleated cell, cells that divided more than once will appear as polynucleated cells (trinucleated and tetranucleated cells) and finally the cells that did not divide at all, will remain mononucleates. For each substance at least three concentrations (2 cultures per concentration) and 2000 cells per concentration were analyzed. Per culture 1000 binucleated lymphocytes were analysed for the presence of one, two or more micronuclei (MN) and expressed as per thousand micronucleated binucleates (MNBN). Also the percentage of binucleates, polynucleated cells, mitotic cells and mononucleated cells with or without MN were recorded.

Cytotoxicity parameter: As a measure of cytotoxicity and cell cycle delay the relative division index was used. The criteria for determining a positive result is based on the reproducible increase in the number of MNBN and on statistical differences between controls and treated samples as determined with the Chi-square test (95% confidence limit is used).

Acknowledgements

We thank Ulrich Schopfer, Tewis Bouwmeester, Fabrizio Gasparini and Ricardo Dolmetsch for their advice and support of our project. We thank Neil Press for critically reading the manuscript and helpful comments. We thank Eric Martin for his virtual screening model and advice for our project. The crystallographic experiments were performed on the X10SA beamline at the Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland.

The authors declare that they have no conflict of interests with the contents of this article.

Key words: CLK2 • CLK-family • DYRK1A • kinase inhibitor • X-ray structure • virtual screening • mobility shift assay • indazoles • Phelan-McDermid syndrome • genotoxicity • alternative splicing

The atomic coordinates and structure factors (codes 6FYI, 6FYK, 6FYL, 6FYO, 6FYP, 6FYR, 6FYV) have been deposited in the Protein Data Bank (http://wwpdb.org/).

References

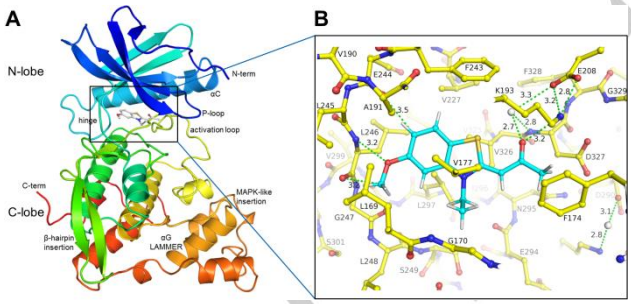
[1] M. Bidinosti, P. Botta, S. Kruttner, C. C. Proenca, N. Stoehr, M. Bernhard, I. Fruh, M. Mueller, D. Bonenfant, H. Voshol, W. Carbone, S. J. Neal, S. M. McTighe, G. Roma, R. E. Dolmetsch, J. A. Porter, P. Caroni, T. Bouwmeester, A. Luthi, I. Galimberti, *Science* **2016**, *351*(6278), 1199-1203.

[2] M. Muraki, B. Ohkawara, T. Hosoya, H. Onogi, J. Koizumi, T. Koizumi, K. Sumi, J. Yomoda, M. V. Murray, H. Kimura, K. Furuichi, H. Shibuya, A. R. Krainer, M. Suzuki, M. Hagiwara, *J Biol Chem* **2004**, *279*(23), 24246-24254.

[3] a) T. C. Coombs, C. Tanega, M. Shen, J. L. Wang, D. S. Auld, S. W. Gerritz, F. J. Schoenen, C. J. Thomas, J. Aube, *Bioorg Med Chem Lett* **2013**, *23*(12), 3654-3661; b) O. Fedorov, K. Huber, A. Eisenreich, P. Filippakopoulos, O. King, A. N. Bullock, D. Szklarczyk, L. J. Jensen, D.

Fabbro, J. Trappe, U. Rauch, F. Bracher, S. Knapp, *Chem Biol* **2011**, *18*(1), 67-76; c) K. Prak, J. Kriston-Vizi, A. W. Chan, C. Luft, J. R. Costa, N. Pengo, R. Ketteler, *Biochemistry* **2016**, *55*(3), 608-617.
[4] D. Fabbro, S. W. Cowan-Jacob, H. Moebitz, *Br J Pharmacol* **2015**, *172*(11), 2675-2700.
[5] H. Kim, K. Choi, H. Kang, S. Y. Lee, S. W. Chi, M. S. Lee, J. Song, D. Im, Y. Choi, S. Cho, *PLoS One* **2014**, *9*(4), e94978.
[6] H. Kim, K. S. Lee, A. K. Kim, M. Choi, K. Choi, M. Kang, S. W. Chi, M. S. Lee, J. S. Lee, S. Y. Lee, W. J. Song, K. Yu, S. Cho, *Dis Model Mech* **2016**, *9*(8), 839-848.
[7] S. Araki, R. Dairiki, Y. Nakayama, A. Murai, R. Miyashita, M. Iwatani, T. Nomura, O. Nakanishi, *PLoS One* **2015**, *10*(1), e0116929.
[8] T. Funnell, S. Tasaki, A. Oloumi, S. Araki, E. Kong, D. Yap, Y. Nakayama, C. S. Hughes, S. G. Cheng, H. Tozaki, M. Iwatani, S. Sasaki, T. Ohashi, T. Miyazaki, N. Morishita, D. Morishita, M. Ogasawara-Shimizu, M. Ohori, S. Nakao, M. Karashima, M. Sano, A. Murai, T. Nomura, N. Uchiyama, T. Kawamoto, R. Hara, O. Nakanishi, K. Shumansky, J. Rosner, A. Wan, S. McKinney, G. B. Morin, A. Nakanishi, S. Shah, H. Toyoshiba, S. Aparicio, *Nat Commun* **2017**, *8*(1), 7.
[9] *Structure of human CDC2-like kinase 2 (CLK2)*, http://www.rcsb.org/structure/3NR9.
[10] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res* **2000**, *28*(1), 235-242.
[11] R. Abagyan, M. Totrov, D. Kuznetsov, *Journal of Computational Chemistry* **1994**, *15*(5), 488-506.
[12] a) R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin, *J Med Chem* **2004**, *47*(7), 1739-1749; b) T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, J. L. Banks, *J Med Chem* **2004**, *47*(7), 1750-1759.
[13] E. J. Martin, V. R. Polyakov, L. Tian, R. C. Perez, *J Chem Inf Model* **2017**, *57*(8), 2077-2088.
[14] Y. Wang, H. Ma, *Drug Discov Today Technol* **2015**, *18*, 1-8.
[15] E. Petsalaki, G. Zachos, *Nat Commun* **2016**, *7*, 11451.
[16] H. Falke, A. Chaikuad, A. Becker, N. Loaec, O. Lozach, S. Abu Jhaisha, W. Becker, P. G. Jones, L. Preu, K. Baumann, S. Knapp, L. Meijer, C. Kunick, *J Med Chem* **2015**, *58*(7), 3131-3143.
[17] W. Kabsch, *Acta Crystallogr D Biol Crystallogr* **2010**, *66*(Pt 2), 125-132.
[18] N. Collaborative Computational Project, *Acta Crystallogr D Biol Crystallogr* **1994**, *50*(Pt 5), 760-763.
[19] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr D Biol Crystallogr* **2010**, *66*(Pt 4), 486-501.
[20] M. Fenech, A. A. Morley, *Mutat Res* **1985**, *147*(1-2), 29-36.
[21] S. B. Carter, *Nature* **1967**, *213*(5073), 261-264.

Entry for the Table of Contents



CLK2 inhibition has been proposed as a potential mechanism to improve autism and neuronal functions in Phelan-McDermid syndrome (PMDS). Herein, we report the discovery of a very potent indazole CLK inhibitor series, and the CLK2 X-ray structure of its most potent analog. Novel high resolution X-ray structures of all CLKs, including the first CLK4 X-ray structure, bound to known CLK2 inhibitor tool compounds (e.g. TG003, CX-4945), are also shown and yield insights into inhibitor selectivity in the CLK family. Efficacy of new CLK2 inhibitors from the indazole series was demonstrated in the mouse brain slice assay, and potential safety concerns were investigated. We show genotoxicity findings in the human lymphocyte MNT assay using two structurally different CLK inhibitors, which reveals a major concern for pan-CLK inhibition in PMDS.