NMR Spectroscopy

DOI: 10.1002/anie.201411658

Fragment Screening and Druggability Assessment for the CBP/p300 KIX Domain through Protein-Observed ^{19}F NMR Spectroscopy**

*Clifford T. Gee, Edward J. Koleski, and William C. K. Pomerantz**

Abstract: *^{19}F NMR spectroscopy of labeled proteins is a sensitive method for characterizing structure, conformational dynamics, higher-order assembly, and ligand binding. Fluorination of aromatic side chains has been suggested as a labeling strategy for small-molecule ligand discovery for protein–protein interaction interfaces. Using a model transcription factor binding domain of the CREB binding protein (CBP)/p300, KIX, we report the first full small-molecule screen using protein-observed ^{19}F NMR spectroscopy. Screening of 508 compounds and validation by ^{1}H–^{15}N HSQC NMR spectroscopy led to the identification of a minimal pharmacaphore for the MLL-KIX interaction site. Hit rate analysis for the CREB-KIX and MLL-KIX sites provided a metric to assess the ligandability or "druggability" of each interface informing future medicinal chemistry efforts. The structural information from the simplified spectra and data collection speed, affords a new screening tool for analysis of protein interfaces and discovery of small molecules.*

Protein–protein interactions (PPIs) are essential nodes in almost all biological processes within the cell. Modulation of these biomolecular interactions through small molecules is now a validated approach for therapeutic intervention[1] and chemical probe development to improve our understanding of biological systems.[2] Despite this validation, disruption of these interactions is difficult due to the large and plastic interfaces at the protein surface and lack of deep hydrophobic binding sites for accommodating traditional drug-like molecules.[3] This is particularly the case for the transient and dynamic interactions found for many transcription factor–protein binding events that occur with modest affinity (K_d = 0.1–100 μM).

We previously used protein-observed fluorine NMR spectroscopy (PrOF NMR) for characterizing ligand binding at PPI sites.[4] We employed singly labeled fluorinated aromatic amino acids due to the enrichment of aromatic amino acids at PPI interfaces.[5] In many cases, the modest protein structural perturbations from a single aryl hydrogen to fluorine substitution, the fast data acquisition for small to medium sized proteins, and the orthogonality of fluorine with biological nuclei, leads to a simplified and readily obtainable 1D-NMR spectrum reporting directly on protein structure.[6] We used this method to characterize small-molecule binding to three different transcription factor domains: bromodomains Brd4, BrdT, and BPTF.[4b] We also studied the protein interaction domain of CBP/p300, KIX, by PrOF NMR and identified small-molecule ligands 1-10 and 1G7.[4a] Beyond model studies, the PrOF NMR method has yet to be tested in a full discovery format. Herein, we report the first full small-molecule screen using PrOF NMR, evaluate its effectiveness for assessing protein druggability at two different binding sites within the same protein, KIX, and report structure activity relationships (SAR) from several newly discovered small molecules.

Molecules of low molecular weight and complexity, called fragments, have helped increase successful screening outcomes against difficult protein targets.[7] NMR spectroscopy has become a preferred method for screening to identify low affinity ligands, in many cases using 1D-NMR ligand-based methods.[8] Analysis of fragment screen hit rates provides valuable information about the ligandability or "druggability" of the proteins,[9] and further development of fragments has led to molecules with more efficacious binding affinity and improved physicochemical properties compared to molecules obtained from traditional high-throughput screenings.[10] To test the feasibility of using PrOF NMR in a fragment screening format, we chose the KIX domain.

KIX is a useful model protein for characterizing protein–small molecule interactions and studying the conformational plasticity of this dynamic protein domain. KIX binds to over a dozen transcription factors (e.g., MLL, E2A-PBX1, and CREB) through two binding sites.[11] These sites as well as new cryptic sites cannot readily be distinguished in a direct binding 1D-NMR ligand-based experiment. Prior screening efforts offer valuable ligand information for comparison with our PrOF NMR experiments. Two natural product depsides, sekikaic acid and lobaric acid, are reported as the most potent KIX inhibitors at the MLL-KIX PPI site (IC_{50} = 34 and 17 μM respectively). These compounds isolated from natural products were discovered after a screen of 50000 commercial library compounds had resulted in no hits.[12] The only HSQC NMR screen resulted in two lead compounds, KG-122 (pamoic acid) and KG-501 (naphthol AS-E phosphate) from a 762 member peptidomimetic library (Figure 1).[13] Isoxazolidine derivatives, discovered first as transcriptional activation domain mimics,[14] were also later shown to be KIX ligands.[15] Despite the paucity of potent ligands, both CREB-KIX, and E2A-KIX interactions have been proposed as important PPI targets for treating blood cancers due to their

[*] C. T. Gee, E. J. Koleski, Prof. W. C. K. Pomerantz
Department of Chemistry, University of Minnesota
207 Pleasant St. SE, Twin Cities (USA)
E-mail: wcp@umn.edu

[**] This project was supported by the University of Minnesota, the NSF-CAREER Award CHE-1352091, and the NIH training grant T32-GM08700 (C.T.G). We also appreciate the generous support of the Garber family and relatives.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201411658.

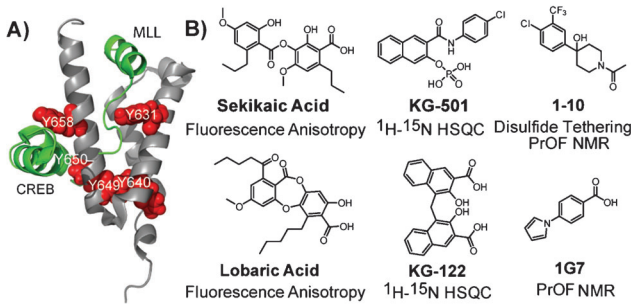


Figure 1. Solution structure of KIX and known ligands. A) NMR solution structure of KIX (PDB: 2LXT) with peptides MLL and pKID (CREB) in green. Tyrosine residues are red. B) KIX ligands and their corresponding methods of discovery.

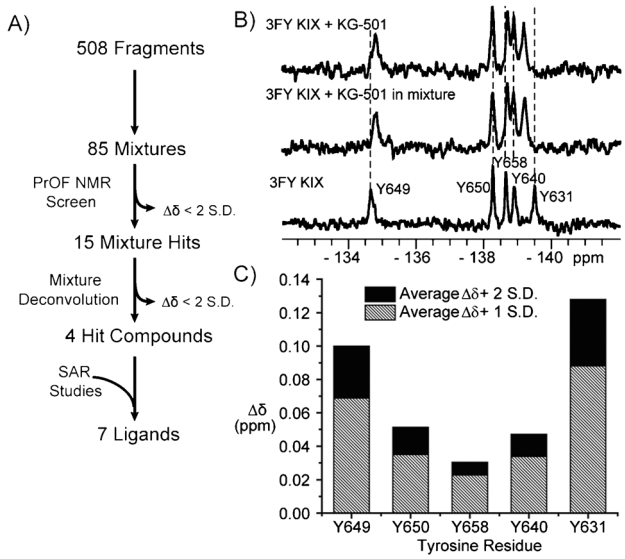


Figure 2. Fragment screening by PrOF NMR spectroscopy. A) Flowchart for screening of the fragment library for KIX ligands. 508 fragments were screened by PrOF NMR, yielding 15 mixture hits and seven final ligands after deconvolution and SAR studies. B) KG-501 was used to test for KIX ligand identification in isolation or in a mixture of six total compounds. C) Statistical cutoffs for chemical-shift perturbations found in the screen. Samples with perturbations greater than two standard deviations above were identified as hits.

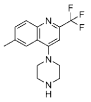
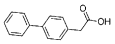
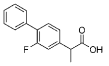
regulatory role in blood cell proliferation further motivating our research.[11b,16]

PPIs involving the KIX domain occur through two distinct and allosterically coupled binding sites on this 3-helical bundle protein. These sites are often referred to as the MLL and CREB/c-Myb sites.[12] To conduct the NMR screen, we replaced native tyrosine residues of KIX with 3-fluorotyrosine (3FY). The MLL site contains one tyrosine (Y631), which makes a critical contact with the native protein. A sequence analysis by Näär et al. of KIX domains shows this residue is conserved in humans, *C. elegans*, and *drosophila*.[17] Y649 is evolutionarily conserved in all KIX domains. Fluorine resonances from Y649, Y650, and Y658 experience significant chemical shifts when CREB binds to KIX as well as Y631 from allostery.[4a] Consistent with natural ligands, our computed druggability analysis using SiteMap identified two druggable sites near the tyrosine residues at both the MLL (10.8 Å from Y631) and CREB/Myb sites (4–10 Å from Y649, Y650, and Y658).[18] Y640, which structurally stabilizes KIX by a cation–π interaction with R600,[19] is found in 97% of KIX domains, and may represent a new small-molecule site for regulating protein conformation.[17] Importantly, singly fluorinated aromatic residues in KIX only modestly perturb secondary structure and ligand binding.[4a]

We expressed 3FY-labeled KIX (12 kDa) in good yield (70 mg L^{-1}) with high labeling efficiency (>98%) for the screen.[18] Our fragment library was generated from the Maybridge rule of three commercial set, combining compounds into 85 mixtures of five or six compounds at a total stock concentration of 33.3 mM per compound in DMSO. As a positive control, we tested a known ligand, KG-501 (833 μM), and identified a binding interaction both in isolation, consistent with prior results,[4a] and in a mixture based on chemical shift perturbation of Y631 (Figure 2B). For the NMR screen we used 40 μM KIX (≈20 mg total). Chemical shift information was acquired in five minutes yielding approximately 510 min of experiment time for screening the 85 mixtures, including additional short reference experiments for each mixture. This experiment time is faster than ^{1}H–^{15}N SOFAST HMQC NMR experiments for similar sized proteins.[20] All mixtures were screened at 833 μM small molecule and 2.5% DMSO. Statistical cutoffs for chemical shift perturbation were set to two standard deviations from the average perturbations from the screen yielding 15 mixtures (Figure 2C). Each mixture was subsequently deconvoluted by individual analysis of each compound leading to four verified ligands (**1**–**4**, 0.8% hit rate). The reduction in hits was due to apparent additive effects of fragments which prevented the identification of a sole compound responsible for chemical shift perturbations in a given mixture.

Ligand titration experiments by PrOF NMR spectroscopy were performed to determine the dissociation constants for the small molecules obtained by monitoring changes in chemical shift ($\Delta\delta$). Three of the four ligands (**2**, **3**, and **4**) discovered from the screen were found to have low mM binding affinities for KIX. These compounds contained either an aryl or phenylacetic acid group (Table 1). Small molecules **3** and **4** exhibited a binding isotherm consistent with one-to-one binding, whereas **2** potentially exhibited higher-order binding above 2 mM based on the binding isotherm generated. Consequently the K_d for **2** was estimated based on fitting the data up to 2 mM. We ruled out activity due to small-molecule aggregation above 2 mM based on well-resolved small-molecule resonances in the presence and the absence of detergent.[18] Given that KG-122 is a diacid compound that binds to KIX in the MLL site, it is possible that multiple copies of the monoacid **2** could bind in the same region on the protein. Y631, which is featured in the MLL site, was the most sensitive reporter for binding and was therefore used to calculate the K_d for **2**, **3**, and **4**. In contrast, Y649, Y650, and Y658, which are featured in the CREB binding site, were significantly less responsive to ligand binding with the exception of **1**, which perturbed resonances

Table 1: Small-molecule ligands discovered by PrOF NMR spectroscopy.

Compound		^1H–^{15}N HSQC	K_d (PrOF) [mM]	Ligand efficiency [kcal mol^{-1} NHA^{-1}]
1		yes	>10	>0.13
2		yes	1.4±0.3[a]	0.32
3		yes	5.3±0.9	0.24
4		yes	1.7±0.6	0.29
5		n.d.	1.6±0.3	0.27
PAA5		n.d.	6.9±0.7	0.18
Flurbiprofen		n.d.	5.9±0.6	0.17

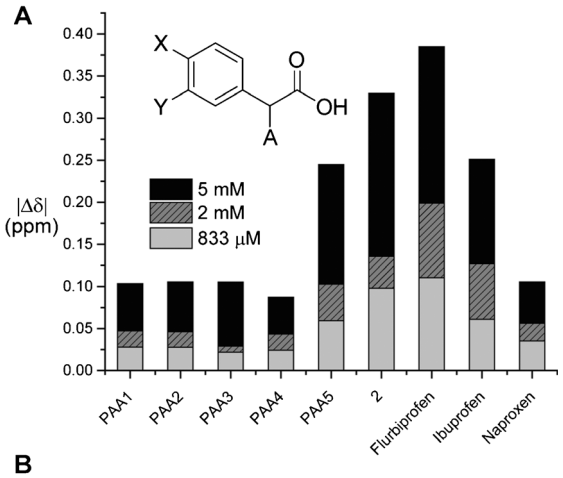
[a] K_d estimate based on data fit up to 2 mM, not accounting for higher-order binding at greater concentrations, n.d. = not determined.

for Y649, Y650, and Y631, potentially indicating binding at the CREB site.

As a follow-up to the screen, the SARs from the four hits were examined to identify structural motifs important for binding. Four trifluoromethylquinoline derivatives of **1** were purchased. These compounds were of interest due to small but significant chemical-shift perturbations from **1**, which were similar to the chemical shift effects observed upon CREB binding to KIX suggesting a possible allosteric binding mode.[4a] One of the four compounds (KM07508) yielded comparable chemical-shift perturbations as **1**.[18] However, as **1** displayed poor binding affinity in titration experiments (K_d > 10 mM), due to the lack of saturated binding, and in 2D NMR experimental results explained below, SAR studies with this ligand class were not pursued further in lieu of the investigation into molecules **2**–**4**.

The other three discovered ligands contained an aryl or phenylacetic acid motif as a potential pharmacophore. The presence of a carboxylic acid has been found in other KIX ligands that bind to the MLL site (sekikaic acid, lobaric acid, KG-122, and 1G7, Figure 1). To evaluate the specificity over other acidic molecules and to probe for false negatives, several compounds from library mixtures were retested. One of the four compounds (9B11, **5**), a biaryl carboxylic acid, was anticipated to have been a false negative due to its similar structure to KIX ligand 1G7. The other three carboxylic acids were selected to test if these aryl and phenylacetic acids represented a class of ligands that bound KIX, whereas other carboxylic acids would not (i.e., a phenethyl, a thiazole, and an aliphatic carboxylic acid). Only **5** yielded significant changes in chemical shift, whereas the other three exhibited minimal to no apparent binding. Titration of **5** yielded a K_d of 1.6±0.3 mM (Table 1). Low mM K_d values and ligand efficiencies from 0.19 to 0.32 kcal mol^{-1} non-hydrogen atom^{-1} were calculated for each fragment. As a comparison, the reported median ligand efficiency was 0.32 for a sample set of 480 ligand–target pairs.[21] Our ligand efficiencies are equal to or slightly lower than this value, supporting the difficulty of this target site.[12]

To expand on the investigation of aryl + acid SAR, nine arylacetic and benzoic acid derivatives were tested as analogues of molecule **2** to inform optimization studies, based on the affinity of **2** and the readily modifiable structure of phenylacetic acids. Of the nine compounds, only the biphenyl (PAA5) yielded a pronounced chemical shift perturbation close to one standard deviation above the average at 5 mM,[18] despite the structural similarity of these analogues to **2**, e.g., 3-chloro-4-hydroxy phenylacetic acid (PAA2), phenylacetic acid (PAA3), and 4-hydroxyphenylacetic acid (PAA4) (Figure 3). Each analogue was also



B

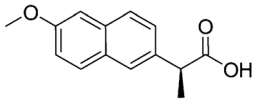
	A	X	Y
PAA1	=O	H	H
PAA2	H	OH	Cl
PAA3	H	H	H
PAA4	H	OH	H
PAA5	H	C_6H_5	H
2	H	Cl	Cl
Flurbiprofen	CH_3	C_6H_5	F
Ibuprofen	CH_3	$CH_2CH(CH_3)_2$	H
Naproxen			

Figure 3. Representative SAR analysis of phenylacetic acids. A) Chemical-shift perturbation plot at three different concentrations of small molecule. B) Phenylacetic acid structures.

analyzed at 2 mM and 833 μM, due to the potential for higher-order binding at high concentrations (>2 mM). The trend remained consistent with all compounds aside from PAA5 yielding insignificant changes in chemical shift. The K_d for PAA5 binding to KIX was 6.9±0.7 mM. These results suggest that hydrophobic groups at the *para* position on the phenylacetic acid play a key role in ligand binding.

Phenylacetic acids are a common scaffold found in several nonsteroidal anti-inflammatory drugs. As an additional analysis of SAR, naproxen, ibuprofen, and flurbiprofen were used to further support the model that *para*-substituted

phenylacetic acid derivatives would bind to KIX. All three compounds were analyzed by PrOF NMR spectroscopy. The chemical-shift perturbations from enantiopure naproxen were under the statistical cutoff at low concentrations. Racemic ibuprofen yielded significant chemical-shift perturbations, whereas racemic flurbiprofen yielded the largest $\Delta\delta$ for Y631 observed thus far (0.38 ppm at 5 mM, $K_d = 5.9 \pm 0.6$ mM, Figure 3). Substitution at the α-position affords an additional site of diversity and both stereochemistry and elaboration of side chains can be investigated as starting points for structure activity studies.

As a complementary method to evaluate the ligand binding site on KIX and rule out concerns of generating false positives from fluorine incorporation, we studied the binding of the original four ligands (**1**–**4**) by ^{1}H–^{15}N HSQC NMR spectroscopy using uniformly ^{15}N-labeled KIX. Experiments were performed at 0, 1, and 5 mM ligand concentration for **1**, **2**, and **3**, whereas **4** was characterized at 0, 0.5, and 1 mM due to low solubility above 2 mM. Residues whose chemical shifts were perturbed by one and two standard deviations above the average perturbation in a residue were mapped onto the protein structure to infer the corresponding binding site of each ligand. Consistent with PrOF NMR measurements, perturbed residues in each case showed concentration-dependent shifts and localized near the MLL binding site.[18] Additional residues were perturbed near an unstructured loop connecting helices α2 and α3 of KIX near the N-terminal portion of the protein. These perturbations may result from a conformational change induced by ligand in the MLL site rather than direct ligand binding near this loop region but will need further investigation. We conclude from this study that use of PrOF NMR led to the identification of a related class of ligands all targeting the MLL site.

Despite few hits, the hit rates from the PrOF NMR screen could be used to assess the "druggability" of the MLL and CREB sites on KIX. Consistent with the literature, the CREB interface is a challenging target, whereas the MLL site is more amenable to ligand binding. For example, disulfide tethering screens reported fewer and less efficacious ligands for the CREB versus the MLL site.[22] These results support designing larger peptidomimetic inhibitors against the CREB site rather than using small molecules.[23] Both naturally occurring acidic transcription factors as well as synthetic ligands (Figure 1) have a higher propensity for interacting at or near the MLL region, providing a starting point for MLL-site ligand development. The KIX-interacting peptide of MLL contains the sequence DIMDFVL in which both aspartic acid residues and phenylalanine are important for binding.[11c] The aryl and phenylacetic acid derivatives reported here, may be minimal mimics of these motifs.

PrOF NMR spectroscopy has several advantages as a screening method due to its sensitivity to detect changes in chemical environment including dynamic structural information of the protein, low background signal, and ease of use. Here we demonstrated the first full small-molecule screen using PrOF NMR spectroscopy, evaluated its efficacy for druggability determination, and reported on a class of small molecules (biaryl and phenylacetic acids) as KIX ligands. SAR analysis guided our investigation to discover three additional KIX-binding small molecules. The biaryl acid and phenylacetic acid structural motif is corroborated by several of the reported KIX ligands (Figure 1) providing a starting point for small-molecule derivatization and development. This study also demonstrates preferential targetability for the MLL site over the CREB site. Although additional optimization is required to develop the discovered ligands into potent inhibitors of KIX interactions, this study has demonstrated the applicability of PrOF NMR spectroscopy as a useful tool for library screening, ligand discovery, and druggability assessment.

Received: December 3, 2014
Published online: February 4, 2015

Keywords: biomolecular NMR · fluorine · fragment screening · protein–protein interactions · transcription

[1] a) A. J. Souers, J. D. Leverson, E. R. Boghaert, S. L. Ackler, N. D. Catron, J. Chen, B. D. Dayton, H. Ding, S. H. Enschede, W. J. Fairbrother, et al., *Nat. Med.* **2013**, *19*, 202–208; b) O. Mirguet, R. Gosmini, J. Toum, C. A. Clément, M. Barnathan, J.-M. Brusq, J. E. Mordaunt, R. M. Grimes, M. Crowe, O. Pineau, et al., *J. Med. Chem.* **2013**, *56*, 7501–7515.

[2] M. R. Arkin, Y. Tang, J. A. Wells, *Chem. Biol.* **2014**, *21*, 1102–1114.

[3] A. D. Thompson, A. Dugan, J. E. Gestwicki, A. K. Mapp, *ACS Chem. Biol.* **2012**, *7*, 1311–1320.

[4] a) W. C. Pomerantz, N. Wang, A. K. Lipinski, R. Wang, T. Cierpicki, A. K. Mapp, *ACS Chem. Biol.* **2012**, *7*, 1345–1350; b) N. K. Mishra, A. K. Urick, S. W. Ember, E. Schonbrunn, W. C. Pomerantz, *ACS Chem. Biol.* **2014**, *9*, 2755–2760.

[5] a) A. A. Bogan, K. S. Thorn, *J. Mol. Biol.* **1998**, *280*, 1–9; b) A. M. Watkins, P. S. Arora, *ACS Chem. Biol.* **2014**, *9*, 1747–1754.

[6] a) J. L. Kitevski-LeBlanc, R. S. Prosser, *Prog. Nucl. Magn. Reson. Spectrosc.* **2012**, *62*, 1–33; b) E. N. Marsh, Y. Suzuki, *ACS Chem. Biol.* **2014**, *9*, 1242–1250; c) X. Ge, C. A. MacRaild, S. M. Devine, C. O. Debono, G. Wang, P. J. Scammells, M. J. Scanlon, R. F. Anders, M. Foley, R. Norton, *J. Med. Chem.* **2014**, *57*, 6419–6427.

[7] P. J. Hajduk, J. Greer, *Nat. Rev. Drug Discovery* **2007**, *6*, 211–219.

[8] a) C. A. Lepre, *Methods Enzymol.* **2011**, *493*, 219–239; b) B. Wu, Z. Zhang, R. Noberini, E. Barile, M. Giulianotti, C. Pinilla, R. A. Houghten, E. B. Pasquale, M. Pellecchia, *Chem. Biol.* **2013**, *20*, 19–33; c) C. Dalvit, P. E. Fagerness, D. T. Hadden, R. W. Sarver, B. J. Stockman, *J. Am. Chem. Soc.* **2003**, *125*, 7696–7703.

[9] P. J. Hajduk, J. R. Huth, S. W. Fesik, *J. Med. Chem.* **2005**, *48*, 2518–2525.

[10] G. Siegal, E. Ab, J. Schultz, *Drug Discovery Today* **2007**, *12*, 1032–1039.

[11] a) I. Radhakrishnan, G. C. Pérez-Alvarado, D. Parker, H. J. Dyson, M. R. Montminy, P. E. Wright, *Cell* **1997**, *91*, 741–752; b) C. M. Denis, S. Chitayat, M. J. Plevin, F. Wang, P. Thompson, S. Liu, H. L. Spencer, M. Ikura, D. P. LeBrun, S. P. Smith, *Blood* **2012**, *120*, 3968–3977; c) P. Ernst, J. Wang, M. Huang, R. H. Goodman, S. J. Korsmeyer, *Mol. Cell. Biol.* **2001**, *21*, 2249–2258.

[12] C. Y. Majmudar, J. W. Højfeldt, C. J. Arevang, W. C. Pomerantz, J. K. Gagnon, P. J. Schultz, L. C. Cesa, C. H. Doss, S. P. Rowe, V. Vásquez, G. Tamayo-Castillo, T. Cierpicki, C. L. Brooks, D. H.

Sherman, A. K. Mapp, *Angew. Chem. Int. Ed.* **2012**, *51*, 11258 – 11262; *Angew. Chem.* **2012**, *124*, 11420 – 11424.
[13] a) J. L. Best, C. A. Amezcua, B. Mayr, L. Flechner, C. M. Murawsky, B. Emerson, T. Zor, K. H. Gardner, M. Montminy, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 17622 – 17627; b) Further studies indicate that the active form of Naphthol AS-E phosphate is its dephosphorylated form. B. X. Li, X. Xiao, *ChemBioChem* **2009**, *10*, 2721 – 2724.
[14] S. P. Rowe, R. J. Casey, B. B. Brennan, S. J. Buhrlage, A. K. Mapp, *J. Am. Chem. Soc.* **2007**, *129*, 10654 – 10655.
[15] S. J. Buhrlage, C. A. Bates, S. P. Rowe, A. R. Minter, B. B. Brennan, C. Y. Majmudar, D. E. Wemmer, H. Al-Hashimi, A. K. Mapp, *ACS Chem. Biol.* **2009**, *4*, 335 – 344.
[16] D. B. Shankar, J. C. Cheng, K. M. Sakamoto, *Cancer* **2005**, *104*, 1819 – 1824.
[17] F. Yang, B. W. Vought, J. S. Satterlee, A. K. Walker, Z. Y. Jim Sun, J. L. Watts, R. DeBeaumont, R. M. Saito, S. G. Hyberts, S. Yang, C. Macol, L. Iyer, R. Tjian, S. van den Heuvel, A. C. Hart, G. Wagner, A. M. Naar, *Nature* **2006**, *442*, 700 – 704.
[18] Please see the Supporting Information.
[19] Y. Wei, J.-C. Horng, A. C. Vendel, D. P. Raleigh, K. J. Lumb, *Biochemistry* **2003**, *42*, 7044 – 7049.
[20] M. J. Harner, B. A. Chauder, J. Phan, S. W. Fesik, *J. Med. Chem.* **2014**, *57*, 9687 – 9692.
[21] A. L. Hopkins, G. M. Keseru, P. D. Leeson, D. C. Rees, C. H. Reynolds, *Nat. Rev. Drug Discovery* **2014**, *13*, 105 – 121.
[22] J. M. Lodge, T. Justin Rettenmaier, J. A. Wells, W. C. Pomerantz, A. K. Mapp, *MedChemComm* **2014**, *5*, 370 – 375.
[23] B. N. Bullock, A. L. Jochim, P. S. Arora, *J. Am. Chem. Soc.* **2011**, *133*, 14220 – 14223.

