VIP Drug Discovery Very Important Paper

International Edition: DOI: 10.1002/anie.201508055
German Edition: DOI: 10.1002/ange.201508055

De Novo Fragment Design for Drug Discovery and Chemical Biology

*Tiago Rodrigues, Daniel Reker, Martin Welin, Michael Caldera, Cyrill Brunner, Gisela Gabernet, Petra Schneider, Björn Walse, and Gisbert Schneider**

Abstract: *Automated molecular de novo design led to the discovery of an innovative inhibitor of death-associated protein kinase 3 (DAPK3). An unprecedented crystal structure of the inactive DAPK3 homodimer shows the fragment-like hit bound to the ATP pocket. Target prediction software based on machine learning models correctly identified additional macromolecular targets of the computationally designed compound and the structurally related marketed drug azosemide. The study validates computational de novo design as a prime method for generating chemical probes and starting points for drug discovery.*

Computer-based de novo design of new chemical entities (NCEs) is an emerging concept for generating small, synthetically tractable molecules with desired physicochemical properties and pharmacological profiles.[1] At the same time, fragment-based drug discovery represents a key technology for hit identification and ligand optimization.[2] Considering the sheer size of drug-like chemical space,[3] computational means assist in the exploration of fragment-like chemical matter and the identification of starting points for rational ligand design.[4] Aside from prototyping effective therapeutics, the design of innovative chemical fragments will also improve our understanding of protein structure and function. Herein, we present the concept of rapid fragment prototyping by ligand-based de novo design and macromolecular-target prediction. Automated software enabled the identification of a fragment-like NCE that inhibits human death-associated protein kinase 3 (DAPK3), which was computationally predicted and experimentally confirmed by the first X-ray crystal structure of the inactive DAPK3 homodimer in complex with the designed ligand. Furthermore, we identified DAPK3 and carbonic anhydrase IX as hitherto unknown targets of the anti-hypertensive drug azosemide. The chemical structure of azosemide represents a grown version of the de novo designed DAPK3 fragment-like inhibitor. The results of this study demonstrate that automated molecular design in combination with computational target prediction and structure-based validation allows for rapid NCE prototyping and bears exceptional potential for future drug discovery and chemical biology. In this pioneering setup, we successfully completed the full de novo design cycle, encompassing computational ligand design and target prediction, chemical synthesis, biochemical testing, and biophysical determination of the ligand–target complex.

DAPK3 is a serine/threonine kinase involved in a multitude of (patho)physiological processes.[5] For example, the selective control of myosin phosphorylation exerted by DAPK3 renders this enzyme a target for innovative therapeutics to treat hypertension and smooth muscle disorders.[5a] However, structural information to understand the kinase dynamics and drive the rational structure-based design of efficient NCEs is scarce. To date, only three high-resolution (≤ 2 Å) DAPK3 crystal structures have been reported.[6] Previous data indicate that phosphorylation of Ser50 regulates the dimerization of the DAPK3 catalytic domain.[6b] Our findings suggest that a druggable inactive DAPK3 homodimer may exist.

We employed the software DOGS (Design of Genuine Structures) for ligand-based de novo design. It comprises a set of 25144 commercially available building blocks for rapid chemical-space exploration by performing virtual organic syntheses of NCEs with up to 58 reaction schemes.[7] The software generates virtual chemical entities that mimic a given template, usually a known drug, with regard to size and pharmacophore features. This scaffold-hopping method has already led to the discovery of numerous NCEs with the desired biochemical and pharmacological effects, but has never been employed for the de novo design of bioisosteric fragments. To probe our design hypothesis, we selected the marketed fragment-like drug fasudil (**1**, Figure 1a) as the design template for DOGS.

Although fasudil is a potent Rho-associated protein kinase 2 (ROCK2) inhibitor, it also possesses ligand efficiency (LE)[8] for DAPK3 ($K_i = 1.2$ μM, LE = 0.41),[9] thus constituting an ideal reference point for the exploration of DAPK3 binding fragments. DOGS suggested a total of 521 potential DAPK3 inhibitors, from which 357 were fragment-like chemical entities (molar weight $< 300\ g\,mol^{-1}$) with structurally diverse, mostly bicyclic cores (Scheme 1). The designed compounds were flagged for undesirable groups that may lead to assay interference according to the PAINS,[10] Rishton,[11] and Hann[12] filters.

Fragment prioritization was achieved with an in-house target prediction model (see the Supporting Information). This quantitative machine learning model relies on target affinity data for small molecules annotated in the ChEMBL v16 database.[13] Compounds were characterized in terms of their physicochemical properties and Morgan substructure fingerprints (*radius* = 4, 2048 bit), yielding 2224 dimensional

[*] Dr. T. Rodrigues,[+] D. Reker,[+] M. Caldera, C. Brunner, G. Gabernet, Dr. P. Schneider, Prof. Dr. G. Schneider
Swiss Federal Institute of Technology (ETH Zürich)
Vladimir-Prelog-Weg 4, 8093 Zürich (Switzerland)
E-mail: gisbert.schneider@pharma.ethz.ch
Dr. M. Welin, Prof. Dr. B. Walse
SARomics Biostructures AB, Medicon Village
22381 Lund (Sweden)
[+] These authors contributed equally to this work.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201508055.

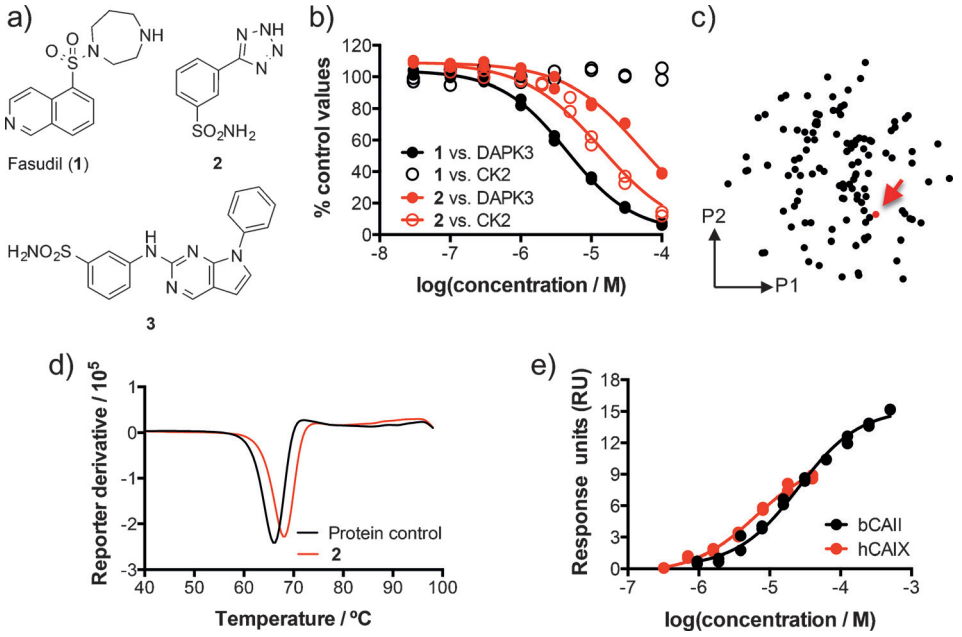
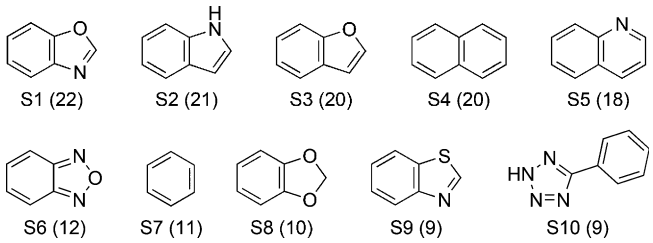


Figure 1. Biochemical and biophysical profiling of fasudil and compound **2**. a) Chemical structures of fasudil, de novo designed compound **2**, and ChEMBL nearest neighbor DAPK3 inhibitor **3**. b) IC_{50} curves of fasudil and **2** against DAPK3 and CK2 ($n=2$). DAPK3: IC_{50} (fasudil) = 4.5 μM ± 0.03 log units, IC_{50} (**2**) = 52 μM ± 0.03 log units; CK2: IC_{50} (fasudil) > 100 μM, IC_{50} (**2**) = 15 μM ± 0.05 log units. Controls: DAPK3: staurosporine, IC_{50} = 0.004 μM (nHill = 1.7); CK2: heparin, IC_{50} = 0.001 μM (nHill = 1.9). c) 2D projection of known fragment-like DAPK3 inhibitors from ChEMBLv16 (K_i or $IC_{50} \leq 100$ μM, black dots) and compound **2** (red dot with arrow) using stochastic neighbor embedding for dimensionality reduction based on CATS2 topological pharmacophore descriptors (new co-ordinates: P1, P2).[14] d) DSF assay of compound **2** (500 μM) against bCAII ($\Delta T_m = 2.40 \pm 0.02$ K, $n=3$). Controls: protein: $T_m = 65.93 \pm 0.02$ °C, acetazolamide (10 μM): $\Delta T_m = 5.20 \pm 0.03$ K. e) SPR assay of **2** against bCAII (K_D = 25 μM ± 0.04 log units, $n=2$) and hCAIX (K_D = 7 μM ± 0.05 log units, $n=2$). Controls: CBS: K_D (bCAII) = 0.95 μM ± 0.06 log units, acetazolamide: K_D (hCAIX) = 0.018 μM ± 0.12 log units.



Scheme 1. Frequently computationally generated scaffolds. Numbers in parentheses are absolute frequencies.

molecular descriptors from which we fitted random forest regression models. Affinity predictions (p*Affinity*) for up to 980 macromolecular drug targets were standard-deviation-corrected (p*Affinity*−σ) to obtain conservative affinity estimates. With the DAPK3 affinity predictions in hand, we selected the de novo designed compound **2** for synthesis and biochemical profiling, given its fragment-like nature and high-ranked predicted affinity (ranked 4th amongst all designs when sorted according to the conservative affinity prediction for DAPK3, p*Affinity*−σ = 4.0). We did not apply any additional metrics or considered a kinase binding hypothesis for compound **2**.

Experimental testing confirmed the affinity of **2** towards DAPK3 with an IC_{50} value of 52 μM ± 0.03 log units (pIC_{50} = 4.3, Figure 1b) and excellent ligand efficiency (LE = 0.40), in full accordance with the regression model. We ruled out unspecific binding that could lead to false positive readouts by colloidal aggregation of **2** at concentrations of up to 500 μM. Cell-based assays revealed the negligible cytotoxicity of **2** at 100 μM (see the Supporting Information). To the best of our knowledge, **2** had not been studied thus far (SciFinder, American Chemical Society). The design template fasudil was screened in parallel against DAPK3 and found to have an IC_{50} value of 4.5 μM ± 0.03 log units (Figure 1b). The results corroborate the scaffold-hopping capabilities of the pharmacophore similarity approach of the DOGS algorithm. Fasudil and the designed compound **2** lack pronounced substructure similarities (Tanimoto similarity = 0.16; Morgan fingerprints, *radius* = 2, 2048 bit) but share a common inhibitory function. The structurally closest known DAPK3 inhibitor annotated in ChEMBL also presents negligible similarity (CHEMBL1988163,[9a] **3**, Tanimoto similarity = 0.28; Figure 1a), suggesting that conventional substructure-based similarity searching in combinatorial virtual libraries would have been inappropriate for identifying **2** as a DAPK3 modulator. A two-dimensional projection of the fragment-like DAPK3 inhibitors together with **2** confirms the presence of analogous pharmacophore features in the known inhibitors, despite their substructural dissimilarity (Figure 1c). We screened **2** against a panel of 27 representative serine/threonine kinases, of which only casein kinase 2 (CK2) was equally inhibited (IC_{50} = 15 μM ± 0.05 log units, LE = 0.45; Figure 1b). The satisfactory kinase selectivity of the designed fragment (Gini coefficient = 0.70)[15] is surprising, taking into account that "flat" molecules are generally considered propitious for promiscuous target engagement.[16]

With the aim of unveiling additional macromolecular targets of **2**, we applied our prediction software to all human drug targets for which ChEMBL ligand data allowed model development. Strong binding was predicted for several carbonic anhydrase isoforms, which was not surprising given the presence of the arene-linked sulfonamide group. Therefore, according to the top-ranking target-binding predictions (see the Supporting Information), we profiled **2** against

bovine carbonic anhydrase II (bCAII) as a surrogate for its human counterpart (p*Affinity*−σ = 6.0). Differential scanning fluorimetry (DSF) revealed a bCAII melting temperature shift of 2.40 ± 0.02 K upon co-incubation with 500 μM of **2** (Figure 1d). Using surface plasmon resonance (SPR) as an orthogonal assay, we confirmed protein binding with a K_D value of 25 μM ± 0.04 log units (Figure 1e). We also determined its affinity towards the cancer-relevant[17] human carbonic anhydrase IX (hCAIX, p*Affinity*−σ = 6.5). Binding to hCAIX (Figure 1e) was observed with an affinity in the range of the predicted value (K_D = 6 μM ± 0.05 log units), suggesting that **2** may serve as a starting point for developing hCAIX inhibition based therapeutics or diagnostic tools. Fasudil showed no binding affinity towards hCAIX, which is in agreement with the computational target prediction.

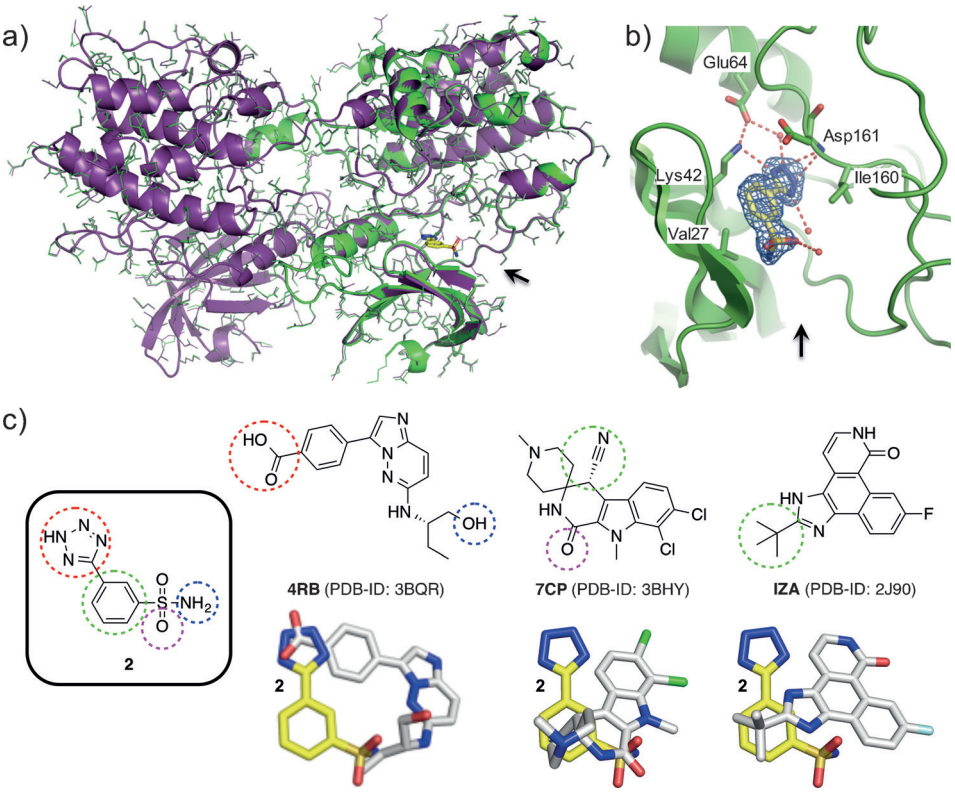


Figure 2. X-ray data and analysis. a) Ligand-free *apo* structure (purple) and ligand-bound complex (green) of DAPK3 and compound **2**. Structures were aligned with the PyMOL (v1.7) "align" function, resulting in a near-perfect superpositioning (RMSD = 0.14 Å) of the *apo* and ligand-bound conformations. Minor structural shifts are visible only in the ligand-binding domain. b) Active site of DAPK3 with bound compound **2** (yellow sticks) and water molecules (red spheres). Interacting side chains (green sticks) and hydrogen bonds (red dotted lines) are highlighted. The mesh outlines the observed electron density of the $2m|F_o| - d|F_c|$ map contoured at 1.0σ around compound **2**. c) Pharmacophore comparison of **2** and three DAPK3 inhibitors (PDB-IDs 2BQR, 3BHY, and 2J90). The colored dotted circles highlight isosteric feature pairs. The structural alignments correspond to the ligand binding poses observed in the superimposed kinase–ligand complexes.

DAPK3 is pro-apoptotic and acts as a tumor suppressor.[5b,18] Blocking inactive DAPK3 would likely abolish the tumor-suppressing activity. However, because of its adequate kinase selectivity, compound **2** is a potentially useful starting point for developing a chemical probe for studying DAPK3-dependent signaling pathways. Importantly, DAPK3 is a novel cardiovascular pharmaceutical target. Blockage of the DAPK3 activity results in vascular structural remodeling, as shown for the prevention of neointimal hyperplasia in vivo.[19]

Having biochemically confirmed that compound **2** inhibits DAPK3, we initiated efforts to obtain a structural model that could rationalize the biochemical data and drive a hit-to-lead program. An X-ray crystal structure of the inactive DAPK3 homodimer in complex with **2** was solved at a resolution of 1.7 Å (see the Supporting Information). Electron density maps calculated from data collected on DAPK3 crystals soaked with compound **2** revealed unambiguous density for the bound fragment in the ATP binding pocket, with very little change in conformation compared to the *apo*-DAPK3 structure (Figure 2a). Compound **2** is inserted in a perpendicular binding mode with respect to the nucleotide-binding site (Figure 2b). The tetrazole ring forms a bidentate hydrogen bridge to the backbone amide of Asp161 (the D of the "DFG" motif) and an additional hydrogen bridge to the side chain of Lys42, which also forms a characteristic salt bridge to Glu64 on the C-helix. The phenyl moiety of **2** is wedged between Val27 situated on the N-lobe and Ile160 on the C-lobe of the kinase. The sulfonamide moiety resides in a solvent-exposed area near the pocket exit. It is unknown whether fasudil actually adopts a similar or rather dissimilar binding pose in DAPK3, considering its different chemical structure and physicochemical properties compared to **2**. There are four crystal structures of DAPK3 deposited in the PDB (IDs: 3BHY[6a] [resolution: 1.24 Å], 3BQR[6b] [1.75 Å], 2J90 [2.0 Å], and 1YRP [3.1 Å]). We superimposed our new ligand-bound structure of DAPK3 with the three high-resolution structures with root-mean-square deviations below 1 Å, which indicates almost identical backbone folds. A comparison of the ligand-binding modes revealed disparate ligand–receptor interaction patterns (Figure 2c). Compound **2** shares individual pharmacophores with all of the three reference inhibitors. The most similar pattern was found with compound 4RB (PDB ID: 3BQR), with two common pharmacophore features but oppositely oriented mid sections. The de novo designed ligand features a tetrazole head group as a bioisosteric replacement of the carboxylic function in 4RB, both interacting with Lys42 of the kinase. To the best of

our knowledge, a tetrazole hinge-binding motif has not been explored in kinase inhibitors to date.

With a structural basis for DAPK3 inhibition by **2** in hand, we recognized that azosemide (**4**, Figure 3) could inhibit DAPK3 in a similar fashion as it contains the former as an exact substructure. Furthermore, azosemide presents a side chain that can explore one of the exit vectors that we had predicted to be beneficial for DAPK3 affinity. Thus azosemide afforded the opportunity to simulate the fragment growing of **2** towards an already pharmacodynamically validated small molecule, which is an approved drug in Japan for the treatment of hypertension. However, despite acting primarily at Henle's loop, to the best of our knowledge, direct molecular targets have remained elusive.[21] Profiling of azosemide against DAPK3 showed concentration-dependent reversible inhibition and pronounced ligand efficiency (IC_{50} = 2.1 μM ± 0.06 log units, LE = 0.35; Figure 3a). Significantly, DAPK3 is the first disclosed direct molecular target for azosemide. Arguably, its inhibition may partly explain the anti-hypertensive effects of **4**. Azosemide also showed binding affinity towards hCAIX (K_D = 1.4 μM ± 0.03 log units; Figure 3b), suggesting that it might be equally useful as a lead structure for hCAIX-based therapeutics or the formulation of targeted drug delivery systems. Azosemide is known to be teratogenic.[22] Considering the long-standing evidence for teratogenicity by carbonic anhydrase inhibitors,[23] the findings of our study provide a hypothesis for the underlying molecular mechanism of these observations.

Both drug discovery and chemical biology benefit from rapid access to NCEs. This necessity can be partially met by the computer-based molecular design of synthetically accessible bioactive agents and reliable macromolecular-target predictions. Herein, we have successfully prototyped such an approach. The results of this study pinpoint the ability of ligand-based de novo design to generate synthesizable, fragment-like starting points that provide innovative chemical structures for "growth" into efficacious leads.

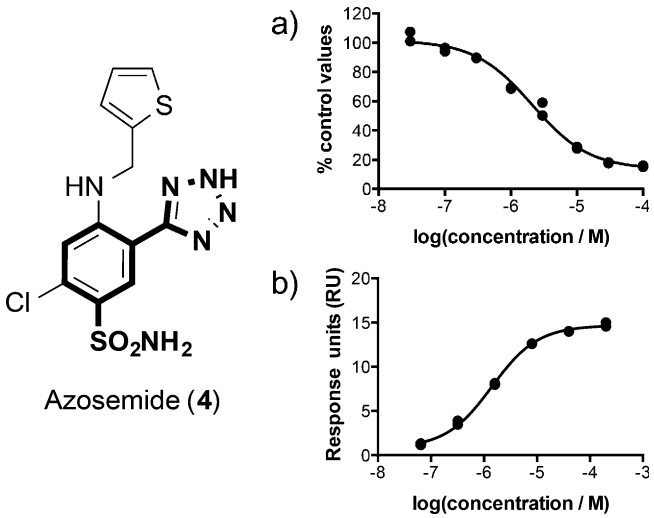


Figure 3. Biochemical and biophysical profiling of azosemide (**4**). The fragment **2** substructure is highlighted. a) IC_{50} curve of azosemide against DAPK3 (n = 2). IC_{50} = 2.1 μM ± 0.06 log units. Control: staurosporine, IC_{50} = 0.015 μM (nHill = 1.3). b) SuPR assay of azosemide binding to hCAIX (n = 2; K_D = 1.4 μM ± 0.03 log units). Control: acetazolamide, K_D = 0.018 μM ± 0.12 log units.

Acknowledgements

The study was financially supported by the ETH Zürich and the OPO Foundation, Zurich, Switzerland. Mirjam Menzi and Moritz Stoltz are thanked for technical support, and Dr. Jörg Scheuermann and Dr. Moreno Wichert for providing access to human carbonic anhydrase IX. P.S. and G.S. are the founders of inSili.com LLC, Zurich, Switzerland.

Keywords: chemical biology · computational chemistry · kinases · molecular design · target prediction

How to cite: *Angew. Chem. Int. Ed.* **2015**, *54*, 15079–15083
Angew. Chem. **2015**, *127*, 15294–15298

[1] a) G. Schneider, U. Fechner, *Nat. Rev. Drug Discovery* **2005**, *4*, 649–663; b) J. Besnard, G. F. Ruda, V. Setola, K. Abecassis, R. M. Rodriguiz, X. P. Huang, S. Norval, M. F. Sassano, A. I. Shin, L. A. Webster, F. R. Simeons, L. Stojanovski, A. Prat, N. G. Seidah, D. B. Constam, G. R. Bickerton, K. D. Read, W. C. Wetsel, I. H. Gilbert, B. L. Roth, A. L. Hopkins, *Nature* **2012**, *492*, 215–220; c) B. Spänkuch, S. Keppner, L. Lange, T. Rodrigues, H. Zettl, C. P. Koch, M. Reutlinger, M. Hartenfeller, P. Schneider, G. Schneider, *Angew. Chem. Int. Ed.* **2013**, *52*, 4676–4681; *Angew. Chem.* **2013**, *125*, 4774–4779; d) T. Rodrigues, T. Kudoh, F. Roudnicky, Y. F. Lim, Y. C. Lin, C. P. Koch, M. Seno, M. Detmar, G. Schneider, *Angew. Chem. Int. Ed.* **2013**, *52*, 10006–10009; *Angew. Chem.* **2013**, *125*, 10190–10193; e) M. Reutlinger, T. Rodrigues, P. Schneider, G. Schneider, *Angew. Chem. Int. Ed.* **2014**, *53*, 4244–4248; *Angew. Chem.* **2014**, *126*, 4330–4334; f) T. Rodrigues, N. Hauser, D. Reker, M. Reutlinger, T. Wunderlin, J. Hamon, G. Koch, G. Schneider, *Angew. Chem. Int. Ed.* **2015**, *54*, 1551–1555; *Angew. Chem.* **2015**, *127*, 1571–1575.
[2] a) C. W. Murray, D. C. Rees, *Nat. Chem.* **2009**, *1*, 187–192; b) M. Baker, *Nat. Rev. Drug Discovery* **2013**, *12*, 5–7.
[3] R. S. Bohacek, C. McMartin, W. C. Guida, *Med. Res. Rev.* **1996**, *16*, 3–50.
[4] C. Sheng, W. Zhang, *Med. Res. Rev.* **2013**, *33*, 554–598.
[5] a) T. A. Haystead, *Cell. Signalling* **2005**, *17*, 1313–1322; b) J. Brognard, Y. W. Zhang, L. A. Puto, T. Hunter, *Cancer Res.* **2011**, *71*, 3152–3161; c) G. Page, D. Kogel, V. Rangnekar, K. H. Scheidtmann, *Oncogene* **1999**, *18*, 7265–7273.
[6] a) K. Huber, L. Brault, O. Fedorov, C. Gasser, P. Filippakopoulos, A. N. Bullock, D. Fabbro, J. Trappe, J. Schwaller, S. Knapp, F. Bracher, *J. Med. Chem.* **2012**, *55*, 403–413; b) A. C. Pike, P. Rellos, F. H. Niesen, A. Turnbull, A. W. Oliver, S. A. Parker, B. E. Turk, L. H. Pearl, S. Knapp, *EMBO J.* **2008**, *27*, 704–714.
[7] M. Hartenfeller, H. Zettl, M. Walter, M. Rupp, F. Reisen, E. Proschak, S. Weggen, H. Stark, G. Schneider, *PLoS Comput. Biol.* **2012**, *8*, e1002380.
[8] A. L. Hopkins, C. R. Groom, A. Alex, *Drug Discovery Today* **2004**, *9*, 430–431.
[9] a) N. Ono-Saito, I. Niki, H. Hidaka, *Pharmacol. Ther.* **1999**, *82*, 123–131; b) Y. Wang, T. Suzek, J. Zhang, J. Wang, S. He, T. Cheng, B. A. Shoemaker, A. Gindulyte, S. H. Bryant, *Nucleic Acids Res.* **2014**, *42*, D1075–1082.
[10] J. B. Baell, G. A. Holloway, *J. Med. Chem.* **2010**, *53*, 2719–2740.
[11] G. M. Rishton, *Drug Discovery Today* **1997**, *2*, 382–384.

[12] M. Hann, B. Hudson, X. Lewell, R. Lifely, L. Miller, N. Ramsden, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 897 – 902.
[13] A. P. Bento, A. Gaulton, A. Hersey, L. J. Bellis, J. Chambers, M. Davies, F. A. Kruger, Y. Light, L. Mak, S. McGlinchey, M. Nowotka, G. Papadatos, R. Santos, J. P. Overington, *Nucleic Acids Res.* **2014**, *42*, D1083 – D1090.
[14] a) M. Reutlinger, C. P. Koch, D. Reker, N. Todoroff, P. Schneider, T. Rodrigues, G. Schneider, *Mol. Inf.* **2013**, *32*, 133 – 138; b) M. Reutlinger, W. Guba, R. E. Martin, A. I. Alanine, T. Hoffmann, A. Klenner, J. A. Hiss, P. Schneider, G. Schneider, *Angew. Chem. Int. Ed.* **2011**, *50*, 11633 – 11636; *Angew. Chem.* **2011**, *123*, 11837 – 11840.
[15] P. P. Graczyk, *J. Med. Chem.* **2007**, *50*, 5773 – 5779.
[16] a) I. Nobeli, A. D. Favia, J. M. Thornton, *Nat. Biotechnol.* **2009**, *27*, 157 – 167; b) P. A. Clemons, N. E. Bodycombe, H. A. Carrinski, J. A. Wilson, A. F. Shamji, B. K. Wagner, A. N. Koehler, S. L. Schreiber, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 18787 – 18792.
[17] M. Wichert, N. Krall, W. Decurtins, R. M. Franzini, F. Pretto, P. Schneider, D. Neri, J. Scheuermann, *Nat. Chem.* **2015**, *7*, 241 – 249.
[18] T. Kawai, M. Matsumoto, K. Takeda, H. Sanjo, S. Akira, *Mol. Cell. Biol.* **1998**, *18*, 1642 – 1651.
[19] T. Usui, T. Sakatsume, R. Nijima, K. Otani, K. Kazama, T. Morita, S. Kameshima, M. Okada, H. Yamawaki, *Clin. Sci.* **2014**, *127*, 539 – 548.
[20] V. B. Chen, W. B. Arendall III, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr. Sect. D* **2010**, *66*, 12 – 21.
[21] O. K. Suh, S. H. Kim, M. G. Lee, *Biopharm. Drug Dispos.* **2003**, *24*, 275 – 297.
[22] I. Hayasaka, K. Uchiyama, K. Murakami, Z. Kato, F. Tamaki, T. Shibata, T. Sugawara, M. Hayashi, *Congenital Anomalies* **1984**, *24*, 111 – 121.
[23] J. G. Wilson, T. H. Maren, K. Takano, A. Ellison, *Teratology* **1968**, *1*, 51 – 60.

Received: August 27, 2015
Published online: October 21, 2015

