Inhibitors Hot Paper

Deutsche Ausgabe: DOI: 10.1002/ange.201505749
Internationale Ausgabe: DOI: 10.1002/anie.201505749

Structural, Biochemical, and Computational Studies Reveal the Mechanism of Selective Aldehyde Dehydrogenase 1A1 Inhibition by Cytotoxic Duocarmycin Analogues

Maximilian F. Koch, Sabrina Harteis, Iris D. Blank, Galina Pestel, Lutz F. Tietze, Christian Ochsenfeld, Sabine Schneider, and Stephan A. Sieber**

Abstract: *Analogues of the natural product duocarmycin bearing an indole moiety were shown to bind aldehyde dehydrogenase 1A1 (ALDH1A1) in addition to DNA, while derivatives without the indole solely addressed the ALDH1A1 protein. The molecular mechanism of selective ALDH1A1 inhibition by duocarmycin analogues was unraveled through cocrystallization, mutational studies, and molecular dynamics simulations. The structure of the complex shows the compound embedded in a hydrophobic pocket, where it is stabilized by several crucial π-stacking and van der Waals interactions. This binding mode positions the cyclopropyl electrophile for nucleophilic attack by the noncatalytic residue Cys302, thereby resulting in covalent attachment, steric occlusion of the active site, and inhibition of catalysis. The selectivity of duocarmycin analogues for ALDH1A1 is unique, since only minor alterations in the sequence of closely related protein isoforms restrict compound accessibility.*

Since their discovery in the 1970s, duocarmycin natural products such as (+)-CC-1065 and Duocarmycin SA have attracted much attention owing to their promising anticancer activity (Figure 1 A).[1] The duocarmycin mode of action is based on its characteristic curved indole structure and a spirocyclopropylcyclohexadienone electrophile.[2] This alkylating moiety exists in conjugation with a vinylogous amide that tames its intrinsic reactivity. Duocarmycins are thus remarkably unreactive in solution.[3] Shape-selective recognition by the narrow AT-rich minor groove of DNA induces a conformational change that results in activation of the cyclopropyl moiety for nucleophilic attack by the 3′ adenine N3 atom (Figure 1 C).[2,4] Because of this unique target-based activation mechanism, duocarmycins were believed to solely address DNA; however, this notion was challenged when activity-based protein profiling (ABPP) with a duocarmycin-derived prodrug probe (seco drug **1**, Figure 1 A) in A459 lung cancer cells revealed aldehyde dehydrogenase 1A1 as an additional target.[5] Moreover, cytotoxic derivatives lacking the DNA-binding unit have been discovered[5,6] that do not exhibit pronounced in situ or in vitro DNA interaction. The affinity for ALDH1A1 was increased, for example, with seco drug **2** (Figure 1 A), and imaging studies confirmed cytosolic localization without pronounced nuclear staining.[6b] Several independent RNA interference (RNAi) studies showed an important role for ALDH1A1 in cancer cell proliferation.[6b,7] However, the

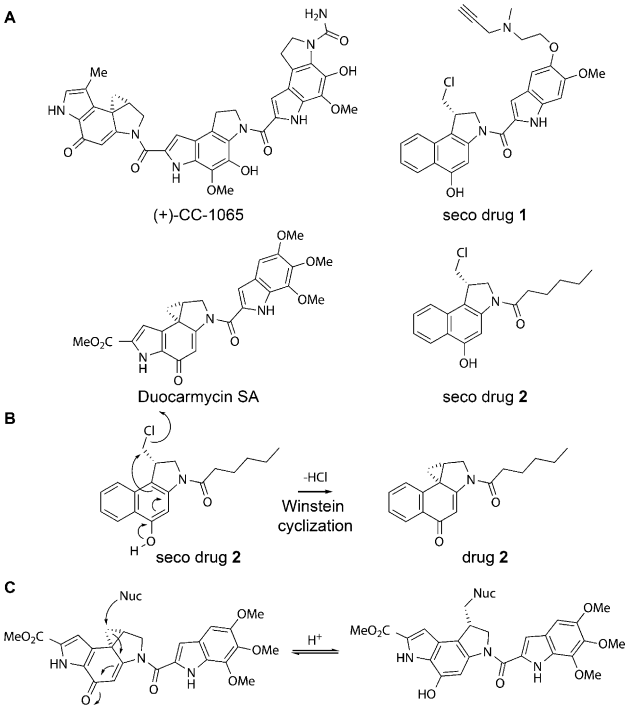

Figure 1. A) Structures of (+)-CC-106, Duocarmycin SA, and seco drugs **1** and **2**. B) Conversion of seco drug **2** into the active molecule via Winstein cyclization. C) Nucleophilic attack at the cyclopropyl ring of Duocarmycin SA by DNA or protein (Nuc = Adenine N3 in DNA or Cys in ALDH1A1).

[*] M. F. Koch,[+] S. Harteis,[+] Dr. S. Schneider, Prof. Dr. S. A. Sieber
Center for Integrated Protein Science Munich CIPS^{M}Department of Chemistry, Technische Universität München
Lichtenbergstrasse 4, 85747 Garching (Germany)
E-mail: sabine.schneider@mytum.de
stephan.sieber@tum.de

I. D. Blank, Prof. Dr. C. Ochsenfeld
Center for Integrated Protein Science Munich CIPS^{M}Department of Chemistry, University of Munich
Butenandtstrasse 7 (C), 81377 München (Germany)

G. Pestel, Prof. Dr. L. F. Tietze
Institute of Organic and Biomolecular Chemistry
Georg-August-Universität Göttingen
Tammannstrasse 2, 37033 Göttingen (Germany)

[+] These authors contributed equally to this work.

Supporting information (including experimental details and structure analysis) and ORCID(s) from the author(s) for this article are available on the WWW under http://dx.doi.org/10.1002/anie.201505749.

exact function of ALDH1A1 in human biology is still not fully understood and tools to specifically manipulate its activity in the presence of related ALDH isoforms are needed.[8] Recent progress on the development of isoform-specific inhibitors in vitro has revealed some promising candidates. However, their selectivity in the context of the whole proteome has not yet been investigated.[9] Up to now, no structural data on the binding of seco drug **2** to ALDH1A1 were available and the proposed specific interaction between this protein and a molecule known to bind DNA was questioned.[10]

Since duocarmycin analogues are currently explored as drug–antibody conjugates (ADCs) a precise characterization of all possible target interactions would be desirable.[11] Herein, we reveal by high-resolution X-ray cocrystal structures of human and sheep ALDH1A1, a perfectly shaped pocket in the active site that precisely interacts with seco drug **2**. Noncatalytic Cys302 covalently traps the compound and several aromatic amino acids engage in crucial interactions through van der Waals forces and π-stacking. The functional role of these residues was analyzed through mutational, kinetic, and computational studies.

The potency and DNA-binding ability of seco drug **2** were evaluated prior to crystallization through MTT cell toxicity and MS-based DNA interaction assays. Seco drug **2** effectively killed A549 cells with a half-maximal inhibitory concentration (IC_{50}) of 28 nM (Figure S1 in the Supporting Information). In contrast to seco drug **1** and duocarmycin SA, both of which covalently modify AT-rich double-stranded DNA, seco drug **2** showed no such modification, thus suggesting a DNA-independent mode of action (Figure S2).[12] Since ALDH1A1 is a confirmed protein target with an important role in cell proliferation, we investigated this binding in greater detail. Recombinant ALDH1A1 from sheep[13] and human[9a] were cocrystallized with seco drug **2** (after Winstein cyclization, Figure 1B) in different crystal forms and the X-ray structures were determined to 1.8 and 2.1 Å resolution, respectively (Table S1 in the Supporting Information). The electron density maps of the compounds reveal binding in the preformed hydrophobic substrate-binding pocket next to the NAD^+ cofactor without the introduction of conformational changes (Figure 2 and Figure S3). The apo and holo forms of the enzymes superimpose with an RMSD of 0.4 Å in both sheep and human structures (Figure S4). The flexible alkyl chain of seco drug **2** points towards the entrance of the active-site pocket and its definition by the electron density varies in the different structures (Figure S5). This is in agreement with inhibition results for indole-bearing duocarmycin analogues (seco drug **1**) that need extra space in this region to accommodate their bulkier substituent.[5,6b] In fact, modeling shows that duocarmycin SA could fit into the channel without the induction of steric clashes (Figure S6). Surprisingly, in both crystal structures the electrophilic cyclopropyl moiety of seco drug **2** did not alkylate the nucleophilic active-site Cys303 (numbering according to the human structure[9a]) but solely the neighboring noncatalytic Cys302.[14] The covalent bond positions the adjacent tricyclic aromatic ring system in a tight-fitting hydrophobic binding pocket that stabilizes the ligand through interactions involving Phe171, Trp178 and Tyr297

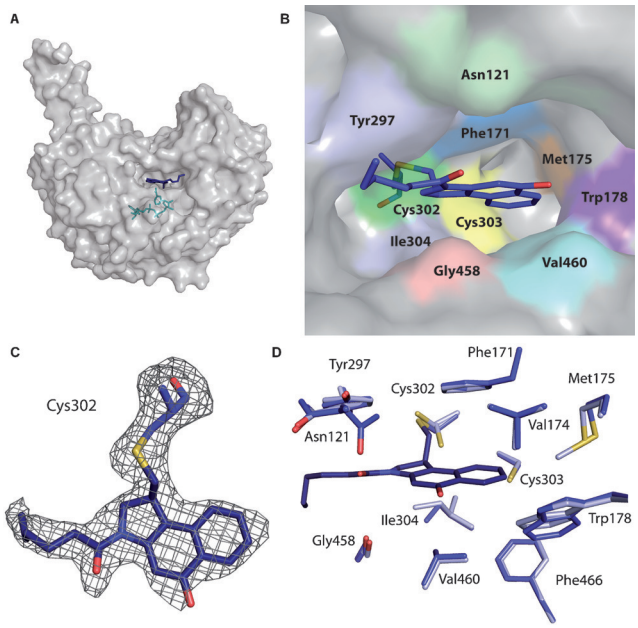

Figure 2. X-ray crystal structure of seco drug **2** bound to human ALDH1A1 (PDB ID: 5AC2, this work). A) Overall structure of the complex, showing ALDH1A1 as a surface and the seco drug and NAD^+ cofactor as blue and green stick models, respectively. B) Close-up view of the hydrophobic catalytic pocket, with seco drug **2** bound to Cys302. C) Simulated-annealing F_o-F_c difference omit electron density map of seco drug **2** and Cys302 contoured at 2.5σ. D) Structural superposition of the active sites of apo ALDH1A1 (light blue, PDB ID: 4WJ9) and in complex with seco drug **2** (dark blue).

(Figure 2B). In line with the high sequence identity between the human and sheep enzymes (92% sequence identity, 97% sequence similarity; Figure S7) the two structures can be superimposed with an RMSD of 0.4 Å.

Based on this binding mode, we next analyzed the impact of exposed amino acids on interaction with the ligand. Owing to the high overall structural similarity, we performed all biochemical studies with the human enzyme and independently confirmed the results by MD calculations[15] with the sheep enzyme. We first evaluated the extent of covalent alkylation by using intact protein mass spectrometry (MS) and gel-based fluorescent labeling with an alkyne probe version of seco drug **2**.[6b] Addition of seco drug **2** or the probe (10-fold excess) at pH 7.4 resulted in ALDH1A1 alkylation after more than 2 h incubation (Figure S8), thus demonstrating more-efficient binding compared to duocarmycin SA (Figure S6). MS/MS sequencing with the seco drug **2** probe confirmed the crystallographic results, with only Cys302 being modified (Figure S9). Even slightly higher pH values (8.0–8.5) were associated with the detection of additional modification sites, including Cys456 and Cys464, as we reported earlier.[5] This assignment can be explained by the elevated nucleophilicity of cysteine thiols at the given pH values, which may reflect nonphysiological conditions. Previous mutational studies with ALDH1A1 had highlighted only Cys303 as essential for catalysis and clearly showed that the neighboring Cys302 is dispensable.[14]

It is thus likely that the binding of seco drug **2** to Cys302 obstructs the active site and thereby prevents substrate access and turnover. Unexpectedly, ALDH1A1 Cys302Ala or Cys302Ser mutants were still inhibited by seco drug **2** (Figure 3A). MS/MS sequencing of both compound-treated mutant enzymes revealed exclusive alkylation of the neighboring active-site Cys303, thus suggesting that once the preferred nucleophile is eliminated, this proximal residue can substitute in the alkylation (Figure 3B).

To obtain a more quantitative measure of the binding kinetics, we determined K_i and k_{inact} values according to the method of Kitz and Wilson for covalent inhibitors.[16] The K_i values for the mutants were up to 10 times higher than those for the wild type, thus suggesting reduced affinity, especially in case of the Ser substitution (Figure S10). The binding preferences of seco drug **2** for either Cys302 or Cys303 were independently validated by molecular dynamics (MD) simulations, which confirmed Cys302 as the favored binding site owing to a stable interaction with the cyclopropyl moiety (Figure 3C). However, Cys303 is within reach of the cyclopropyl moiety and could thus easily bind in the mutants upon slight movement of Trp178, as observed in a previous drug–protein cocrystal structure by Hurley et al. (PDB ID: 4X4L; Figure S11).[9a] In line with experimental data, the calculations further revealed that the Cys302Ala mutant provides better accessibility to Cys303 compared to the bulkier Cys302Ser mutant (Figure S12).

With a mechanistic appreciation of the alkylation reaction, we next focused on the hydrophobic pocket and its interactions with the ligand. The structures revealed several aromatic amino acid residues that stabilize the molecule by van der Waals forces and π-stacking (Figure 2B). To analyze their role in seco drug **2** binding, Phe171, Trp178, and Tyr297 were individually exchanged for Ala and the corresponding mutant proteins were tested for compound inhibition. While the k_{inact} values were comparable to wild type and thus indicative of similar covalent enzyme inactivation rates, the K_i values were significantly elevated (up to 17-fold) in the mutant proteins, thus suggesting a reduction in stabilizing interactions (Figure 4A). MD calculations confirmed this hypothesis and revealed π-stacking and van der Waals interactions with Trp178, as well as just van der Waals interactions with Phe171 and Tyr297 (Figure 4B and Figure S13).[17] Interestingly, Tyr297 showed strong and stable van der Waals interactions with the whole aliphatic chain of seco drug **2** (Figure 4B and Figure S14) during all MD simulations. Compared to Trp178 and Phe171, there are more atoms of Tyr297 within reach of seco drug **2**, thus highlighting this residue as the most important interaction partner (Figure 4B). Additional van der Waals interactions with Leu174, Ile304, and Val460 were also observed (Figure S13).

The unique fit into this π-stacking network, as well as the shape, accessibility, and polarity of the corresponding binding pocket, provide an explanation for the previously observed proteome selectivity for ALDH1A1.[5,6b] Interestingly, even the closely related ALDH isoforms ALDH1A2, ALDH1A3, and ALDH2 (all about 70% sequence identity) were not or only slightly inhibited (ALDH1A2) by seco drug **2**. This is likely due to steric clashes, for example, as a result of the substitution of Ile304 (ALDH1A1) by Thr (ALDH1A2/3) or of Gly458/His293 (ALDH1A1) by Asp458/Phe293 (ALDH2) (Figure 4C–E). Duocarmycin analogues are thus

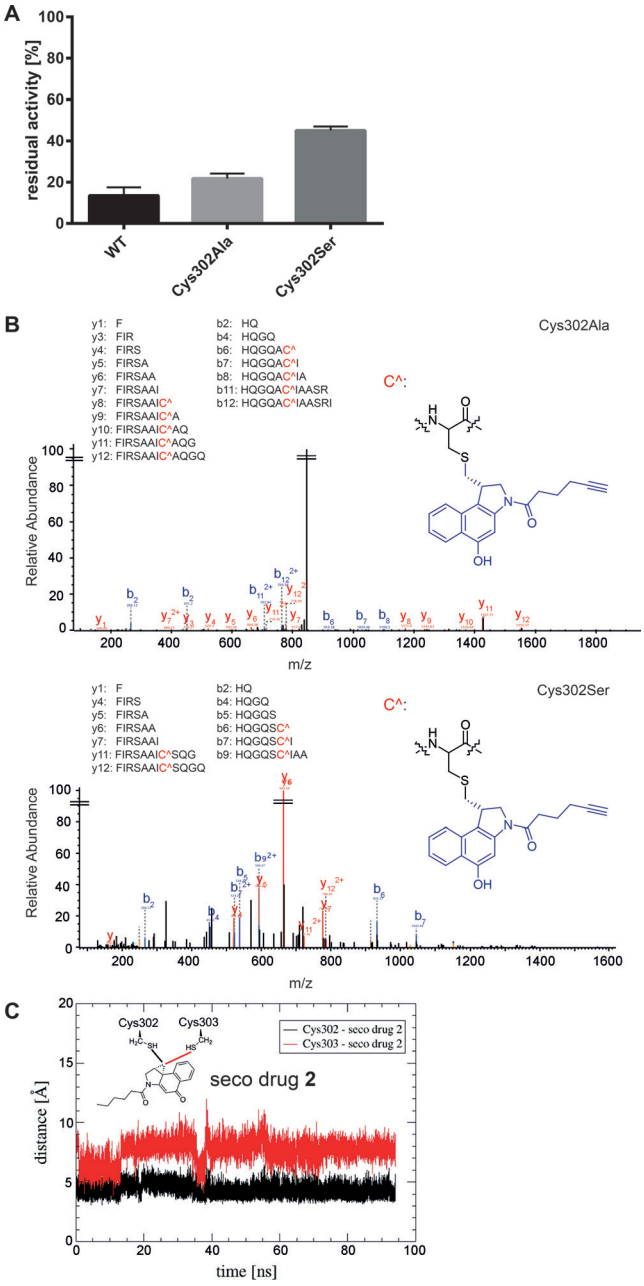

Figure 3. A) Propionaldehyde-based activity assay with wild type (WT) enzyme and the binding-site mutants Cys302Ala and Cys302Ser. Residual activity was measured relative to the uninhibited enzyme activity (shown in %) after 5 h of incubation with a 50-fold excess of seco drug **2** (mean +/− SD, $n = 3$). B) LC–MS/MS binding-site identification in human ALDH1A1 for Cys302Ser and Cys302Ala mutants with a probe version of seco drug **2**. Displayed are fragmented peptides (y and b ions) identified by MS/MS sequencing. The modified cysteine residues are indicated by C^ (highlighted in red). C) Distance of the sulfur atoms of Cys302 and Cys303 to the cyclopropyl moiety of seco drug **2** from MD calculations. Cys302 is always closer to the cyclopropyl moiety than Cys303, which supports the selectivity of seco drug **2** for Cys302.

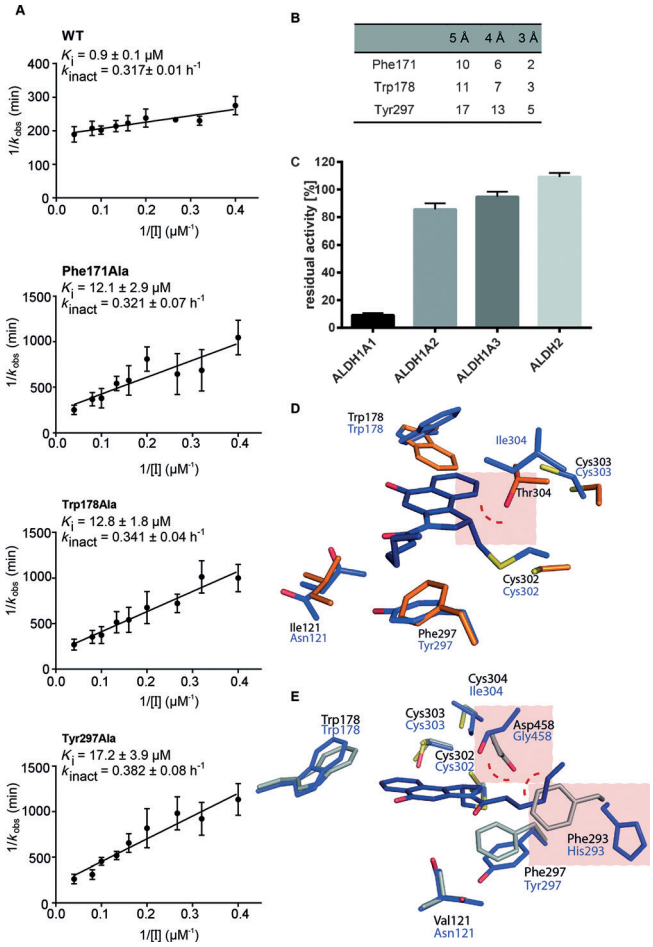

Figure 4. A) Determination of K_i and k_{inact} values according the method of Kitz and Wilson for WT ALDH1A1 and mutants where the interacting residues for seco drug **2** (Phe171Ala, Trp178Ala, and Tyr297Ala) were mutated to alanine. B) Average number of Phe171, Trp178, and Tyr297 atoms interacting with seco drug **2** at the indicated distances (according to MD calculations). C) Propionaldehyde-based activity assay with related human isoforms ALDH1A2, ALDH1A3, and ALDH2. Residual activity was measured after 5 h of incubation with a 50-fold excess of seco drug **2**. D) Active site of human ALDH1A1 with seco drug **2** (blue, PDB ID: 5AC2, this work) and the apo structure of ALDH1A2 (orange, PDB ID: 1BI9). E) Active site of human ALDH1A1 with seco drug **2** (blue, PDB ID: 5AC2, this work) and the apo structure of ALDH2 (gray, PDB ID: 3N80).

important tools for analyzing the role of ALDH1A1 in cancer cells and stem-cell development. In fact, such tools are urgently needed since commercially available ALDH1A1 inhibitors like diethylaminobenzaldehyde and disulfiram are known to be unselective and also bind ALDH2.[18] Recent efforts in broad inhibitor screening combined with rational design have provided attractive starting points for more isoform-specific inhibitors. Among these are reversible binders, which have been shown by structural studies to address the hydrophobic pocket through a different binding mode to that of seco drug **2**.[9a,19] In contrast to seco drug **2**, their inhibitory effect and specificity for ALDH1A1 is based on an extension into a cavity next to the entrance to the binding pocket (Figure S15). Although these molecules exhibit a preference for ALDH1A1 compared to other ALDH isoforms, their proteome-wide selectivity has not yet been evaluated.

In summary, this study elucidates the basis for the selective interaction between a duocarmycin analogue and ALDH1A1. A suite of structural, biochemical, kinetic, and computational methods revealed binding to a hydrophobic pocket, stabilization through van der Waals interactions and π-stacking, and covalent attachment to a cysteine residue directly adjacent to the catalytic residue. This mechanism is unique to ALDH1A1, and even closely related ALDH isoforms are not capable of binding to seco drug **2** owing to steric restrictions at their entrance channel and the catalytic site. The discovery of this unprecedented binding mode is thus not only of upmost importance for drug development, for example, to enable adjustment of the compound design to be exclusively DNA or ALDH1A1 specific (see the Supporting Information), but also provides the first tool for a focused study of ALDH1A1 function in whole proteomes. In addition, future studies will address the question of whether there are noncovalent binders of seco drug **2** that could contribute to the mode of action.

Acknowledgements

S.A.S. acknowledges funding by the Deutsche Forschungsgemeinschaft. S.A.S. and C.O. also acknowledges SFB749 and the excellence cluster CIPSM. L.F.T. was supported by VW-Stiftung and State of Lower Saxony. S.Sch. acknowledges funding by SFB749, CIPSM and Fonds der chemischen Industrie. We would like to thank Prof. Michael Groll for access to the crystallization facility, Megan Wright for critical revision of the manuscript, Volker Kirsch for help with cytotoxicity studies, and Carola Draxler for help with MALDI measurements.

Keywords: ALDH1A1 · duocarmycin · enzymes · inhibitors · protein structure

How to cite: *Angew. Chem. Int. Ed.* **2015**, *54*, 13550–13554
Angew. Chem. **2015**, *127*, 13754–13758

[1] a) D. G. Martin, C. G. Chidester, D. J. Duchamp, S. A. Mizsak, *J. Antibiot.* **1980**, *33*, 902–903; b) T. Yasuzawa, Y. Saitoh, M. Ichimura, I. Takahashi, H. Sano, *J. Antibiot.* **1991**, *44*, 445–447; c) D. L. Boger, D. S. Johnson, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 3642–3649.

[2] K. S. MacMillan, D. L. Boger, *J. Med. Chem.* **2009**, *52*, 5771–5780.

[3] L. F. Tietze, B. Krewer, H. Frauendorf, *Eur. J. Mass Spectrom.* **2009**, *15*, 661–672.

[4] D. L. Boger, R. M. Garbaccio, *Bioorg. Med. Chem.* **1997**, *5*, 263–276.

[5] T. Wirth, K. Schmuck, L. F. Tietze, S. A. Sieber, *Angew. Chem. Int. Ed.* **2012**, *51*, 2874–2877; *Angew. Chem.* **2012**, *124*, 2928–2931.

[6] a) L. F. Tietze, J. M. von Hof, M. Muller, B. Krewer, I. Schuberth, *Angew. Chem. Int. Ed.* **2010**, *49*, 7336–7339; *Angew. Chem.* **2010**, *122*, 7494–7497; b) T. Wirth, G. F. Pestel, V. Ganal, T. Kirmeier, I. Schuberth, T. Rein, L. F. Tietze, S. A. Sieber,

Angew. Chem. Int. Ed. **2013**, *52*, 6921 – 6925; *Angew. Chem.* **2013**, *125*, 7059 – 7063.
[7] L. Yue, Z. M. Huang, S. Fong, S. Leong, J. G. Jakowatz, A. Charruyer-Reinwald, M. Wei, R. Ghadially, *Melanoma Res.* **2015**, *25*, 138 – 148.
[8] P. Marcato, C. A. Dean, C. A. Giacomantonio, P. W. Lee, *Cell Cycle* **2011**, *10*, 1378 – 1384.
[9] a) C. A. Morgan, T. D. Hurley, *J. Med. Chem.* **2015**, *58*, 1964 – 1975; b) A. C. Kimble-Hill, B. Parajuli, C. H. Chen, D. Mochly-Rosen, T. D. Hurley, *J. Med. Chem.* **2014**, *57*, 714 – 722; c) M. Khanna, C. H. Chen, A. Kimble-Hill, B. Parajuli, S. Perez-Miller, S. Baskaran, J. Kim, K. Dria, V. Vasiliou, D. Mochly-Rosen, T. D. Hurley, *J. Biol. Chem.* **2011**, *286*, 43486 – 43494.
[10] a) M. Tercel, S. P. McManaway, E. Leung, H. D. Liyanage, G. L. Lu, F. B. Pruijn, *Angew. Chem. Int. Ed.* **2013**, *52*, 5442 – 5446; *Angew. Chem.* **2013**, *125*, 5552 – 5556; b) L. F. Tietze, S. A. Sieber, *Angew. Chem. Int. Ed.* **2013**, *52*, 5447 – 5449; *Angew. Chem.* **2013**, *125*, 5557 – 5559.
[11] R. C. Elgersma, R. G. Coumans, T. Huijbregts, W. M. Menge, J. A. Joosten, H. J. Spijker, F. M. de Groot, M. M. van der Lee, R. Ubink, D. J. van den Dobbelsteen, D. F. Egging, W. H. Dokter, G. F. Verheijden, J. M. Lemmens, C. M. Timmers, P. H. Beusker, *Mol. Pharm.* **2015**, *12*, 1813 – 1835.
[12] L. Tietze, B. Krewer, J. M. Von Hof, H. Frauendorf, I. Schuberth, *Toxins* **2009**, *1*, 134.
[13] S. A. Moore, H. M. Baker, T. J. Blythe, K. E. Kitson, T. M. Kitson, E. N. Baker, *Structure* **1998**, *6*, 1541 – 1551.
[14] J. Farres, T. T. Wang, S. J. Cunningham, H. Weiner, *Biochemistry* **1995**, *34*, 2592 – 2598.
[15] R. O. Dror, R. M. Dirks, J. P. Grossman, H. Xu, D. E. Shaw, *Annu. Rev. Biophys.* **2012**, *41*, 429 – 452.
[16] R. Kitz, I. B. Wilson, *J. Biol. Chem.* **1962**, *237*, 3245 – 3249.
[17] a) C. R. Martinez, B. L. Iverson, *Chem. Sci.* **2012**, *3*, 2191 – 2201; b) R. S. Paton, J. M. Goodman, *J. Chem. Inf. Model.* **2009**, *49*, 944 – 955.
[18] J. S. Moreb, D. Ucar, S. Han, J. K. Amory, A. S. Goldstein, B. Ostmark, L. J. Chang, *Chem.-Biol. Interact.* **2012**, *195*, 52 – 60.
[19] C. A. Morgan, T. D. Hurley, *Chem.-Biol. Interact.* **2015**, *234*, 29 – 37.

Received: June 24, 2015
Revised: August 5, 2015
Published online: September 16, 2015