JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Ordering a Dynamic Protein Via a Small-Molecule Stabilizer

Ningkun Wang,[†] Chinmay Y. Majmudar,[‡] William C. Pomerantz,[‡] Jessica K. Gagnon,[‡] Jack D. Sadowsky,[§] Jennifer L. Meagher,[∥] Taylor K. Johnson,[⊥] Jeanne A. Stuckey,[∥] Charles L. Brooks III,[†,‡,⊥] James A. Wells,[§] and Anna K. Mapp*[,†,‡,⊥]

[†]Program in Chemical Biology, [‡]Department of Chemistry, [∥]Life Sciences Institute, and [⊥]Department of Medicinal Chemistry, University of Michigan, Ann Arbor, Michigan 48109, United States

[§]Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94158, United States

Supporting Information

ABSTRACT: Like many coactivators, the GACKIX domain of the master coactivator CBP/p300 recognizes transcriptional activators of diverse sequence composition via dynamic binding surfaces. The conformational dynamics of GACKIX that underlie its function also render it especially challenging for structural characterization. We have found that the ligand discovery strategy of Tethering is an effective method for identifying small-molecule fragments that stabilize the GACKIX domain, enabling for the first time the crystallographic characterization of this important motif. The 2.0 Å resolution structure of GACKIX complexed to a small molecule was further analyzed by molecular dynamics simulations, which revealed the importance of specific side-chain motions that remodel the activator binding site in order to accommodate binding partners of distinct sequence and size. More broadly, these results suggest that Tethering can be a powerful strategy for identifying small-molecule stabilizers of conformationally malleable proteins, thus facilitating their structural characterization and accelerating the discovery of small-molecule modulators.

Transcriptional coactivators are among the most conformationally malleable of proteins and contain binding surfaces that undergo rapid remodeling as complexes are formed with their cognate ligands.[1,2] This plasticity is essential to their function, enabling recognition of an often diverse array of transcriptional activator sequences.[3,4] Perhaps the best-studied example of this is the GACKIX domain of the coactivator CBP/p300, a small (90 amino acid) domain that is known to interact with >10 distinct amphipathic sequences at two distinct binding sites (Figure 1a) in order to stimulate transcription at hundreds of genes,[5−9] including those regulating hematopoiesis, memory formation, and the inflammatory response.[10−12] Not surprisingly, the malleability of this class of proteins renders them especially intractable to crystallographic characterization, either alone or in complex with their binding partners. In the case of the GACKIX domain, there are no crystal structures of either the free protein or any complexed form. Here we report that a covalently linked small-molecule ligand of this conformationally dynamic protein has enabled a high-resolution snapshot of the coactivator interacting with a ligand to be obtained for the first time. This first crystal structure of GACKIX provides important insight to the side-chain orientations of this domain in the context of ligand recognition, particularly with regard to small molecules. Furthermore, these results show that the ligand discovery strategy of Tethering[13−16] can be expanded to target conformationally dynamic proteins and enable their structural characterization.

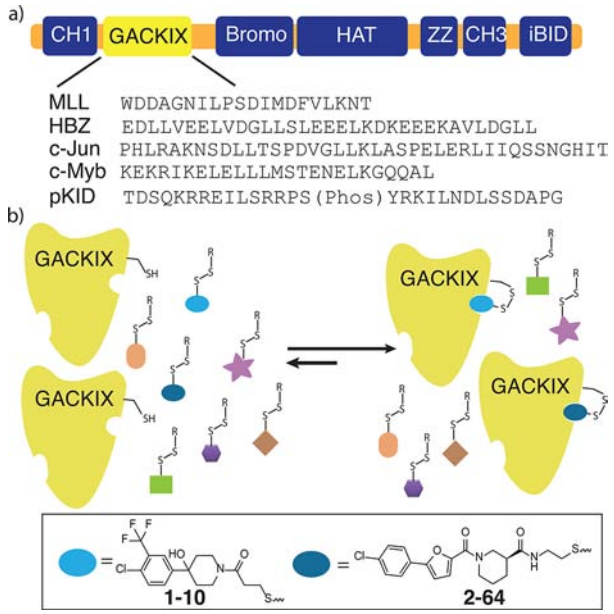

Figure 1. (a) The GACKIX domain is in the N-terminal region of CBP/p300. GACKIX interacts with >10 amphipathic transcriptional activators using two distinct sites.[5−9] MLL, HBZ, and c-Jun target a smaller, deeper site, while the activation domains of c-Myb and CREB (pKID) utilize a second, broader site. (b) Schematic of the Tethering screen used to identify small-molecule fragments (**1-10** and **2-64**) that form a disulfide bond with a cysteine introduced at position 664 (L664C) within GACKIX. See the SI for details.

We screened for small molecules that interact with the GACKIX domain using the Tethering approach,[16] a strategy that provides a mechanism for the rapid discovery of covalent ligands (Figure 1b). Attention was focused on the binding site that is targeted by the transcriptional activation domains of proteins such as the mixed-lineage leukemia (MLL) activator

Received: December 14, 2012
Published: February 5, 2013

   dx.doi.org/10.1021/ja3122334 | J. Am. Chem. Soc. 2013, 135, 3363−3366

and c-Jun; the Tethering approach is a fragment discovery method, and the smaller, deeper MLL/c-Jun binding site appeared to be more targetable by low-molecular-weight compounds.[17,18] Toward this end, a residue at the rim of the binding surface, L664, was mutated to a cysteine, and the resulting GACKIX L664C mutant was fully characterized [see the Supporting Information (SI) for details]. Small-molecule fragments containing a disulfide motif were then screened for the ability to form a disulfide bond with GACKIX L664C in the presence of a competitor, β-mercaptoethanol. Two fragment ligands, denoted as **1-10** and **2-64** (Figure 1b), emerged from the screen with high Tethering efficiency to GACKIX L664C, as quantified by 50% dose response (DR_{50}) values (2−8 μM).

To assess the effect of tethered **1-10** or **2-64** on the binding properties of GACKIX, fluorescence anisotropy binding assays were used to measure the binding affinity of wild-type GACKIX, GACKIX L664C, and fragment-tethered GACKIX L664C complexes to native transcriptional activator ligands that target the two different binding sites (Figure 2a). Consistent

a)

GACKIX construct	K_D (μM) MLL	pKID
GACKIX wt	0.5 ± 0.1	0.9 ± 0.1
GACKIX L664C	2.8 ± 0.3	0.8 ± 0.1
1-10—GACKIX L664C	62 ± 4	1.9 ± 0.2
2-64—GACKIX L664C	94 ± 7	1.1 ± 0.1

b)

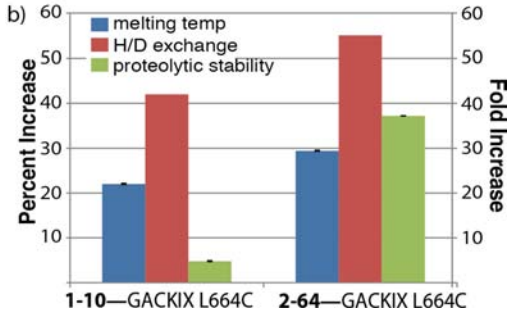

Figure 2. (a) K_D values for GACKIX constructs interacting with fluorescein-labeled MLL and pKID peptides were determined by fluorescence anisotropy assays. Each K_D is a fitted result of experiments performed in triplicate with the indicated error (standard deviation). (b) Bar graph depicting the percent increase in melting temperature (T_M) as monitored by circular dichroism (blue bars), the percentage of backbone amides protected from H−D exchange (red bars), and the fold increase in resistance to thermolysin degradation of the GACKIX mutants (green bars) upon tethering to (left) **1-10** and (right) **2-64**. The data have been normalized to GACKIX L664C.

with the screen design, the presence of **1-10** or **2-64** decreased MLL binding to GACKIX L664C by ∼22 to 33-fold. Also, while tethered **2-64** did not affect GACKIX's binding affinity for pKID, the transcriptional activation domain of CREB that interacts with the distal binding site,[19] GACKIX tethered to fragment **1-10** did exhibit ∼2-fold attenuated binding to pKID. This suggests that **1-10** engages the amino acid side chains constituting the allosteric network connecting the two binding sites.[6,20,21]

The tethered fragments significantly altered the stability of the GACKIX domain. This was assessed for each of the fragment−protein pairs by measuring changes in CD-monitored thermal melting temperature (T_M), amide hydrogen−deuterium (H−D) exchange, and thermolysin-mediated proteolysis (Figure 2b). For example, the **1-10**−GACKIX L664C and **2-64**−GACKIX L664C complexes exhibited a 15−18 °C (≥20%) increase in T_M. In the H−D exchange experiment, the mass of free GACKIX L664C shifted by 29 Da upon exposure to D_2O for 1 min as monitored by mass spectrometry,[13] whereas the mass shifts were 17 and 13 Da when **1-10** and **2-64**, respectively, were tethered to GACKIX L664C, showing that 40−55% of the exchangeable amides were protected from H−D exchange compared with the free protein. The proteolytic stability of the tethered complex (as indicated by the half-life, $t_{1/2}$) increased 5−37 fold relative to the untethered protein (e.g., for **1-10**, $t_{1/2}$ = 10 min vs 2.1 min).[22,23] These findings encouraged us to pursue crystallization of fragment−GACKIX L664C complexes.

Of the various fragment−protein complexes and conditions that were screened, the best results were obtained with **1-10**−GACKIX L664C under the crystallizing conditions of 1.8 M ammonium sulfate and 0.1 M Tris (pH 7.0) at 25 °C, which gave crystals amenable to diffraction. However, only microcrystals of **2-64** tethered to GACKIX L664C were obtained and were of too poor quality for the structure to be solved. Initially, molecular replacement strategies using the NMR structures of GACKIX bound to native transcriptional activation domains were used, but they did not lead to the **1-10**−GACKIX structure.[5,19] Therefore, a selenomethionine-incorporated GACKIX L664C tethered to **1-10** was prepared, and the X-ray structure was solved. From these data, the structure of **1-10**−GACKIX L664C was determined to 2.0 Å resolution.

As illustrated in Figure 3a, the small molecule **1-10** sits within the MLL/c-Jun binding site of GACKIX and is oriented toward the core of the protein between helices α_3 (residues 646−664) and α_2 (residues 623−638). Notably, the aromatic ring of **1-10** is positioned relatively deep in a hydrophobic pocket lined by the side chains of Ile611, Leu628, Leu607, Val635, and Tyr631 (Figure 3b); Leu628 and Tyr631 have previously been shown to be key residues involved in the interaction of GACKIX with MLL.[5,24] In particular, Tyr631 closely contacts the aromatic ring of **1-10** (∼4 Å), as illustrated by the above 2σ deviation of the Tyr631 φ and ψ angles.[25] Consistent with these data, chemical shift perturbation experiments with ^{15}N-labeled KIX L664C (free vs covalently tethered to **1-10**) revealed significant changes in the backbone amide shifts of the residues lining the hydrophobic binding surface (Ile611, Leu628, Leu607, Val635, Tyr631, and Ile660) upon tethering to **1-10** (Figure 4a).

The prevailing structural model of the amphipathic class of activator−coactivator complexes is that the activator forms an amphipathic helix upon binding to the surface of the coactivator.[1,26,27] Although only a limited suite of surfaces have been characterized, the available data suggest that the binding surfaces are often broad,[2,28] making them particularly challenging to target with small molecules that have far less volume and surface area than the typical helix of a transcriptional activator.[29] Overlays of the **1-10**−GACKIX L664C structure with the averages of the previously reported NMR structures of GACKIX−ligand complexes[5,30,31] yielded root-mean-square deviation (rmsd) values of 1.07−1.81 Å, demonstrating the overall similarity of the backbone structures. The exception to this similarity is in the loop region between helices α_1 and α_2 (residues 612−622), which deviates significantly (rmsd = 2.73−3.11 Å). This difference is not surprising, as conformational changes in the loop regions are

dx.doi.org/10.1021/ja3122334 | J. Am. Chem. Soc. 2013, 135, 3363−3366

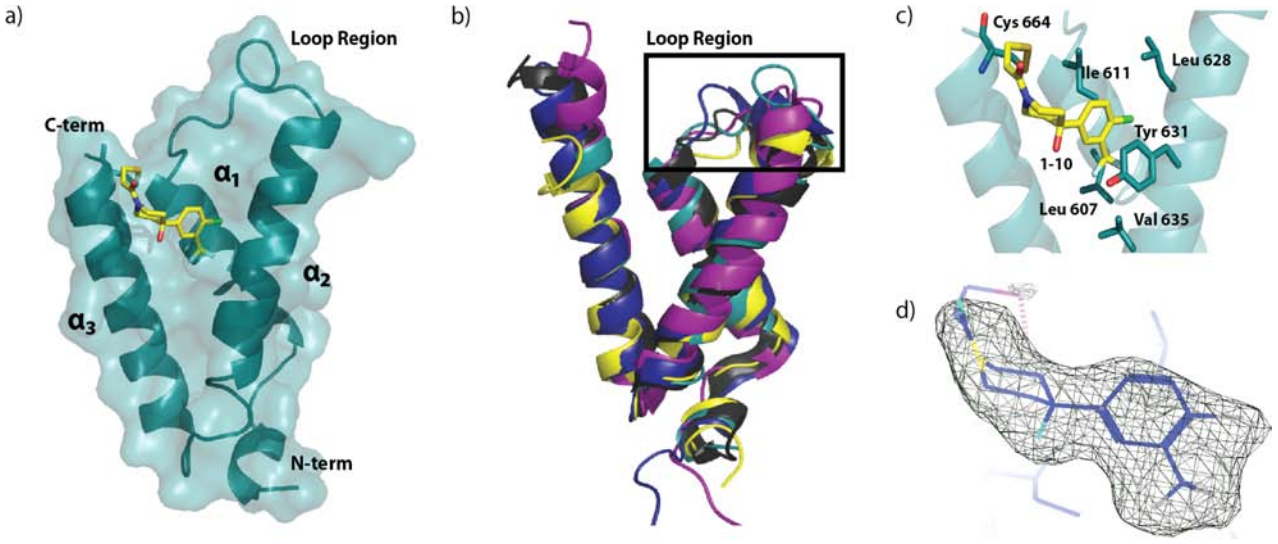

Figure 3. (a) Refined crystal structure of GACKIX L664C covalently tethered to fragment **1-10** (refined resolution = 2.0 Å, R_{work}/R_{free} = 0.2064/0.2329). (b) Crystal structure of GACKIX L664C tethered to **1-10** (teal) superimposed using Coot on the NMR solution structures of GACKIX in complex with cognate transcriptional activation domains pKID (yellow, PDB ID 1KDX, rmsd = 1.40 Å), MLL and c-Myb (deep blue, PDB ID 2AGH, rmsd = 1.80 Å), PCET (purple, PDB ID 2KWF, rmsd = 1.81 Å), and FOXO3A (black, PDB ID 2LQH, rmsd = 1.07 Å). (c) Interactions between **1-10** (yellow) and residue side chains of GACKIX L664C (blue) at the binding surface. (d) 3σ electron density map ($F_o - F_c$) of **1-10** illustrating the fit of the small molecule.

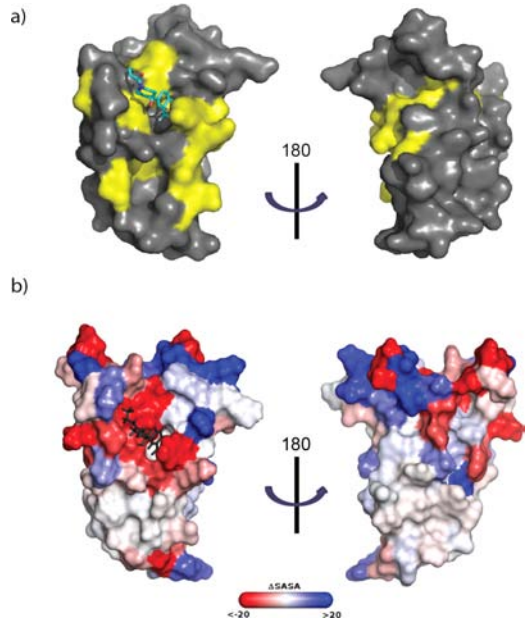

Figure 4. (a) Results of chemical shift perturbation experiment (^{1}H,^{15}N-HSQC) with **1-10**-tethered GACKIX L664C. Residues that shifted by more than 1 standard deviation upon **1-10** tethering are shown in yellow and include Ile611, Leu628, Leu607, Val635, Tyr631, and Ile660. (b) Difference in the average solvent-accessible surface area (SASA, in Å^2) calculated by residue between simulations of untethered and **1-10**-tethered GACKIX L664C. A residue colored red is less solvent-exposed in the **1-10**-tethered structure, with the color intensity indicating the extent of the change; blue residues are more solvent-exposed in the **1-10**-tethered structure.

thought to be integral to the ability of GACKIX to accommodate diverse native ligands.[5,19,21,30]

To dissect in more detail how the GACKIX surface remodels itself to recognize fragment **1-10**, we carried out 40 ns molecular dynamics simulations of the GACKIX crystal structure with or without ligand **1-10**. A gross comparison of the backbone revealed that a change in the loop conformation is the most significant, as shown in the rms fluctuations (Figure S6 in the SI) and in the average structure overlay (Figure S7). These changes are often difficult to visualize by solution methods because the loop region contains several proline residues, but mutagenesis and NMR methods have suggested that conformational plasticity in this region underlies the ability of GACKIX to recognize diverse amphipathic sequences.[5,20,21] It is this movement of the loop and a rotation of helix α_1 that enable the formation of a narrower binding surface to accommodate a molecule that is considerably smaller than a peptidic helix (~77% smaller volume). The binding surface that is targeted by **1-10** is also significantly different, both as a result of loop conformational changes and because of side-chain motions, as demonstrated by the change in solvent-accessible surface area of the residues when the fragment is tethered (Figure 4b). For example, the liganded GACKIX shows a population shift in the Tyr631 side-chain χ angles relative to the untethered protein, leading to a hydrophobic binding surface for deeper interactions (see movie S1 in the SI). Simulations of **2-64** tethered to GACKIX L664C suggested that the binding mode of this ligand is similar to that of **1-10**, further demonstrating the ability of this protein to adapt to different binding partners (Figure S8). The helices α_3 and α_2 must open to accommodate this larger ligand, and corresponding changes in the chemical shifts of residues involved in this opening were observed by NMR spectroscopy (Figures S5 and S7).

In conclusion, we have obtained a 2 Å-resolution snapshot of the conformationally dynamic coactivator GACKIX domain complexed with a small molecule. This will significantly facilitate the use of rational structure-based approaches to design more potent analogues; for example, current efforts include extending molecule **1-10** at the C4 position of the aromatic ring to enable it to engage with the hydrophobic space within the GACKIX site more effectively. From a broader perspective, these results in combination with recent studies

dx.doi.org/10.1021/ja3122334 | J. Am. Chem. Soc. 2013, 135, 3363−3366

showing stabilization of conformationally dynamic proteins by noncovalent interactions with small molecules[32,33] suggest that Tethering may be an exceptionally enabling approach for obtaining long-sought X-ray crystallographic data for conformationally dynamic proteins. This includes not only transcriptional coactivators such as CBP/p300 targeted here but also members of other cellular machines that rely upon conformationally dynamic interfaces to recognize binding partners.[34,35]

■ ASSOCIATED CONTENT

Supporting Information

Details of the Tethering screen, additional structural details, crystallographic data (CIF), and a movie (AVI). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

amapp@umich.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

A.K.M. is grateful for support of this work from NIH 2R01 GM65330 and R01 CA140667. W.C.P. acknowledges NIH F32 GM090550, and N.W. is a fellow of the UM PSTP (NIGMS GM07767). C.L.B. acknowledges NIH GM037554. J.A.W. is grateful for support of this work from NIH R01 AI070292. J.D.S. is a fellow of California TRDRP (Award 110385). Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor for the support of this research program (Grant 085P1000817). We thank Dr. David Smith of LS-CAT for help with remote data collection. Coordinates for the **1-10**–GACKIX L664C complex have been deposited in the PDB (4I9O).

■ REFERENCES

(1) Sugase, K.; Dyson, H. J.; Wright, P. E. *Nature* **2007**, *447*, 1021.
(2) Thompson, A. D.; Dugan, A.; Gestwicki, J. E.; Mapp, A. K. *ACS Chem. Biol.* **2012**, *7*, 1311.
(3) Singh, G. P.; Ganapathi, M.; Dash, D. *Proteins* **2007**, *66*, 761.
(4) Boehr, D. D.; Nussinov, R.; Wright, P. E. *Nat. Chem. Biol.* **2009**, *5*, 789.
(5) De Guzman, R. N.; Goto, N. K.; Dyson, H. J.; Wright, P. E. *J. Mol. Biol.* **2006**, *355*, 1005.
(6) Goto, N. K.; Zor, T.; Martinez-Yamout, M.; Dyson, H. J.; Wright, P. E. *J. Biol. Chem.* **2002**, *277*, 43168.
(7) Ernst, P.; Wang, J.; Huang, M.; Goodman, R. H.; Korsmeyer, S. J. *Mol. Cell. Biol.* **2001**, *21*, 2249.
(8) Cook, P. R.; Polakowski, N.; Lemasson, I. *J. Mol. Biol.* **2011**, *409*, 384.
(9) Vendel, A. C.; McBryant, S. J.; Lumb, K. J. *Biochemistry* **2003**, *42*, 12481.
(10) Kasper, L. H.; Boussouar, F.; Ney, P. A.; Jackson, C. W.; Rehg, J.; van Deursen, J. M.; Brindle, P. K. *Nature* **2002**, *419*, 738.
(11) Wood, M. A.; Attner, M. A.; Oliveira, A. M. M.; Brindle, P. K.; Abel, T. *Learn. Mem.* **2006**, *13*, 609.
(12) Liu, G.-H.; Qu, J.; Shen, X. *Biochim. Biophys. Acta* **2008**, *1783*, 713.
(13) Sadowsky, J. D.; Burlingame, M. A.; Wolan, D. W.; McClendon, C. L.; Jacobson, M. P.; Wells, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 6056.
(14) Hardy, J. A.; Lam, J.; Nguyen, J. T.; O'Brien, T.; Wells, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12461.
(15) Scheer, J. M.; Romanowski, M. J.; Wells, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 7595.
(16) Erlanson, D. A.; Braisted, A. C.; Raphael, D. R.; Randal, M.; Stroud, R. M.; Gordon, E. M.; Wells, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9367.
(17) Two natural products that target the MLL site have recently been identified. See: Majmudar, C. Y.; Højfeldt, J. W.; Arevang, C. J.; Pomerantz, W. C.; Gagnon, J. K.; Schultz, P. J.; Cesa, L. C.; Doss, C. H.; Rowe, S. P.; Vásquez, V.; Tamayo-Castillo, G.; Cierpicki, T.; Brooks, C. L., III; Sherman, D. H.; Mapp, A. K. *Angew. Chem., Int. Ed.* **2012**, *51*, 11258.
(18) A small pilot screen using protein observed by ^{19}F NMR spectroscopy identified a small-molecule fragment that interacts with the MLL site. See: Pomerantz, W. C.; Wang, N.; Lipinski, A. K.; Wang, R.; Cierpicki, T.; Mapp, A. K. *ACS Chem. Biol.* **2012**, *7*, 1345.
(19) Radhakrishnan, I.; Perezalvarado, G.; Parker, D.; Dyson, H.; Montminy, M.; Wright, P. *Cell* **1997**, *91*, 741.
(20) Brüschweiler, S.; Schanda, P.; Kloiber, K.; Brutscher, B.; Kontaxis, G.; Konrat, R.; Tollinger, M. *J. Am. Chem. Soc.* **2009**, *131*, 3063.
(21) Korkmaz, E. N.; Nussinov, R.; Haliloğlu, T. *PLoS Comput. Biol.* **2012**, *8*, No. e1002420.
(22) Park, C.; Marqusee, S. *Nat. Methods* **2005**, *2*, 207.
(23) Hoofnagle, A. N.; Resing, K. A.; Ahn, N. G. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 1.
(24) Arai, M.; Dyson, H. J.; Wright, P. E. *FEBS Lett.* **2010**, *584*, 4500.
(25) This is consistent with data from a solution binding study of untethered **1-10** interacting with GACKIX containing ^{19}F-labeled Tyr631,[18] which showed a dose-dependent change in the ^{19}F chemical shift.
(26) Uesugi, M.; Nyanguile, O.; Lu, H.; Levine, A. J.; Verdine, G. L. *Science* **1997**, *277*, 1310.
(27) Thoden, J. B.; Ryan, L. A.; Reece, R. J.; Holden, H. M. *J. Biol. Chem.* **2008**, *283*, 30266.
(28) Fuller, J. C.; Burgoyne, N. J.; Jackson, R. M. *Drug Discovery Today* **2009**, *14*, 155.
(29) Fry, D. *Pept. Sci.* **2006**, *84*, 535.
(30) Wang, F.; Marshall, C. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 6078.
(31) Zor, T.; De Guzman, R. N.; Dyson, H. J.; Wright, P. E. *J. Mol. Biol.* **2004**, *337*, 521.
(32) Saalau-Bethell, S. M.; Woodhead, A. J.; Chessari, G.; Carr, M. G.; Coyle, J.; Graham, B.; Hiscock, S. D.; Murray, C. W.; Pathuri, P.; Rich, S. J.; Richardson, C. J.; Williams, P. A.; Jhoti, H. *Nat. Chem. Biol.* **2012**, *8*, 920.
(33) Basse, N.; Kaar, J. L.; Settanni, G.; Joerger, A. C.; Rutherford, T. J.; Fersht, A. R. *Chem. Biol.* **2010**, *17*, 46.
(34) Dunker, A. K.; Silman, I.; Uversky, V. N.; Sussman, J. L. *Curr. Opin. Struct. Biol.* **2008**, *18*, 756.
(35) Dyson, H. J.; Wright, P. E. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197.