VIP Computer-Based Drug Design

DOI: 10.1002/anie.201206897

Drugs by Numbers: Reaction-Driven De Novo Design of Potent and Selective Anticancer Leads**

*Birgit Spänkuch, Sarah Keppner, Lisa Lange, Tiago Rodrigues, Heiko Zettl, Christian P. Koch, Michael Reutlinger, Markus Hartenfeller, Petra Schneider, and Gisbert Schneider**

In light of recent analyses of drug development,[1,2] new lead compounds with well-defined pharmacological activity profiles are urgently sought.[3,4] Computer-based de novo design has been suggested as a method of choice to meet this need, as it generates innovative molecular scaffolds and chemotypes by accessing virtually infinite chemical space.[5–7] Here we present the successful application of fully automated chemistry-driven de novo design to discover an innovative low-molecular-weight inhibitor that selectively blocks inactive human Polo-like kinase 1 (hPlk1) with nanomolar potency. This potential anticancer compound was generated by the algorithm "from scratch" and synthesized following the exact reaction scheme suggested by our software. It reduced cancer-cell proliferation without affecting the vitality of nontransformed cells, and exhibited no inhibitory effects against a panel of activated kinases. The computationally designed compound is a derivative of the antidepressant fluoxetine, for which we observed a similar but weaker cellular response profile. This study provides proof-of-concept for de novo design as a leading tool for generating novel chemotypes in the absence of a structural model of the target protein and with minimal experimental effort.

The serine–threonine kinase hPlk1 plays a central role in cell cycle control and is a target for the development of novel cancer therapeutics.[8,9] While type I inhibitors bind to an active conformation of the kinase, type II inhibitors block the inactive kinase, in which the enzyme's activation loop is in the so-called "DFG-out" conformation.[10,11] We have recently identified compound **1** as a strong ($IC_{50} = 0.2$ nM) inhibitor of inactive hPlk1.[12] With compound **1** as a design template, we employed our newly developed software DOGS (*Design Of Genuine Structures*) for generating alternative chemotypes that mimic the pharmacophoric features of the template molecule but contain structurally distinct scaffolds.[13,14] The software constructs molecules by applying a set of 83 chemical reactions to a stock of over 25 000 readily available molecular building blocks.[15] Compounds are prepared in silico through iterative application of motivated synthetic reactions.[16] In each step along the virtual construction path, conserving the pharmacophoric features and maintaining the structural similarity of the growing ligand candidate to a template structure (here compound **1**) guide compound prioritization.[17] Among the best designs, we identified compound **4** as a chemotype extending the structural diversity of known kinase inhibitors (Figure 1 A). Surprisingly, this chemically attractive de novo designed compound is a structural analogue of fluoxetine, a well-known antidepressant.[18,19] This observation points to functional similarities between the two compounds, which motivated us to investigate compound **4** in more detail.

The DOGS software produced a total of 218 compounds from 100 preferred starting fragments. The designs are computed to have druglike properties and are synthetically plausible (mean $\pm\sigma$: molecular weight $= 457 \pm 59$ Da, lipophilicity ($S\log P$) $= 4.5 \pm 1.1$, aqueous solubility ($\log S$) $= -5.9 \pm 1.3$, synthetic plausibility ($rsynth$) $= 0.7 \pm 0.3$). Hopkins' quantitative estimate of druglikeness (QED)[20] for the set of designed compounds ($QED = 0.5 \pm 0.2$) is in agreement with the average value of 0.49 obtained for approved drugs.[21] In total, 57 different molecule scaffolds were generated, with 57% of the designs containing one of the ten most frequent scaffolds (Figure 1 B). This broad scaffold diversity reflects the permissive pharmacophoric similarity measure that was applied during the design process. Among the best ranking designs we observed several branched structure motifs although the backbone of the template **1** is linear. In a preliminary study we had synthesized a de novo designed compound with a linear scaffold from this series, which exhibited pronounced inhibitory activity against inactive hPlk1.[14] Here, we focused on the branched compound **4**.

We computed a fitness landscape representing a probabilistic model of the structure–activity relationships of known kinase inhibitors (Figure 1 C). The landscape represents a visualization of the distribution of 12 647 bioactive com-

[*] Dr. T. Rodrigues, Dr. H. Zettl, C. P. Koch, M. Reutlinger, Dr. M. Hartenfeller, Dr. P. Schneider, Prof. Dr. G. Schneider
Departement Chemie und Angewandte Biowissenschaften
Eidgenössische Technische Hochschule (ETH)
Wolfgang-Pauli-Strasse 10, 8093 Zürich (Switzerland)
E-mail: gisbert.schneider@pharma.ethz.ch
Homepage: http://www.modlab.ethz.ch
Priv.-Doz. Dr. B. Spänkuch, Dr. S. Keppner, L. Lange
Universitätsfrauenklinik, Molekulare Onkologie und Gynäkologie
Eberhard Karls Universität
Calwerstrasse 7, 72076 Tübingen (Germany)

[**] We thank Dr. M. Bieler for computing *QED* values and Sarah Haller for technical support. Dr. T. Geppert performed the computational docking experiment. Dr. M. Rupp contributed the ISOAK similarity function to DOGS. This research was financially supported by the Deutsche Krebshilfe (grant no. 108651), the Wilhelm-Sander-Stiftung (grant no. 2009.024.1), the Messer-Stiftung (to B.S.), the Swiss National Science Foundation (grant no. 205321-134783), and the Deutsche Forschungsgemeinschaft (grant no. FOR1406TP4) (to G.S.). B.S. is grateful to Dr. D. Wallwiener for providing working facilities. G.S. is grateful to the Chemical Computing Group, Inc. (Montreal, Canada) for a research license of MOE.

Supporting information for this article (including full experimental details) is available on the WWW under http://dx.doi.org/10.1002/anie.201206897.

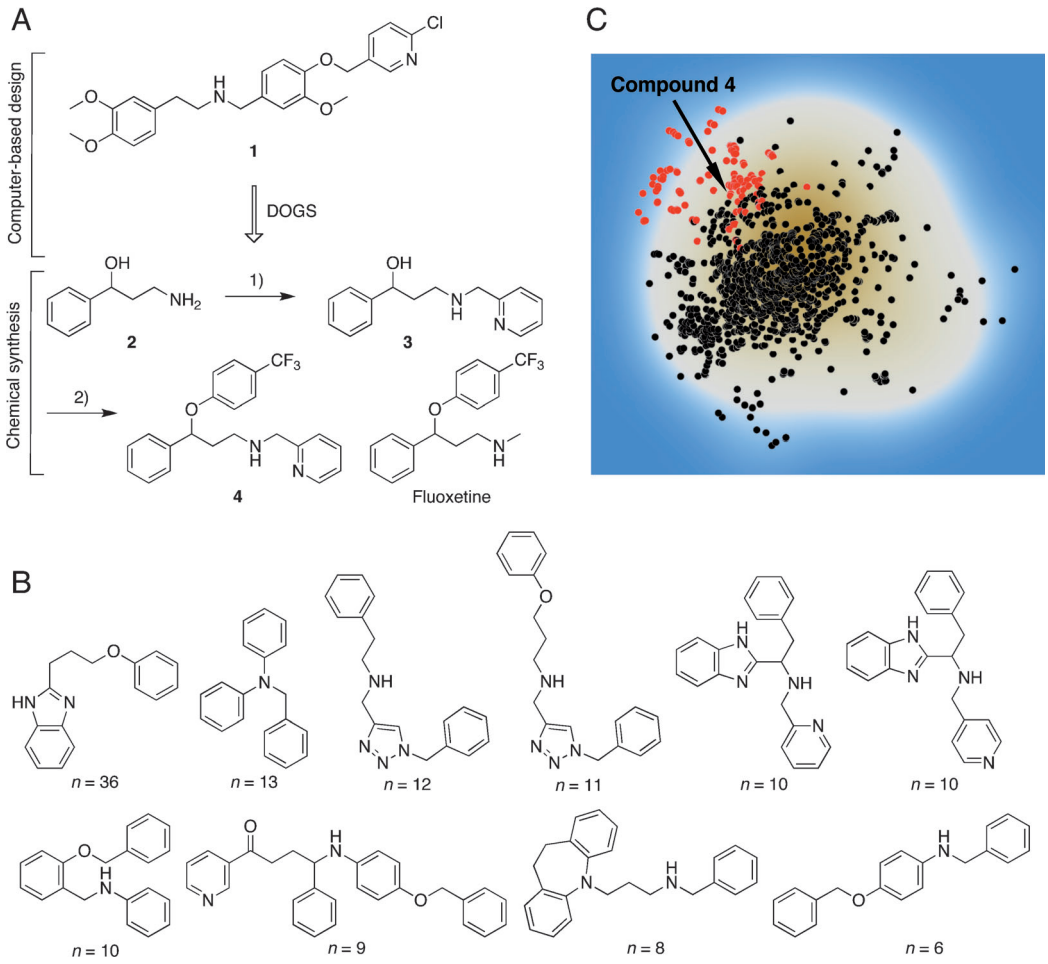

Figure 1. A) De novo design and synthesis of compound **4**: Reagents and conditions: 1) 2-pyridinecarboxaldehyde, 1,2-dichloroethane (DCE), $NaBH(OAc)_3$, RT; 2) 4-trifluoromethylphenol, PPh_3, diethyl azodicarboxylate (DEAD), THF, 0 °C. The designed compound is a derivative of fluoxetine. B) Scaffold diversity: The 10 most frequently generated scaffold structures among the computer-designed compounds (*n*: number of molecules). C) Artificial fitness landscape: The plot presents the distribution of known kinase inhibitors (black dots) and the de novo designed compounds (red dots) in the pharmaceutically relevant chemical space spanned by 12647 diverse druglike bioactive compounds (COBRA database). Overall compound density is indicated by a color gradient from blue (few data points) to orange (many data points). For visualization all molecules were represented by topological pharmacophoric features (CATS descriptor) and projected onto two dimensions by stochastic neighbor embedding (SNE).

pounds affecting 980 macromolecular targets,[22] together with our 218 de novo designed molecules. For landscape modeling we represented all molecules by a topological pharmacophore descriptor (CATS)[23] and projected the resulting 210-dimensional real-valued fingerprints onto the plane by stochastic neighbor embedding.[24] We employed this nonlinear projection method to preserve and focus on the local neighborhood behavior[25] and reduce the risk of projecting artifacts from outliers and globally non-Gaussian distribution of data.[26,27] Apparently, the majority of the designed compounds are located in an area of chemical space that is adjacent to known kinase inhibitors. This finding suggests that these two sets of compounds possess partially overlapping but not identical pharmacophoric features and the designed molecular entities were generated beyond the structural diversity of known kinase inhibitors. The location of compound **4** at the border between known and designed molecules further prompted us to select it as a preferable candidate for detailed investigation.

We also docked compound **4** and fluoxetine into the active site of hPlk1 to obtain a preliminary structural model of the enzyme–ligand complex. Automated flexible docking into the ATP binding cavity of active hPlk1 (Protein Data Bank,[28] PDB ID: 2ou7)[29] and a model of inactive hPlk1, which we had generated from an inactive DFG-out conformation of Aurora A kinase (PDB ID: 2c6e),[14,30] suggests the preferred binding of compound **4** to the inactive enzyme (GOLD software, ASP score[31,32] = 40 vs. 47; greater positive values indicate potentially stronger binding). Fluoxetine docking yielded overall lower scores for the predicted ligand–protein complexes (active hPlk1: 29; inactive hPlk1: 37) but exhibited the same trend. These modeling results suggested that both compound **4** and fluoxetine might actually bind and prefer inactive over active hPlk1 conformations.

The designed compound was readily amenable to chemical synthesis; we followed the two-step synthetic route suggested by the software without further optimization (Figure 1 A). The racemic product **4** was obtained after purification, through sequential reductive amination of 3-amino-1-phenylpropan-1-ol (**2**) yielding the intermediate **3**, which underwent Mitsunobu reaction to form the ether linkage.[33] Hereby we validate the potential of the reaction-based generation of compounds for medicinal chemistry.

We initially analyzed compounds **1**, **4**, and fluoxetine for their inhibitory activity against 48 active kinases. At a concentration of 10 μM none of the compounds significantly interfered with active hPlk1 or any of the other human

kinases tested (not shown). It is of special note that the de novo designed compound **4** did not block the activity of Aurora kinase A (STK6), which interplays with hPlk1 effects in vivo.[34]

Subsequently, we performed kinase assays using synchronized HeLa cells to analyze whether compound **4** inhibits hPlk1 kinase in its inactive form. Immunoprecipitated hPlk1 was subjected to kinase assays with increasing concentrations of compound **4** and with casein as the substrate. We observed significantly reduced hPlk1 activity to levels of 40–60% for all concentrations of compound **4** (Figure 2A, 1 nM: 55%, $p = 0.0031$ (two-way ANOVA with Bonferroni correction) 10 nM: 60%, $p = 0.0006$; 100 nM: 46%, $p = 0.007$; 1 μM: 54%, $p = 0.001$; 33 μM: 41%, $p = 0.01$). We also determined hPlk1 activity after incubation of immunoprecipitated hPlk1 with fluoxetine, which exhibited significant inhibition albeit only at 100-fold higher concentrations of 10 and 100 μM (Figure 2B). At a concentration of 10 nM, fluoxetine reduced hPlk1 kinase activity to 80% ($p = 0.08$), 100 nM to 90% ($p = 0.38$), 1 μM to 71% ($p = 0.09$), 10 μM to 61% ($p = 0.03$), and 100 μM to 37% ($p = 0.0003$). These findings indicate that fluoxetine is indeed a moderate inhibitor of inactive hPlk1, and this might actually account for some of the reported antiproliferative effects of fluoxetine on cancer cells.[35,36]

As an additional test for target selectivity, we performed an inhibition assay with compound **4** and inactive Aurora A kinase. No reduction of kinase activity was observed up to a ligand concentration of 33 μM (Figure 2C). Thus, because the obtained novel kinase inhibitor exhibits nanomolar affinity to the inactive state of hPlk1 and favorable estimated ligand efficiency[37] (ΔG/number of heavy atoms ≈ 0.66), it represents a promising starting point for optimization.

To study the effect of compound **4** and fluoxetine on cell cycle progression, we performed FACScan analysis with HeLa cells after double-thymidine block in the presence of either one of these compounds. When we monitored the cell cycle over 48 h we observed strongly delayed progression throughout all phases (Figure S1A in the Supporting Information). Four hours after release, one-third of the control cells were in the G0/G1 phase, one-third in the S phase, and one-third in the G2/M phase, while cells treated with 10 μM compound **4** were still mainly in the S phase; when treated with 33 μM compound **4** almost all cells were locked in the S phase (Figure S1A, first panel). After six hours, the control cells reached the G2/M phase, but cells treated with 33 μM compound **4** were still in the S phase (Figure S1A, second panel). After eight hours, the cells treated with 10 μM compound **4** displayed approximately the same cell cycle distribution as the control cells, while cells treated with 33 μM compound **4** were still delayed in the S phase (Figure S1A, third panel). Ten to fourteen hours after release, the control cells and the cells treated with 10 μM compound **4** all reached normal cell cycle distribution (Figure S1A, fourth to sixth panel), and cells that had been treated with 33 μM compound **4** completed S phase and arrived in G2/M phase. Within 24–48 h after release, the control cells had completed a normal cell cycle, but cells treated with compound **4** progressed from G2/M arrest (after 14 h) to apoptosis (Figure S1A, seventh to bottom panel). Treatment with fluoxetine led to comparable

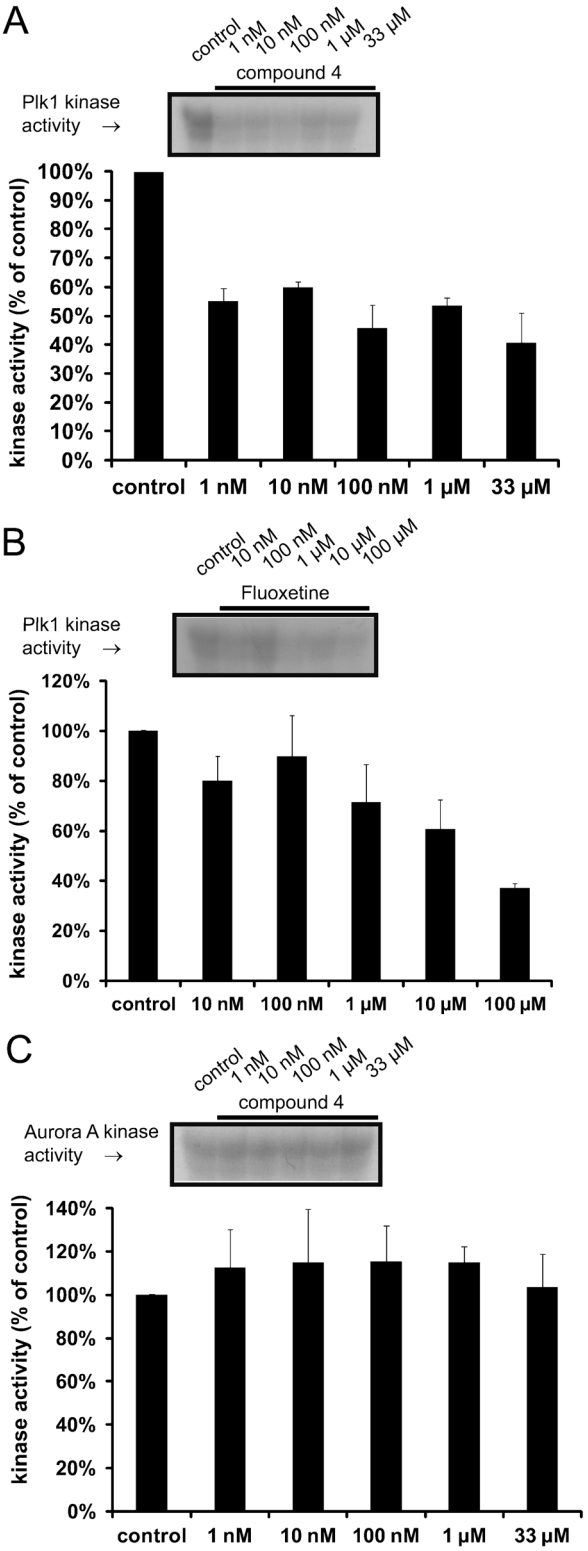

Figure 2. Inhibition of inactive hPlk1 and Aurora A kinase. Kinase assays using immunoprecipitated hPlk1 from HeLa cells after incubation with compound **4** (A) and fluoxetine (B), or compound **4** together with immunoprecipitated human Aurora A kinase from HeLa cells (C). The autoradiograms show representative assays, and the bar graphs show means of three independent experiments measuring hPlk1 or Aurora A kinase activity after immunoprecipitation and incubation with the two compounds, respectively, followed by subsequent kinase assay using casein as a substrate.

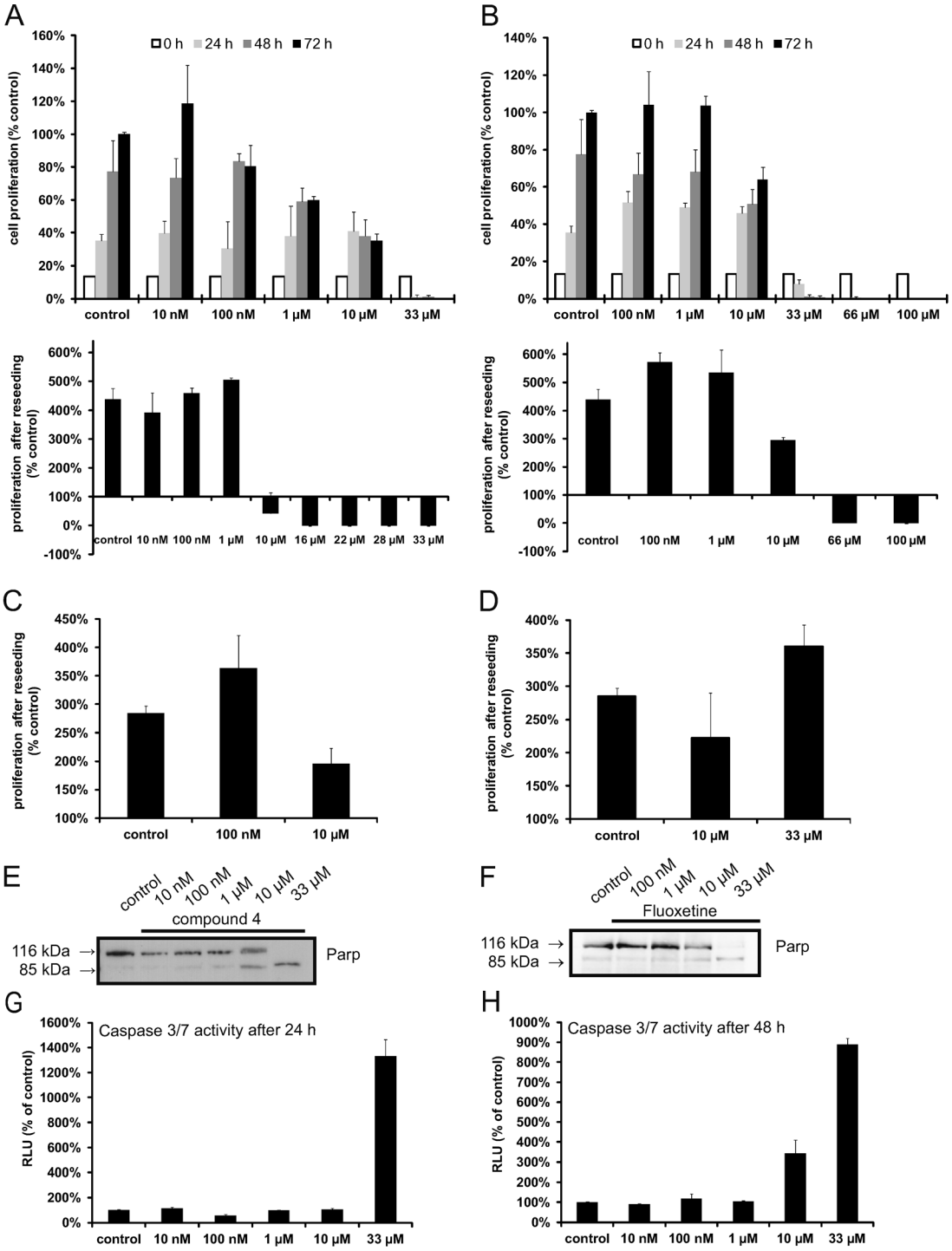

Figure 3. A–D) Cell proliferation and viability. Effect of compound **4** (A) or fluoxetine (B) on cell proliferation and wash-out grow-out assay in HeLa cells, and effect of compound **4** and fluoxetine in wash-out grow-out assays in hTERT-RPE1 cells (C, D). Cell proliferation of HeLa cells was analysed 24 to 72 h after treatment. Control cells were incubated with culture medium alone. Percentage of surviving cells is given as percentage of the number of control cells 72 h after incubation with the drugs. All experiments: $n=3$. Additionally, HeLa cells (A, B) and hTERT-RPE1 cells (C, D) were analysed using wash-out grow-out assays after treatment with compound **4** (A, C) or fluoxetine (B, D). Cells were treated with respective compound for 72 h, harvested and one fifth (HeLa cells) or one third (hTERT-RPE1 cells), respectively, was re-seeded in new six-well-plates with fresh medium without drugs. (E–H) Induction of cellular apoptosis. Western blot analysis of Parp cleavage after treatment of HeLa cells with compound **4** (E) or fluoxetine (F). To determine the full-length Parp protein and the cleavage product in apoptotic cells, Western blot analyses targeting Parp were performed in HeLa cells 48 h after incubation. Caspase 3/7 activation was monitored 24 h (G) and 48 h (H) after treatment with compound **4**. Luminescence is given as relative RLU levels ($n=3$, mean $\pm\sigma$). Control cells were incubated with normal culture medium.

effects as compound **4** concerning cell cycle progression of HeLa cells over the period of 48 h after release from the double thymidine block (Figure S1 B). Cells showed delayed progression in cell cycle and, as observed for compound **4**, they exited from G2/M phase into apoptosis at later points in time.

To determine the impact of compound **4** and fluoxetine on proliferation we counted cells in the period 24–72 h after treatment. Compound **4** induced concentration- and time-dependent reduction of HeLa cell proliferation with an EC_{50} value of 4 ± 1 μM (Figure 3 A, upper panel). After 72 h the reduction was significant for concentrations of 100 nM and higher (100 nM: 81 %, $p = 0.03$; 1 μM: 60 %, $p = 0.02$; 10 μM: 35 %, $p = 0.01$; 33 μM: 0 %, $p = 0.002$). To analyze whether the cells were able to recover from the impairment induced by the inhibitors, we determined cell proliferation a further 72 h after reseeding the cells in new six-well-plates with fresh culture medium without drugs (wash-out grow-out assay). Cells were counted after the first 72 h period, then reseeded and cell numbers were determined after another 72 h period. Incubation with fresh medium without compound **4** led to recovery for the cells that had been first treated with lower concentrations (10 nM–1 μM), and the recovery was comparable to that of the untreated control (Figure 3 A, lower panel). The increase in cell numbers was approximately 400–500 % compared to the number of the reseeded cells in the 72 h period. In contrast, higher concentrations resulted in a reduction of recovery for the cells that had been treated with 10 μM of compound **4** (42 %). The cells treated first with higher concentrations (16 μM and 22 μM) did not recover; after the additional 72 h period with fresh medium without drug the relative cell proliferation was 0 %. Those cells treated with 28 and 33 μM compound **4,** respectively, were completely dead at the time of reseeding, thus leading to 0 % cells after the additional 72 h with fresh medium without drug.

Compared to compound **4**, fluoxetine induced a weaker reduction of HeLa cell proliferation over the period of 24 to 72 h, but it was still significant for concentrations exceeding 10 μM (Figure 3 B, upper panel). The reduction of cell proliferation was concentration- and time-dependent with an EC_{50} value of 14 ± 2 μM [10 μM: reduction to 64 % ($p = 0.02$); 33 μM: 1 % ($p = 0.0001$); 66 μM: 0 % ($p < 0.0001$); and 100 μM: 0 % ($p < 0.0001$)]. Wash-out grow-out assays gave comparable results, with recovery after fluoxetine treatment similar to those of control cells for concentrations up to 1 μM (430–570 % compared to the reseeded cell number), but higher concentrations led to impaired recovery (10 μM: 295 %) (Figure 3 B, lower panel). After treatment with 66 or 100 μM fluoxetine, the cells were completely dead within the first 72 h, leading to 0 % recovery after an additional 72 h with fresh medium without drug, as observed for compound **4**.

We also investigated the impact of compound **4** and fluoxetine on the recovery of hTERT-RPE1 cells (Figure 3 C,D) to determine their capacity as potential anticancer drugs. For that reason, hTERT-RPE1 cells were treated with 100 nM and 10 μM compound **4**, and with 10 and 33 μM fluoxetine, respectively. As described for HeLa cells, hTERT-RPE1 cells were counted after the first 72 h period, then reseeded and the cell numbers were determined after another 72 h period. Both compounds induced slightly impaired recovery, yet all cells were viable and able to proliferate reseeding into fresh medium.

To analyze the induction of apoptosis, we first performed Western blot analyses against Parp in HeLa cells. We were able to detect an increasing amount of cleavage products of 85 kDa, accompanied by a decrease of the full-length protein of 116 kDa, after treatment with compound **4**, starting at concentrations of 100 nM up to 33 μM, where completely degraded full-length protein was observed (Figure 3 E). After incubation with fluoxetine the full-length protein was completely cleaved first at a drug concentration of 33 μM, but at lower drug concentrations we observed only a slight increase of cleaved Parp fragment compared to control cells (Figure 3 F). To confirm the induction of apoptosis after treatment with compound **4** we performed Caspase 3/7 assays after incubation times of 24 and 48 h (Figure 3 G,H); we found concentration- and time-dependent Caspase 3/7 activation indicative of apoptosis.

The most relevant outcome of this study: de novo designed compound **4** proved to be a synthetically feasible, druglike hPlk1 inhibitor with high in vitro potency and selectivity, antiproliferative activity against cancer cells, and with minimal effect on immortalized nontransformed cells. The results confirm our concept of reaction-driven, template-based de novo design as a premier methodology for the rapid identification of novel bioactive molecules exhibiting a desired biological activity spectrum. This approach may well provide an opportunity to jumpstart stalled drug-discovery projects.

Received: August 25, 2012
Published online: November 20, 2012

Keywords: drug design · enzyme inhibitors · fragment-based design · Polo-like kinase · virtual synthesis

[1] M. E. Bunnage, *Nat. Chem. Biol.* **2011**, *7*, 335 – 339.
[2] K. I. Kaitin, J. A. DiMasi, *Clin. Pharmacol. Ther.* **2011**, *89*, 183 – 188.
[3] F. Pammolli, L. Magazzini, M. Riccaboni, *Nat. Rev. Drug Discovery* **2011**, *10*, 428 – 438.
[4] Y. L. Bennani, *Drug Discovery Today* **2011**, *16*, 779 – 792.
[5] G. Schneider, U. Fechner, *Nat. Rev. Drug Discovery* **2005**, *4*, 649 – 663.
[6] M. Hartenfeller, G. Schneider, *WIREs Comp. Mol. Sci.* **2011**, *1*, 742 – 759.
[7] G. Schneider, *J. Comput. Aided Mol. Des.* **2012**, *26*, 115 – 120.
[8] C. McInnes, M. D. Wyatt, *Drug Discovery Today* **2011**, *16*, 619 – 625.
[9] K. Strebhardt, *Nat. Rev. Drug Discovery* **2010**, *9*, 643 – 660.
[10] Y. Liu, N. S. Gray, *Nat. Chem. Biol.* **2006**, *2*, 358 – 364.
[11] I. Akritopoulou-Zanze, P. J. Hajduk, *Drug Discovery Today* **2009**, *14*, 291 – 297.
[12] S. Keppner, E. Proschak, G. Schneider, B. Spänkuch, *ChemMedChem* **2009**, *4*, 1806 – 1809.
[13] M. Hartenfeller, H. Zettl, M. Walter, M. Rupp, F. Reisen, E. Proschak, S. Weggen, H. Stark, G. Schneider, *PLoS Comput. Biol.* **2012**, *8*, e1002380.
[14] G. Schneider, T. Geppert, M. Hartenfeller, F. Reisen, A. Klenner, M. Reutlinger, V. Hähnke, J. A. Hiss, H. Zettl, S.

Keppner, B. Spänkuch, P. Schneider, *Future Med. Chem.* **2011**, *3*, 415 – 424.
[15] M. Hartenfeller, G. Schneider, *Methods Mol. Biol.* **2011**, *672*, 299 – 323.
[16] M. Hartenfeller, M. Eberle, P. Meier, C. Nieto-Oberhuber, K.-H. Altmann, G. Schneider, E. Jacoby, S. Renner, *J. Chem. Inf. Model.* **2012**, *52*, 1167 – 1178.
[17] G. Schneider, M. Hartenfeller, M. Reutlinger, Y. Tanrikulu, E. Proschak, P. Schneider, *Trends Biotechnol.* **2009**, *27*, 18 – 26.
[18] D. T. Wong, J. S. Horng, F. P. Bymaster, K. L. Hauser, B. B. Molloy, *Life Sci.* **1974**, *15*, 471 – 479.
[19] D. T. Wong, F. P. Bymaster, E. A. Engleman, *Life Sci.* **1995**, *57*, 411 – 441.
[20] G. R. Bickerton, G. V. Paolini, J. Besnard, S. Muresan, A. L. Hopkins, *Nat. Chem.* **2012**, *4*, 90 – 98.
[21] P. Leeson, *Nature* **2012**, *481*, 455 – 456.
[22] P. Schneider, G. Schneider, *QSAR Comb. Sci.* **2003**, *22*, 713 – 718.
[23] G. Schneider, W. Neidhart, T. Giller, G. Schmid, *Angew. Chem.* **1999**, *111*, 3068 – 3070; *Angew. Chem. Int. Ed.* **1999**, *38*, 2894 – 2896.
[24] G. Hinton, S. Roweis, *Adv. Neural Inf. Proc. Syst. (NIPS)*, **2002**, *15*, 857 – 864.
[25] D. Horvath, C. Jeandenans, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 680 – 690.
[26] M. Reutlinger, G. Schneider, *J. Mol. Graphics Modell.* **2012**, *34*, 108 – 117.
[27] M. Rupp, P. Schneider, G. Schneider, *J. Comput. Chem.* **2009**, *30*, 2285 – 2296.
[28] 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*, 235 – 242.
[29] M. Kothe, D. Kohls, S. Low, R. Coli, A. C. Cheng, S. L. Jacques, T. L. Johnson, C. Lewis, C. Loh, J. Nonomiya, A. L. Sheils, K. A. Verdries, T. A. Wynn, C. Kuhn, Y. H. Ding, *Biochemistry* **2007**, *46*, 5960 – 5971.
[30] N. M. Heron, *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1320 – 1323.
[31] G. Jones, P. Willett, R. C. Glen, A. R. Leach, R. Taylor, *J. Mol. Biol.* **1997**, *267*, 727 – 748.
[32] G. Jones, P. Willett, R. C. Glen, *J. Mol. Biol.* **1995**, *245*, 43 – 53.
[33] T. Y. But, P. H. Toy, *Chem. Asian J.* **2007**, *2*, 1340 – 1355.
[34] S. M. Lens, E. E. Voest, R. H. Medema, *Nat. Rev. Cancer* **2010**, *10*, 825 – 841.
[35] A. Krishnan, R. Hariharan, S. A. Nair, M. R. Pillai, *Biochem. Pharmacol.* **2008**, *75*, 1924 – 1934.
[36] A. Stepulak, W. Rzeski, M. Sifringer, K. Brocke, A. Gratopp, K. Kupisz, L. Turski, C. Ikonomidou, *Cancer Biol. Ther.* **2008**, *7*, 1685 – 1693.
[37] A. L. Hopkins, C. R. Groom, A. Alex, *Drug Discovery Today* **2004**, *9*, 430 – 431.

