Special
Collection

The Selectivity of Fosfosal for STAT5b over STAT5a is Mediated by Arg566 in the Linker Domain

Julian Gräb[a] and Thorsten Berg*[a]

Dedicated to Professor Horst Kessler, on the occasion of his 80th birthday.

Fosfosal is the O-phosphorylated derivative of salicylic acid, with documented clinical use as a prodrug for the treatment of inflammatory diseases. We recently discovered that fosfosal itself inhibits the protein-protein interaction domain, the SH2 domain, of the tumor-related transcription factor STAT5b. Here, we demonstrate that fosfosal is selective for STAT5b over its close homologue STAT5a. This selectivity is mediated by the STAT5b residue Arg566, located in the SH2 domain-adjacent linker domain. Our data provide further evidence for the role of the STAT linker domain in determining the activity of small molecules against the SH2 domain. We present a refined binding model for fosfosal and STAT5b, which can serve as the basis for the development of fosfosal-based STAT5b inhibitors.

STATs (signal transducers and activators of transcription) are dimeric transcription factors which convey signals from the cell membrane to the nucleus.[1] Two forms of the family member STAT5 exist, STAT5a and STAT5b, which exhibit 94% sequence identity on the amino acid level.[2] Despite this high degree of similarity, STAT5a and STAT5b exhibit some nonredundant functions.[3] Although both STAT5 proteins are frequently constitutively activated in human cancers,[4] STAT5b has been shown to be the main driver of tumor cell proliferation in both squamous cell carcinoma[5] and Bcr-Abl-positive leukemia cells.[6] In NPM1-ALK–expressing T-cell lymphomas, STAT5b signaling promotes tumor growth, but STAT5a acts as a tumor suppressor.[7] In contrast, inhibition of STAT5a, but not STAT5b, increases the differentiation of osteoblasts in human bone marrow-derived stromal cells, suggesting selective inhibition of STAT5a as a therapeutic modality against age-related osteoporosis.[8]

The role of STAT5 in cancer has encouraged the development of a number of STAT5 inhibitors.[9] These are thought to act on the STAT5 protein-protein interaction domain, the Src homology 2 (SH2) domain. However, despite the known non-identical roles of STAT5a and STAT5b, the differential effects of most STAT5 inhibitors on the two proteins have not been tested. Our group recently developed the catechol bisphosphates Stafib-1,[10] Stafib-2,[11] and Capstafin[12] as the first small molecules that inhibit STAT5b with selectivity over STAT5a,[13] and the m-terphenyl phosphate Stafia-1 as the first small molecule that inhibits STAT5a with selectivity over STAT5b.[14]

The starting point for the development of our catechol bisphosphate-based STAT5b inhibitors[10] was fosfosal (O-phosphorylated salicylic acid, Figure 1A), which we identified as a STAT5b SH2 domain inhibitor during high-throughput screening of natural products and known bioactive compounds.[9b] Fosfosal (Disdolen®, Uriach, Spain) has been in clinical use as a phosphate prodrug of salicylic acid for treatment of inflammatory diseases.[15] The selectivity of fosfosal itself for STAT5b over STAT5a was not investigated in the original study,[9b] because the competitive binding assay against STAT5a[10] had not yet been developed.

Here, we provide a detailed activity analysis of fosfosal against STAT5a and STAT5b wild-type and point mutant proteins. Our data lead to a refined binding model for fosfosal and STAT5b, which will support the rational development of fosfosal-based STAT5b inhibitors, and highlights the relevance of the STAT linker domain for the function of the adjacent SH2 domain.

Analysis of fosfosal in a fluorescence polarization (FP)-based competitive binding assay indicated a fivefold weaker activity against STAT5a ($K_i = 148 \pm 9$ μM) than STAT5b ($K_i = 29.3 \pm 1.8$ μM, Figure 1B, Table 1, and Table S1 in the Supporting Information). This result cannot be explained by the previously published docking model of fosfosal and STAT5b (Figure 1C),[9b] which postulated interactions between the negatively charged groups of fosfosal and the STAT5b amino acids Lys600 and Arg618 (Figure 1C) to be the main determinants of protein binding. Because Lys600 and Arg618 are conserved between STAT5b and STAT5a, they cannot govern binding specificity. The published binding pose was generated by docking fosfosal into a rigid STAT5b homology model based on the crystal structure of STAT5a[16] using AutoDock Vina,[17] which does not account for protein flexibility.

We recently identified Arg566 in the STAT5b linker domain as the main determinant of the STAT5b selectivity of catechol bisphosphates.[18] Although Arg566 appears to be too distant to

[a] *J. Gräb, Prof. Dr. T. Berg*
Leipzig University, Institute of Organic Chemistry
Johannisallee 29, 04103 Leipzig (Germany)
E-mail: tberg@uni-leipzig.de
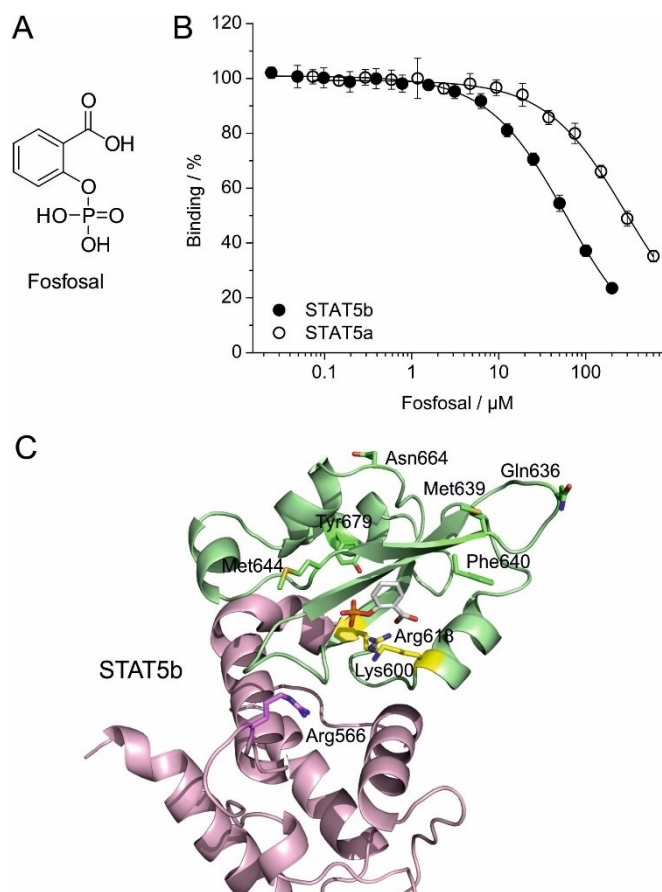
A

Fosfosal

B

C

Figure 1. A) Structure of fosfosal. B) Activity of fosfosal against wild-type STAT5b and STAT5a in competitive FP assays. C) Previously published docking pose of fosfosal into STAT5b generated using AutoDock Vina.[9b] Linker domain: magenta; SH2 domain: green. The side chains of the conserved Lys600 and Arg618 are shown with carbon atoms in yellow; the carbon atoms of the side chains of the divergent amino acids in the SH2 domain and in position 566 of the linker domain are colored according to the corresponding domain.

Table 1. Activities of fosfosal against wild-type and mutant STAT5 proteins.[a]

Protein	K_i [μM] or % inhibition at 600 μM
STAT5b	29.3 ± 1.8 μM
STAT5a	148 ± 9 μM
STAT5b Arg566Trp	$24 \pm 2\%$ inhibition
STAT5a Trp566Arg	10.2 ± 0.7 μM
STAT5b-6M	10.7 ± 0.6 μM
STAT5b-7M	189 ± 3 μM
STAT5b Arg566Ala	$22 \pm 3\%$ inhibition
STAT5b Arg566Glu	no inhibition
STAT5b Met644Lys	17.2 ± 0.9 μM
STAT5a Trp566Arg/Lys644Met	12.3 ± 0.6 μM

[a] Inhibition constants K_i have been calculated from the IC_{50} values shown in Supporting Table S1, using the published equation.[19]

selectivity of fosfosal for STAT5b over STAT5a, we tested the activity of fosfosal against the point mutant protein STAT5b Arg566Trp in a fluorescence polarization assay. Activity against STAT5b Arg566Trp, in which the arginine at position 566 of STAT5b was swapped for the tryptophan at position 566 of STAT5a (Figure S1), was drastically decreased ($24 \pm 2\%$ inhibition at 600 μM, the highest concentration tested) compared to wild-type STAT5b ($K_i = 29.3 \pm 1.8$ μM, Figure 2A). Conversely, the activity of fosfosal against the reverse cross-over mutant STAT5a Trp566Arg was increased by 14-fold ($K_i = 10.2 \pm 0.7$ μM) compared to wild type STAT5a ($K_i = 148 \pm 9$ μM, Figure 2A, Tables 1 and S1). These data indicate that the amino acid at position 566 of the linker domain is crucial for inhibition by fosfosal. We also tested fosfosal against the STAT5b mutant Gln636Pro/ Met639Asn/Phe640Leu/Met644Lys/Asn664Ser/Tyr679Phe (dubbed STAT5b-6M), in which all six of the amino acids which differ from the SH2 domain of STAT5a were mutated to the corresponding STAT5a residues (Figure S1).[18] The activity of fosfosal against STAT5b-6M ($K_i = 10.7 \pm 0.6$) was not reduced to the level of wild type STAT5a ($K_i = 148 \pm 9$ μM), but instead increased almost 3-fold compared to wild type STAT5b ($K_i =$
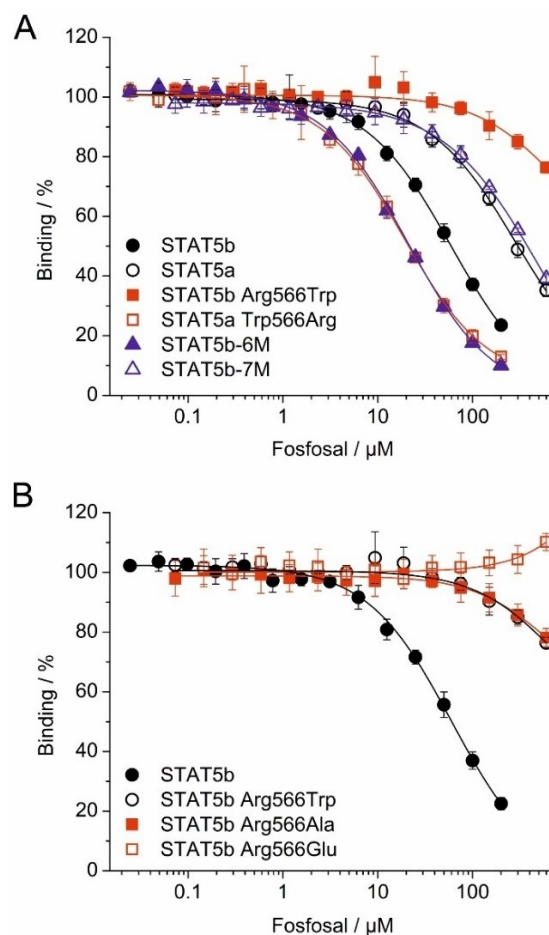
A

B

Figure 2. A) Activity of fosfosal in FP assays against wild-type STAT5b, wild-type STAT5a, STAT5b Arg566Trp, STAT5a Trp566Arg, STAT5b-6M, and STAT5b-7M. B) Activity of fosfosal in FP assays against wild-type STAT5b, STAT5b Arg566Trp, STAT5b Arg566Ala, and STAT5b Arg566Glu.

interact with fosfosal as bound to the SH2 domain of our original homology model (Figure 1C), Arg566 is located in a flexible loop which can shift towards the SH2 domain.[18] To investigate whether Arg566 is also responsible for the moderate

ChemBioChem

Communications
doi.org/10.1002/cbic.202000111

Chemistry Europe
European Chemical
Societies Publishing

29.3 ± 1.8 μM, Figure 2A, Tables 1 and S1). Additionally, we tested fosfosal against the STAT5b mutant Arg566Trp/ Gln636Pro/Met639Asn/Phe640Leu/Met644Lys/Asn664Ser/ Tyr679Phe (dubbed STAT5b-7M),[18] in which the Arg566Trp mutation was introduced in STAT5b-6M. STAT5b-7M showed substantially decreased affinity for fosfosal ($K_i = 189 \pm 3$ μM) compared to STAT5b-6M ($K_i = 10.7 \pm 0.6$ μM, Figure 2A, Tables 1 and S1). Taken together, these results indicate that the preference of fosfosal for STAT5b is mediated by the amino acid in position 566 of the STAT5a/b linker domain, and not by the divergent amino acids in the SH2 domain.

The strongly reduced activity of fosfosal against STAT5b Arg566Trp (24 ± 2% inhibition at 600 μM) as compared to wild-type STAT5b ($K_i = 29.3 \pm 1.8$ μM, Figure 2A) may result from the absence of an electrostatic interaction between one of the negatively charged groups of fosfosal and the guanidinium group of the side chain of arginine, or be caused by the presence of the bulky indole side chain of tryptophan. To differentiate between these two possibilities, we tested the mutant proteins STAT5b Arg566Ala and STAT5b Arg566Glu.[18] While the activity of fosfosal against STAT5b Arg566Ala (22 ± 3% inhibition at 600 μM) was similarly low to its activity against STAT5b Arg566Trp (24 ± 2% inhibition at 600 μM), there was no inhibition of STAT5b Arg566Glu (Figure 2B, Tables 1 and S1). These data strongly suggest that the reduced binding of fosfosal to STAT5b Arg566Trp is not caused by the presence of the indole side chain of Trp566, but by the absence of the positively charged side chain of STAT5b Arg566.

To rationalize our experimental data, we docked fosfosal into a previously-described refined STAT5b homology model, in which the flexible loop in the STAT5b linker domain had been shifted towards the SH2 domain.[18] We allowed side chain flexibility for amino acids in the vicinity of the putative binding site, using the newly developed docking program AutoDock FR.[20] This identified a docking pose for fosfosal which was consistent with the experimental data (Figure 3). In this docking pose, the negatively charged groups of fosfosal interact not only with Lys600 and Arg618 in the SH2 domain, but also with Arg566 in the linker domain of STAT5b. The conserved SH2 domain amino acids Ser620 and Ser622, which form hydrogen bonds with the phosphate group of phosphotyrosine-containing peptide sequences in crystal structures of STAT proteins,[21] are expected to be engaged in hydrogen bond networks with the carboxylate and the phosphate group of fosfosal.

An interesting observation is that the activity of fosfosal against the mutants which comprise the STAT5a SH2 domain together with an arginine in position 566, STAT5a Trp566Arg ($K_i = 10.2 \pm 0.7$ μM) and STAT5b-6M ($K_i = 10.7 \pm 0.6$ μM), is higher than the activity against wild-type STAT5b ($K_i = 29.3 \pm 1.8$ μM, Figure 2A, Tables 1 and S1). Similar observations have been made for catechol bisphosphates[18] and the nucleoside triphosphates ATP and GTP,[23] which are also more active against these two mutants than against wild-type STAT5b. From the new docking model of fosfosal (Figure 3), it appeared that the side chain of the divergent STAT5a/b amino acid in position 644 is in proximity to the small-molecule binding site. Although the Met644 of STAT5b might provide a contribution to binding of
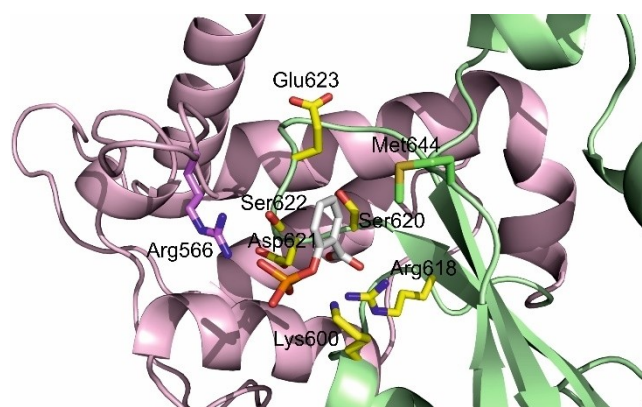

Figure 3. Docking of fosfosal into a homology model of STAT5b, in which the linker domain loop bearing Arg566 is shifted towards the SH2 domain, using AutoDock FR. The colored amino acid side chains were defined as flexible in the docking process. Linker domain: magenta; SH2 domain: green. The side chains of the conserved amino acids are shown with carbon atoms in yellow; the side chains of the divergent amino acids Arg566 and Met644 are shown with carbon atoms colored like the corresponding domain. The figure was generated using PyMOL.[22]

fosfosal, the Lys644 present in STAT5a seems more likely to do so, either through strong cation-π interactions with fosfosal's benzene ring, or through electrostatic interactions with the negatively charged groups of fosfosal. In order to investigate the role of the divergent amino acids in position 644, we created the point mutant protein STAT5b Met644Lys, and found that this was indeed inhibited by fosfosal to a higher extent ($K_i = 17.2 \pm 0.9$ μM) than wild-type STAT5b ($K_i = 29.3 \pm 1.8$ μM, Figure 4, Tables 1 and S1). Conversely, the double mutant STAT5a Trp566Arg/Lys644Met was inhibited to a slightly lesser extent ($K_i = 12.3 \pm 0.6$ μM) than the single mutant STAT5a Trp566Arg ($K_i = 10.2 \pm 0.7$ μM, Figure 4, Tables 1 and S1). This indicates that the amino acid in position 644 does provide a contribution to fosfosal binding. However, the residual activity difference between both STAT5a Trp566Arg ($K_i = 10.2 \pm 0.7$ μM)
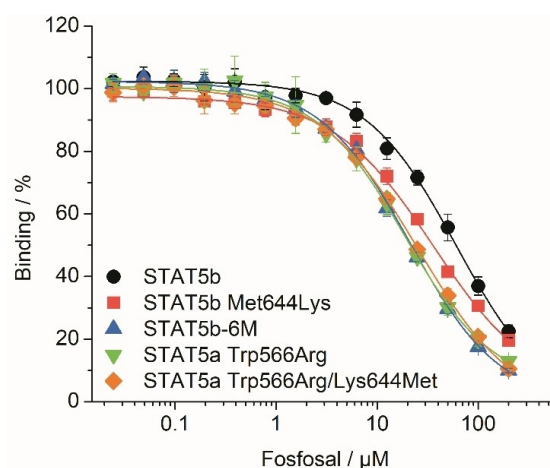

Figure 4. Activity of fosfosal in FP assays against wild-type STAT5b, STAT5b Met644Lys, STAT5b-6M, STAT5a Trp566Arg, and STAT5a Trp566Arg/Lys644-Met.

ChemBioChem

Communications
doi.org/10.1002/cbic.202000111

Chemistry Europe
European Chemical
Societies Publishing

and STAT5b-6M ($K_i = 10.7 \pm 0.6$ μM), and STAT5b Met644Lys ($K_i = 17.2 \pm 0.9$ μM) suggests that one or more of the remaining five divergent amino acids in the STAT5a SH2 domain also contributes to the higher activity of fosfosal against STAT5a Trp566Arg and STAT5b-6M.

In summary, we show that the STAT5b inhibitor fosfosal exhibits fivefold selectivity over the highly homologous STAT5a, and that this selectivity is mediated by Arg566 in the STAT5b linker domain. The refined binding mode of fosfosal presented here will facilitate the development of improved fosfosal-based STAT5b inhibitors, and is consistent with the binding mode of catechol bisphosphate-based STAT5b inhibitors.[18] In both cases, an additional negatively charged group in the *ortho*-position of the phenyl phosphate moiety (a carboxylate in case of fosfosal, or a second phosphate group in case of catechol bisphosphates) mediates STAT5b selectivity by allowing concomitant targeting of both the SH2 domain and Arg566. Our study reveals that the STAT5b-selectivity of catechol bisphosphate-based inhibitors has its origins in the lead structure fosfosal, and provides further evidence for a contribution of the STAT linker domain to the function of the SH2 domain.[24]

Conflict of Interest

The authors declare no conflict of interest.

[1] a) D. E. Levy, J. E. Darnell Jr., *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 651–662; b) G. Miklossy, T. S. Hilliard, J. Turkson, *Nat. Rev. Drug Discovery* **2013**, *12*, 611–629.

[2] a) L. Hennighausen, G. W. Robinson, *Genes Dev.* **2008**, *22*, 711–721; b) I. Wittig, B. Groner, *Curr. Drug Targets Immune Endocr. Metab. Disord.* **2005**, *5*, 449–463.

[3] a) X. Liu, G. W. Robinson, K. U. Wagner, L. Garrett, A. Wynshaw-Boris, L. Hennighausen, *Genes Dev.* **1997**, *11*, 179–186; b) G. B. Udy, R. P. Towers, R. G. Snell, R. J. Wilkins, S. H. Park, P. A. Ram, D. J. Waxman, H. W. Davey, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 7239–7244; c) B. Basham, M. Sathe, J. Grein, T. McClanahan, A. D'Andrea, E. Lees, A. Rascle, *Nucleic Acids Res.* **2008**, *36*, 3802–3818.

[4] T. Berg, *ChemBioChem* **2008**, *9*, 2039–2044.

[5] a) S. Xi, Q. Zhang, W. E. Gooding, T. E. Smithgall, J. R. Grandis, *Cancer Res.* **2003**, *63*, 6763–6771; b) P. L. Leong, S. Xi, S. D. Drenning, K. F. Dyer, A. L. Wentzel, E. C. Lerner, T. E. Smithgall, J. R. Grandis, *Oncogene* **2002**, *21*, 2846–2853.

[6] a) L. Casetti, S. Martin-Lanneree, I. Najjar, I. Plo, S. Auge, L. Roy, J. C. Chomel, E. Lauret, A. G. Turhan, I. Dusanter-Fourt, *Cancer Res.* **2013**, *73*, 2052–2058; b) M. Schaller-Schonitz, D. Barzan, A. J. Williamson, J. R. Griffiths, I. Dallmann, K. Battmer, A. Ganser, A. D. Whetton, M. Scherr, M. Eder, *PLoS One* **2014**, *9*, e97243; c) S. Kollmann, E. Grundschober, B. Maurer, W. Warsch, R. Grausenburger, L. Edlinger, S. Huuhtanen, S. Lagger, L. Hennighausen, P. Valent, T. Decker, B. Strobl, M. Mueller, S. Mustjoki, A. Hoelbl-Kovacic, V. Sexl, *Leukemia* **2019**.

[7] a) Q. Zhang, H. Y. Wang, X. B. Liu, M. A. Wasik, *Nat. Med.* **2007**, *13*, 1341–1348; b) M. Nieborowska-Skorska, A. Slupianek, L. Xue, Q. Zhang, P. N. Raghunath, G. Hoser, M. A. Wasik, S. W. Morris, T. Skorski, *Cancer Res.* **2001**, *61*, 6517–6523.

[8] K. M. Lee, K. H. Park, J. S. Hwang, M. Lee, D. S. Yoon, H. A. Ryu, H. S. Jung, K. W. Park, J. Kim, S. W. Park, S. H. Kim, Y. M. Chun, W. J. Choi, J. W. Lee, *Cell Death Dis.* **2018**, *9*, 1136.

[9] a) J. Müller, B. Sperl, W. Reindl, A. Kiessling, T. Berg, *ChemBioChem* **2008**, *9*, 723–727; b) M. Gräber, W. Janczyk, B. Sperl, N. Elumalai, C. Kozany, F. Hausch, T. A. Holak, T. Berg, *ACS Chem. Biol.* **2011**, *6*, 1008–1014; c) B. D. Page, H. Khoury, R. C. Laister, S. Fletcher, M. Vellozo, A. Manzoli, P. Yue, J. Turkson, M. D. Minden, P. T. Gunning, *J. Med. Chem.* **2012**, *55*, 1047–1055; d) A. A. Cumaraswamy, A. M. Lewis, M. Geletu, A. Todic, D. B. Diaz, X. R. Cheng, C. E. Brown, R. C. Laister, D. Muench, K. Kerman, H. L. Grimes, M. D. Minden, P. T. Gunning, *ACS Med. Chem. Lett.* **2014**, *5*, 1202–1206; e) Z. Liao, L. Gu, J. Vergalli, S. A. Mariani, M. De Dominici, R. K. Lokareddy, A. Dagvadorj, P. Purushottamachar, P. A. McCue, E. Trabulsi, C. D. Lallas, S. Gupta, E. Ellsworth, S. Blackmon, A. Ertel, P. Fortina, B. Leiby, G. Xia, H. Rui, D. T. Hoang, L. G. Gomella, G. Cingolani, V. Njar, N. Pattabiraman, B. Calabretta, M. T. Nevalainen, *Mol. Cancer Ther.* **2015**, *14*, 1777–1793; f) A. Berg, T. Berg, *Bioorg. Med. Chem. Lett.* **2017**, *27*, 3349–3352; g) E. L. Wong, E. Nawrotzky, C. Arkona, B. G. Kim, S. Beligny, X. Wang, S. Wagner, M. Lisurek, D. Carstanjen, J. Rademann, *Nat. Commun.* **2019**, *10*, 66.

[10] N. Elumalai, A. Berg, K. Natarajan, A. Scharow, T. Berg, *Angew. Chem. Int. Ed.* **2015**, *54*, 4758–4763; *Angew. Chem.* **2015**, *127*, 4840–4845.

[11] N. Elumalai, A. Berg, S. Rubner, L. Blechschmidt, C. Song, K. Natarajan, J. Matysik, T. Berg, *Sci. Rep.* **2017**, *7*, 819.

[12] N. Elumalai, A. Berg, S. Rubner, T. Berg, *ACS Chem. Biol.* **2015**, *10*, 2884–2890.

[13] N. Elumalai, K. Natarajan, T. Berg, *Bioorg. Med. Chem.* **2017**, *25*, 3871–3882.

[14] K. Natarajan, D. Müller-Klieser, S. Rubner, T. Berg, *Chem. Eur. J.* **2020**, *26*, 148–154.

[15] J. Ramis, I. Gich, J. Torrent, R. Mis, F. Jane, J. Forn, *Int. J. Clin. Pharmacol. Ther. Toxicol.* **1989**, *27*, 352–357.

[16] D. Neculai, A. M. Neculai, S. Verrier, K. Straub, K. Klumpp, E. Pfitzner, S. Becker, *J. Biol. Chem.* **2005**, *280*, 40782–40787.

[17] O. Trott, A. J. Olson, *J. Comput. Chem.* **2010**, *31*, 455–461.

[18] J. Gräb, A. Berg, L. Blechschmidt, B. Klüver, S. Rubner, D. Y. Fu, J. Meiler, M. Gräber, T. Berg, *ACS Chem. Biol.* **2019**, *14*, 796–805.

[19] Z. Nikolovska-Coleska, R. Wang, X. Fang, H. Pan, Y. Tomita, P. Li, P. P. Roller, K. Krajewski, N. G. Saito, J. A. Stuckey, S. Wang, *Anal. Biochem.* **2004**, *332*, 261–273.

[20] P. A. Ravindranath, S. Forli, D. S. Goodsell, A. J. Olson, M. F. Sanner, *PLoS Comput. Biol.* **2015**, *11*, e1004586.

[21] a) S. Becker, B. Groner, C. W. Müller, *Nature* **1998**, *394*, 145–151; b) X. Mao, Z. Ren, G. N. Parker, H. Sondermann, M. A. Pastorello, W. Wang, J. S. McMurray, B. Demeler, J. E. Darnell Jr., X. Chen, *Mol. Cell* **2005**, *17*, 761–771.

[22] Schrodinger, LLC, *The PyMol Molecular Graphics System*, Version 0.99, **2015**.

[23] A. Berg, B. Sperl, T. Berg, *ChemBioChem* **2019**, *20*, 2227–2231.

[24] a) E. Yang, M. A. Henriksen, O. Schaefer, N. Zakharova, J. E. Darnell Jr., *J. Biol. Chem.* **2002**, *277*, 13455–13462; b) E. Yang, Z. Wen, R. L. Haspel, J. J. Zhang, J. E. Darnell Jr., *Mol. Cell. Biol.* **1999**, *19*, 5106–5112; c) C. Mertens, B. Haripal, S. Klinge, J. E. Darnell, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 14811–14816; d) A. T. Namanja, J. Wang, R. Buettner, L. Colson, Y. Chen, *J. Mol. Biol.* **2016**, *428*, 579–589.