▌**Drug Discovery**

KRAS Binders Hidden in Nature

Andreas Bergner,[a] Xiaoling Cockcroft,[a] Gerhard Fischer,[a] Andreas Gollner,[a] Wolfgang Hela,[a] Roland Kousek,[a] Andreas Mantoulidis,[a] Laetitia J. Martin,[a, b] Moriz Mayer,[a] Barbara Müllauer,[a] Gabriella Siszler,[a] Bernhard Wolkerstorfer,[a] Dirk Kessler,[a] and Darryl B. McConnell*[a]

Abstract: Natural products have proven to be a rich source of molecular architectures for drugs. Here, an integrated approach to natural product screening is proposed, which uncovered eight new natural product scaffolds for KRAS—the most frequently mutated oncogenic driver in human cancers, which has remained thus far undrugged. The approach combines aspects of virtual screening, fragment-based screening, structure-activity relationships (SAR) by NMR, and structure-based drug discovery to overcome the limitations in traditional natural product approaches. By using our approach, a new "snugness of fit" scoring function and the first crystal-soaking system of the active form of KRASG12D, the protein–ligand X-ray structures of a tricyclic indolopyrrole fungal alkaloid and an indoloisoquinolinone have been successfully elucidated. The natural product KRAS hits discovered provide fruitful ground for the optimization of highly potent natural-product-based inhibitors of the active form of oncogenic RAS. This integrated approach for screening natural products also holds promise for other "undruggable" targets.

Natural products have proven to be a rich source of molecular architectures which bind to drug targets.[1] In contrast to synthetic small molecules, these secondary metabolites, which have evolved to protect host organisms, tend to feature more stereogenic centers, contain rigid cores and multiple ring systems with low degrees of conformational entropy (compounds
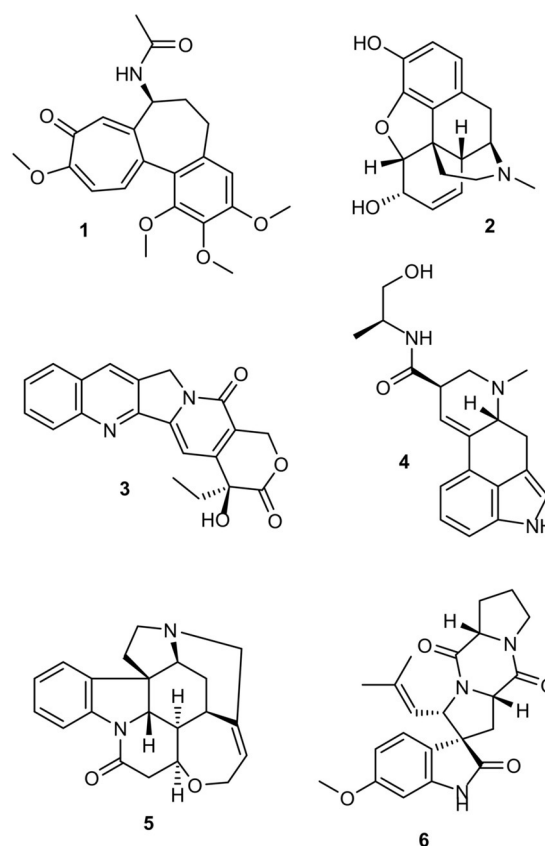
[a] Dr. A. Bergner, Dr. X. Cockcroft, Dr. G. Fischer, Dr. A. Gollner, W. Hela, R. Kousek, Dr. A. Mantoulidis, Dr. L. J. Martin, Dr. M. Mayer, B. Müllauer, G. Siszler, B. Wolkerstorfer, Dr. D. Kessler, Dr. D. B. McConnell
Boehringer Ingelheim RCV GmbH & Co KG
Doktor-Boehringer-Gasse 5-11, 1120 Vienna (Austria)
E-mail: darryl.mcconnell@boehringer-ingelheim.com

[b] Dr. L. J. Martin
Present address: F. Hoffmann-La Roche Ltd.
Grenzacherstrasse 124, 4070 Basel (Switzerland)

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:
https://doi.org/10.1002/chem.201902810.

Figure 1. Secondary metabolites with multiple fused ring architectures including the tricyclic cycloheptatrienone colchicine (**1**), the tetrahydroisoquinoline derived pentacycle morphine (**2**), the pentacyclic quinolone camptothecin (**3**), the tetracyclic indole ergonovine (**4**), the heptacyclic indoline strychnine (**5**) and the spiroindolinone spyrotryprostatin A (**6**).

1–**6**, Figure 1).[2] However, the future role of natural products in drug discovery is constrained due to the limitations in isolation of pure natural products for screening, hit validation and hit characterization.[3] KRAS is the most frequently mutated oncogene and belongs to the protein family of small GTPases that function as binary molecular switches involved in cell signaling.[4] Although KRAS could serve as an excellent target for many cancers, no therapeutic agent directly targeting RAS has been clinically approved. The main reason for this is the lack of druggable pockets on the surface of RAS. In recent years, the existence of two pockets on the surface of KRAS has been discovered, which could potentially be amenable to small-molecule drug discovery.[5] The shallow, polar pocket situated be-

tween the switch I and switch II region of KRAS (SI/II-pocket) is of particular interest because it is present in the active, oncogenic form of KRAS. However, molecular scaffolds capable of binding to the active form of KRAS remain scarce and are of weak affinity.[6] In this paper, we describe an integrated approach to screening natural products in which virtual screening of natural products has been combined with elements of fragment-based screening,[7] the use of SAR by NMR,[8] and structure-based drug discovery[9] to uncover natural-product-based KRAS binders.

Our previous attempts to identify starting points through high-throughput screening of 1.6 million compounds, including a diversity-oriented synthesis library (150 000 compounds), failed to deliver any inhibitors of GTP-KRAS. This demonstrates that the chemical space covered by our historical compound collection was not sufficiently complementary to the pockets on KRAS in the affinity range tested by biochemical and cellular assays. Natural products, however, could represent an area of chemical space containing molecular architectures with sufficient complementarity to KRAS.

Considering that many natural product structures are not ideal for virtual screening due to their large size and high degrees of substitution, a chemoinformatics based PipelinePilot workflow, analogous to that used by other groups,[10] deconstructed all structures in the CRC *Dictionary of Natural Products* (DNP)[11] (180 000 entries at the time of application). This produced a library of 226 000 original and deconstructed natural products amenable to virtual screening, referred to subsequently as the virtual DNP (vDNP). At the time of the virtual screening campaign, only one X-ray co-crystal structure of a ligand bound to the active form of KRAS had been reported.[5b] Due to ambiguity in the electron density for the proposed binding mode in this structure, we decided instead to use the published GDP-KRAS structure (PDB code 4EPV) with indole **7** bound as published by the Fesik group,[5a] as a surrogate for GTP-KRAS.

The vDNP was docked by using Glide into the SI/II-pocket of GDP-KRAS. The compounds with the best 500 molecular-weight-normalized docking scores[12] (Glide score/MW) were visually inspected, and compounds that mimicked the known indole occupation of the SI/II-pocket (Figure 3 A) and formed H-bonds with D54 were selected. Compounds deemed chemically unstable were removed. 3-Substituted indoles were also
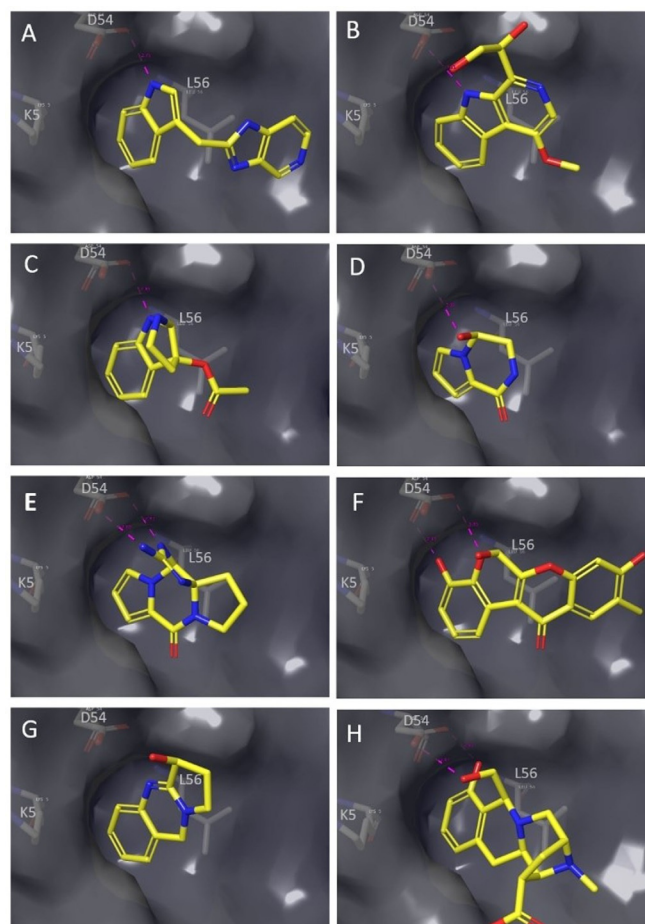

Figure 3. Docking poses of virtual hits in the switch I/II-pocket. (A) **7**, (B) **8**, (C) **9a**, (D) **10b**, (E) **11b**, (F) **12b**, (G) **13**, (H) **14b**.
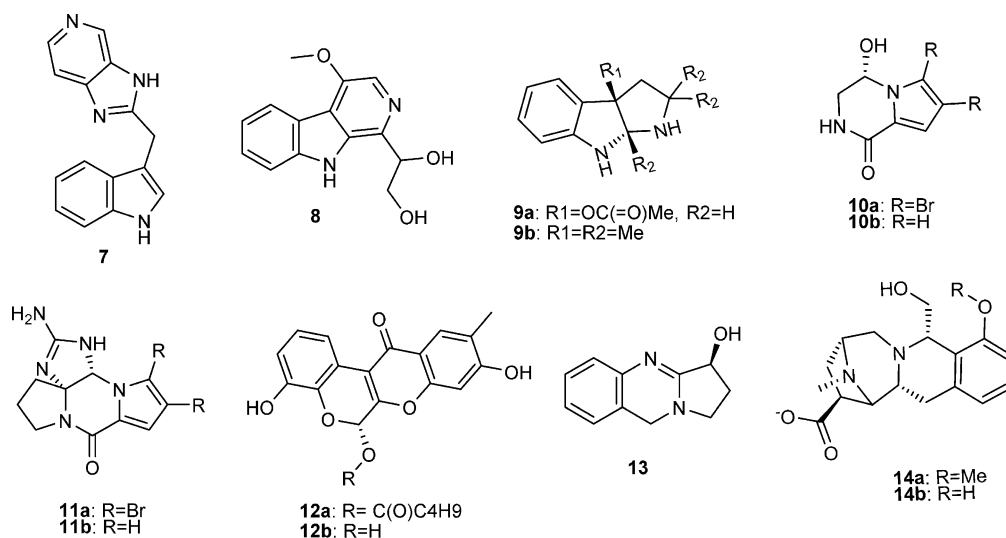

Figure 2. Selected natural product virtual KRAS hits and indole **7** present in the GDP-KRAS co-crystal structure (PDB code 4EPV) used for docking.

not pursued further because they did not provide sufficient structural novelty as a starting point versus **7** (Figure 2).

From the virtual screen of the vDNP, natural product hits isolated from fungi, sponges, plants and bacteria were identified (Figure 2). β-Carbolines and γ-carbolines, which are relatively ubiquitous in nature, substituted at the 1-, 2- and 3-positions were found. β-Carboline **8** (Figure 2), similar to indole **7**, filled the lipophilic SI/II-pocket flanked by L56 and K5 and formed two H-bonds to D54: one through the indole NH and a second through the glycerol substituent (Figure 3B). The tricyclic indolopyrrole alkaloid **9a** (Figure 2) first isolated from the fungus *Fusarium incarnatum* mimics the indole binding mode of **7** and **8**, while forming a second H-bond to D54 through the pyrrole nitrogen (Figure 3C). The marine sponges *Agelas longissima* and *Phakellia flabellate* provided two KRAS virtual hits derived from bromopyrrole alkaloids, namely, the debrominated analogue **10b** of longamide (**10a**) (Figure 2) and the tetracyclic guanidine scaffold Phakellin **11b**, which co-occurs with its dibromo analogue **11a** ((+)-dibromophakellin; Figure 2). Compound **10b** forms one H-bond to D54 through a hydroxyl group (Figure 3d), whereas **11b** forms two H-bonds to D54

through the guanidine group (Figure 3e). Two virtual hits were also uncovered from plant secondary metabolites, namely, diffusarotenoid **12a** and vasicine **13**, a tricyclic quinazoline (Figure 2). The *cis*-fused tetrahydrochromenochromene ring system of the de-esterified diffusarotenoid **12b** (Figure 2) complemented the shape of the SI/II-pocket while forming two H-bonds to D54: one through the phenolic oxygen and the backbone carbonyl, and one between the 6-hydroxyl and the side chain (Figure 3f). Compound **13** also occupies the SI/II-pocket and forms an H-bond with D54 through the 3-hydroxy group (Figure 3g). Deconstruction of the methyl ether of quinocarcinol **14a,** isolated from the bacteria Streptomyces melanovinaceus, led to the virtual hit **14b** (Figure 2) for which the iminoazepinoisoquinoline skeleton occupies the SI/II-pocket and forms two H-bonds to the side chain of D54 with the phenolic and alcoholic oxygens (Figure 3h).

To validate the virtual screen, **9b**, a close analogue of the tricyclic indolopyrrole alkaloid **9a** was tested for binding to GCP-KRAS by using NMR. Compound **9b** demonstrated cross peak shifts in the 2D ^{1}H/^{15}N HSQC NMR spectra of GCP-KRAS^{G12D} (Figure 4a), confirming that these virtual hits do indeed bind to

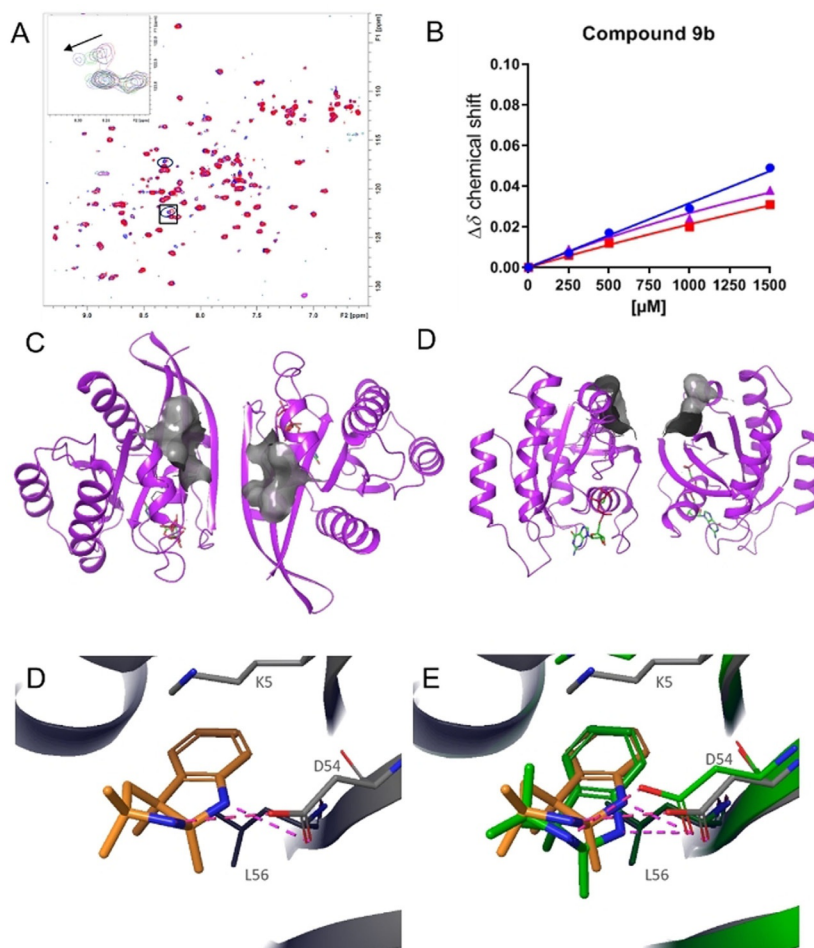**Figure 4.** NMR and X-ray crystal structure of indolopyrrole **9b**. (A) Dose-dependent cross peak shifts in the 2D ^{1}H/^{15}N HSQC NMR spectra of GCP-KRAS^{G12D} on addition of **9b**. (B) NMR K_D titration of **9b** binding to GCP-KRAS^{G12D}. (C) Top view and (D) side view of the high-throughput soaking crystallization system for KRAS^{G12D} with the SI/II-pocket surface depicted in grey. (E) X-Ray co-crystal structure of **9b** binding to the SI/II-pocket of GCP-KRAS^{G12D}. (F) X-Ray co-crystal structure of **9b** binding to the SI/II-pocket of GCP-KRAS^{G12D} in yellow overlaid with the docking pose in green.

the active form of KRAS. The dissociation constant (K_D) for **9b**, as measured by NMR, was in the millimolar range with no saturation obtained with the concentrations tested (up to 1.5 mM, Figure 4b). To be able to perform high throughput crystallization of very weakly binding fragments ($K_D > 1$ mM) with KRAS, such as those reported here, we developed a robust crystal-soaking system for active KRAS[G12D]. By using GDP-KRAS[G12D], we exchanged the nucleotide with the non-hydrolysable GTP analogue GMP-PCP (GCP) to yield a novel crystal form amenable to high-throughput crystallization, which mimics the beta sheet dimeric form of active KRASG12D[13] (Figure 4c). With this soaking system, we were able to determine the co-crystal structure of **9b** with GCP-KRAS at a resolution of 1.2 Å (Figure 4c). The observed binding mode confirmed the docking prediction (Figure 4d), namely binding to the SI/II-pocket and the formation of two H-bonds to D54.

The docking pose of rotenoid **12b** displayed striking complementarity to the shape of the SI/II-pocket (Figure S2A, Supporting Information); however, semiempirical quantum mechanics calculations (AM1, PM3, MNDO) and crystal structure analysis of structurally related compounds in the Cambridge

Structural Database suggest that rotenoids exist in a flat conformation. The complementary "kinked" docking pose of **12b** in the SI/II-pocket of KRAS encouraged us to generate a sub-library of 1383 rotenoid-like "kinked" molecules with contiguous fused rings from our corporate database for a second virtual screening campaign.

To ensure that the shape complementarity between the rigid three-dimensional topology of our "kinked" library and the KRAS protein were assessed accurately, we developed a "snugness-of-fit" scoring function. The shapes of a protein and a ligand were calculated as envelopes around their respective atoms and were represented as 3D grids. Envelope overlaps that occurred at small atomic radii corresponded to those areas where the protein and ligand were close, that is, "snug", whereas overlaps that only occurred at larger radii represented fewer snug areas. Addition of all envelope overlap grids (at different radii) provided a quantitative measure for the "snugness of fit" of the ligand at each point in space. The grids can be also used for visualizing the snugness of fit of a docked or crystallized compound using iso-contour surfaces or color-coded surfaces (Figure S2, Supporting Information).
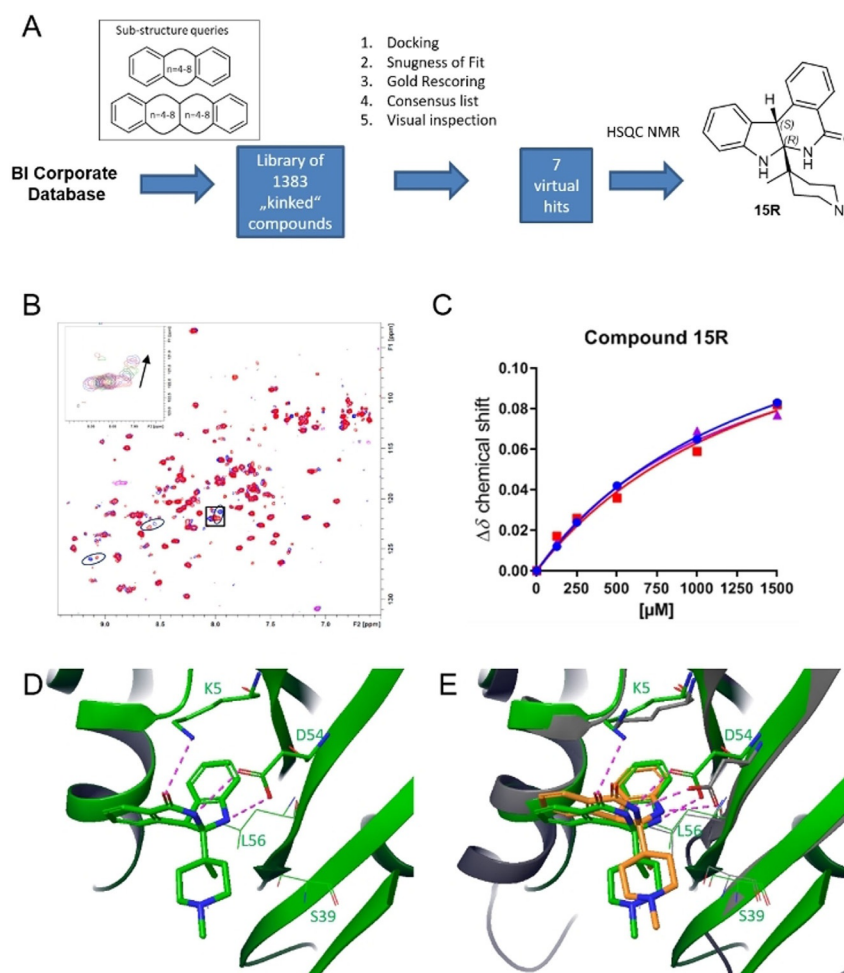

Figure 5. "Kinked" compound library virtual screen and crystal structure of indoloisoquinolinone **15**. (A) Schematic representation of the generation of the "kinked" compound library and virtual screen leading to the hit **15**. (B) Dose dependent cross peak shifts in the 2D ^{1}H/^{15}N HSQC NMR spectra of GCP-KRAS[G12D] on addition of **15R**. (C) NMR K_D titration of **15R** binding to GCP-KRAS[G12D] showing a K_D of 1 mM. (D) X-Ray co-crystal structure of **15R** binding to the SI/II-pocket of GCP-KRAS[G12D]. (E) X-Ray co-crystal structure of **15R** binding to the SI/II-pocket of GCP-KRAS[G12D] in green overlaid with the docking pose in orange.

The "kinked" library was virtually screened in an analogous procedure to that used for the vDNP, with the addition of the "snugness of fit" scoring function for pose post-processing (Figure 5 a). After visual inspection, 7 compounds were selected for NMR testing and the racemic compound **15** showed NMR cross peak shifts. Both enantiomers **15 R** and **15 S** docked with conserved indole binding modes (Figure S3 A,B, Supporting Information); hence, the racemate **15** was separated into its two enantiomers **15 R** (2*R*,3*S*) and **15 S** (2*S*,3*R*). **15 R** displayed a K_D of 1 mм (Figure 5 B,C), as determined by the 2D $^1H/^{15}N$ HSQC NMR spectra to GCP-KRAS^{G12D}, whereas **15 S** bound more weakly with a dissociation constant in the millimolar range (Figure S3 C and Figure S3 D). Enantiomer **15 R** was successfully soaked into the crystallography system (Figure 5 D) with the crystal structure confirming the docking pose (Figure 5 E). **15 R** binds to the SI/II-pocket of active KRAS and forms two chelating H-bonds to D54 with an N—O distance of 2.9 and 3.0 Å, respectively, and a weaker third H-bond to the K9 side chain through the carbonyl group with an O—N distance of 3.4 Å.

In conclusion, from the two virtual screens conducted, virtual hits stemming from seven classes of natural products from fungi, sponges, plants, bacteria and one natural-product-like class from our corporate compound collection were identified. Of these, we confirmed biophysical binding to the active form of KRAS^{G12D} for the tricyclic indolopyrrole alkaloid **9 b** and the indoloisoquinolinones **15 R** and **15 S**. Importantly, we elucidated the binding mode by determining the protein–ligand complex structures of **9 b** and **15 R** with the first X-ray crystal-soaking system for the active form of KRAS^{G12D}. The wealth of natural product KRAS hits, discovered by using this integrated approach of natural product screening, should act as inspiration for the design of more potent and selective RAS inhibitors. This approach also promises to be applicable to other "undruggable" targets beyond KRAS, for which classical screening methods have proven to be unsuccessful.

Acknowledgements

Conflict of interest

The authors declare no conflict of interest.

[1] J. Clardy, C. Walsh, *Nature* **2004**, *432*, 829–837.
[2] a) K. Grabowski, G. Schneider, *Curr. Chem. Biol.* **2007**, *1*, 115–127; b) T. Henkel, R. M. Brunne, H. Müller, F. Reichel, *Angew. Chem. Int. Ed.* **1999**, *38*, 643–647; *Angew. Chem.* **1999**, *111*, 688–691; c) M. Feher, J. M. Schmidt, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 218–227; d) P. Ertl, S. Roggo, A. Schuffenhauer, *J. Chem. Inf. Model.* **2008**, *48*, 68–74.
[3] a) D. J. Newman, G. M. Cragg, K. M. Snader, *Nat. Prod. Rep.* **2000**, *17*, 215–234; b) A. L. Harvey, *Drug Discovery Today* **2008**, *13*, 894–901; c) J. W.-H. Li, J. C. Vederas, *Science* **2009**, *325*, 161–165.
[4] J. Colicelli, *Sci. STKE* **2004**, *250*, re13.
[5] a) Q. Sun, J. P. Burke, J. Phan, M. C. Burns, E. T. Olejniczak, A. G. Waterson, T. Lee, O. W. Rossanese, S. W. Fesik, *Angew. Chem. Int. Ed.* **2012**, *51*, 6140–6143; *Angew. Chem.* **2012**, *124*, 6244–6247; b) T. Maurer, L. S. Garrenton, A. Oh, K. Pitts, D. J. Anderson, N. J. Skelton, B. P. Fauber, B. Pan, S. Malek, D. Stokoe, M. J. Ludlam, K. K. Bowman, J. Wu, A. M. Giannetti, M. A. Starovasnik, I. Mellman, P. K. Jackson, J. Rudolph, W. Wang, G. Fang, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 5299–5304.
[6] a) A. Cruz-Migoni, P. Canning, C. E. Quevedo, C. J. Bataille, N. Bery, A. Miller, A. J. Russell, S. E. Phillips, S. B. Carr, T. H. Rabbitts, *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 2545–2550; b) C. E. Quevedo, A. Cruz-Migoni, N. Bery, A. Miller, T. Tanaka, D. Petch, C. J. Bataille, L. Y. Lee, P. S. Fallon, H. Tulmin, *Nat. Commun.* **2018**, *9*, 3169.
[7] a) P. J. Hajduk, J. Greer, *Nat. Rev. Drug Discovery* **2007**, *6*, 211–219; b) H. Jhoti, A. Cleasby, M. Verdonk, G. Williams, *Curr. Opin. Chem. Biol.* **2007**, *11*, 485–493.
[8] S. B. Shuker, P. J. Hajduk, R. P. Meadows, S. W. Fesik, *Science* **1996**, *274*, 1531–1534.
[9] W. C. Ripka, W. J. Sipio, J. Blaney, *Lect. Heterocycl. Chem.* **1987**, *9*, 95–104.
[10] a) H. Prescher, G. Koch, T. Schuhmann, P. Ertl, A. Bussenault, M. Glick, I. Dix, F. Petersen, D. E. Lizos, *Bioorg. Med. Chem.* **2017**, *25*, 921–925; b) S. Wetzel, A. Schuffenhauer, S. Roggo, P. Ertl, H. Waldmann, *CHIMIA* **2007**, *61*, 355–360.
[11] Version 14.1 ed. (Ed.: C. H. Informa), London, **2005**.
[12] a) A. T. García-Sosa, C. Hetenyi, U. Maran, *J. Comput. Chem.* **2010**, *31*, 174–184; b) M. M. Cavalluzzi, G. F. Mangiatordi, O. Nicolotti, G. Lentini, *Expert Opin. Drug Discovery* **2017**, *12*, 1087–1104.
[13] S. Muratcioglu, T. S. Chavan, B. C. Freed, H. Jang, L. Khavrutskii, R. N. Freed, M. A. Dyba, K. Stefanisko, S. G. Tarasov, A. Gursoy, *Structure* **2015**, *23*, 1325–1335.