JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
COMMUNICATION
pubs.acs.org/JACS

Assessing Helical Protein Interfaces for Inhibitor Design

Brooke N. Bullock, Andrea L. Jochim, and Paramjit S. Arora*

Department of Chemistry, New York University, New York, New York 10003, United States

S *Supporting Information*

ABSTRACT: Structure-based design of synthetic inhibitors of protein−protein interactions (PPIs) requires adept molecular design and synthesis strategies as well as knowledge of targetable complexes. To address the significant gap between the elegant design of helix mimetics and their sporadic use in biology, we analyzed the full set of helical protein interfaces in the Protein Data Bank to obtain a snapshot of how helices that are critical for complex formation interact with the partner proteins. The results of this study are expected to guide the systematic design of synthetic inhibitors of PPIs. We have experimentally evaluated new classes of protein complexes that emerged from this data set, highlighting the significance of the results described herein.

Interactions of proteins with partner proteins control essential cellular processes, and misregulation of these interactions is often implicated in disease states.[1] However, despite their fundamental role, protein−protein interactions (PPIs) are generally not considered attractive targets for drug design because of their large and often flat contact surfaces.[2−4] A promising rational design approach for the discovery of PPI inhibitors is centered on the role of protein secondary structures at protein interfaces. Analysis suggests that although protein interfaces are large, often a small subset of the residues contributes significantly to the free energy of binding.[5−8] Secondary structures are common scaffolds for the organization of these "hot spots" in proteins.[4,9,10] It has been demonstrated that synthetic molecules that reproduce key elements of energetically significant protein secondary structures can inhibit chosen interfaces with high affinity and specificity.[11−23]

We recently analyzed the full set of helical protein interfaces in the Protein Data Bank (PDB) to identify potentially suitable candidates for inhibition by small molecules or helix mimetics.[24,25] We began by identifying protein complexes that feature helical segments at interfaces and computationally evaluating the energetic contribution of helices to complex formation (Figure 1). Although several examinations of PPIs have been performed, our approach is unique in its focus on interfaces involving a specific secondary structure. The key motivation behind this structure-based dissection of interfaces is to aid the systematic design of synthetic inhibitors of PPIs.

In earlier reports, we categorized helical protein interfaces identified with our algorithm by cellular functions[24] and proposed a predictive scale for inhibition of PPIs by synthetic ligands.[25] These studies focused on the disposition and energetic contributions of "hot spot" residues within interfacial helices and provided a list of interactions that have not previously been

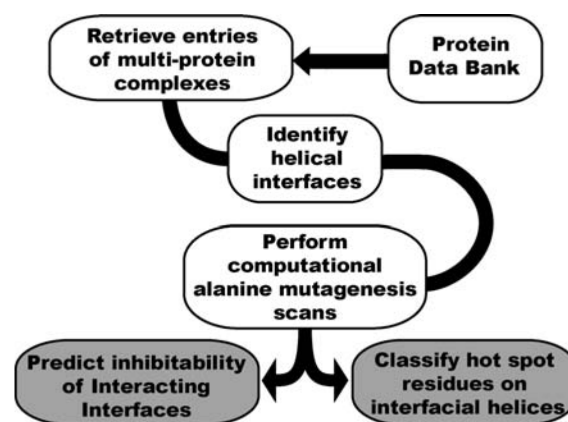

Figure 1. Evaluation of structures from the Protein Data Bank to identify and assess helical interfaces in protein−protein interactions. The helical interfaces were evaluated by computational alanine scanning mutagenesis.

inhibited along with candidate helices whose mimics may serve as potent inhibitors. On the basis of these predictions, we designed cell-permeable synthetic α-helices that interfere with PPIs that control transcription of hypoxia-inducible genes and Ras signaling.[13,14] Here we examine the composition and characteristics of helical domains identified to be critical for protein complex formation. We analyzed the full set of available protein complexes in the PDB to assess amino acid propensity at helical interfaces, the location and positioning of hot spot residues on helices, and contact residues on partner proteins.

Examination of entries in the PDB (version August 2009) showed that multiprotein complexes constitute roughly 15% of the databank.[24,25] Of these, 62% feature a helix at the interface, highlighting the role of α-helices in PPIs. However, presence of a helix at the interface does not imply a critical role for the particular helix in the interaction. To evaluate the energetic contribution of each helix to the complex formation, we employed computational alanine scanning mutagenesis scans within Rosetta to identify residues that contribute most strongly to complex formation.[26,27] Alanine scanning mutagenesis is a standard approach for identifying hot spot residues.[28] The results of this analysis have been reported along with a full list of filtered PPIs.[25]

Three general strategies have been used to develop helix mimetics: helix stabilization, helical foldamers, and helical surface mimetics.[29,30] Helix stabilization methods based on side-chain cross-links[18,31] and hydrogen-bond surrogates (HBSs)[32] preorganize amino acid residues and initiate helix formation. Helical

Received: June 30, 2011
Published: August 16, 2011

ACS Publications © 2011 American Chemical Society

dx.doi.org/10.1021/ja206074j | *J. Am. Chem. Soc.* 2011, 133, 14220–14223

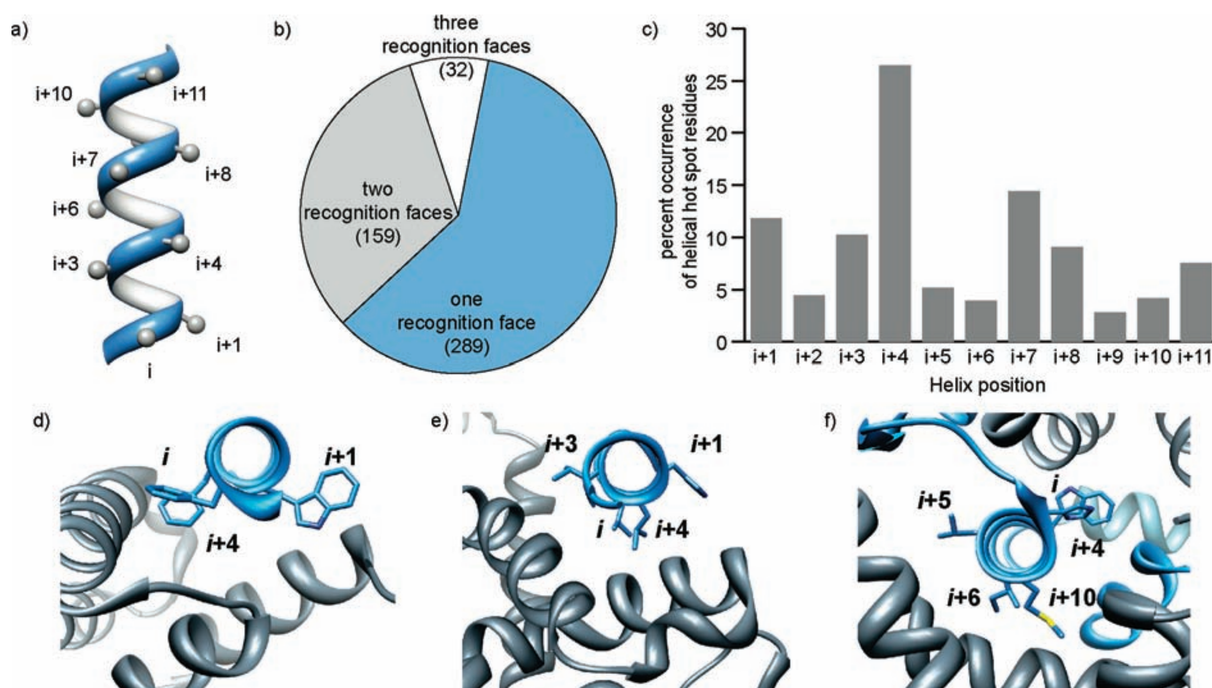

Figure 2. Energetic contributions of residues on different faces of interfacial helices. (a) Positioning of side-chain residues on a canonical α-helix. (b) Percent occurrence of hot spot residues on one, two, or three helical faces (the total number of helices in each category is shown in parentheses). (c) Percent occurrence of hot spot residues as a function of helix position. (d−f) examples of protein complexes with hot spot residues on one face, two faces, and three faces (PDB entries 1xl3, 1xiu, and 1or7).

foldamers,[11,33] such as β-peptides[34−36] and peptoids,[37] are composed of amino acid analogues and are capable of adopting conformations similar to those found in natural proteins. Helical surface mimetics utilize conformationally restricted scaffolds with attached functional groups that resemble the i, $i + 3$, $i + 4$, and $i + 7$ pattern of side-chain positioning on an α-helix (Figure 2a). Surface mimetics typically impart functionality from one face of the helix,[38] while stabilized peptide helices and foldamers are able to reproduce functionality present on multiple faces of the target helix. A key advantage of the helix surface mimicry is that it affords low-molecular-weight compounds as modulators of protein interactions.[39−44]

A catalog of PPIs that predicts energetic contributions of residues on different faces of interfacial helices should provide an invaluable starting point for the design of synthetic inhibitors of protein complex formation. Such a data set would enable the design of an appropriate mimic for a particular interface of interest. On the basis of this hypothesis, we analyzed the occurrence of hot spot residues on different helical faces. Hot spot residues are defined as residues that upon mutation to alanine are predicted to decrease the binding energy by a threshold value $\Delta\Delta G_{\text{bind}} \geq 1.0$ kcal mol^{-1}, as measured in Rosetta energy units.[5,7,8,26] We used a cutoff value of $\Delta\Delta G_{\text{avg}} \geq 2.0$ kcal mol^{-1} to define strongly and weakly interacting interfaces.[25] This average binding energy difference accounts for all hot spot residues at an interface. Our current data set consists of 480 "strongly interacting" interfaces, which were closely examined. The number of such complexes will grow as new entries are deposited in the PDB.

Analysis revealed that roughly 60% of the helical interfaces in the data set feature helices with hot spot residues on one face of the helix (Figure 2b,d), a third of the complexes utilize helices with hot spots on two faces (Figure 2b,e), and roughly 10% require all three faces for interaction with the target protein partner (Figure 2b,f). The full list of PPIs that correspond to each category is included in the Supporting Information (SI). Residues i, $i + 1$, and $i + 2$ reside on different faces of a single helical turn; we examined models of each interfacial helix individually, as the noninteger number of residues per helical turn makes it difficult to classify locations of noncontiguous residues on helical faces. The overall percent occurrence of hot spot residues at the first 12 positions in interfacial helices is depicted in Figure 2c. Our inquiry suggests that helix surface mimetics may prove to be a highly effective class of synthetic inhibitors; however, a significant fraction of PPIs will require mimetics that array protein-like functionality on multiple faces. Figure 3 shows the

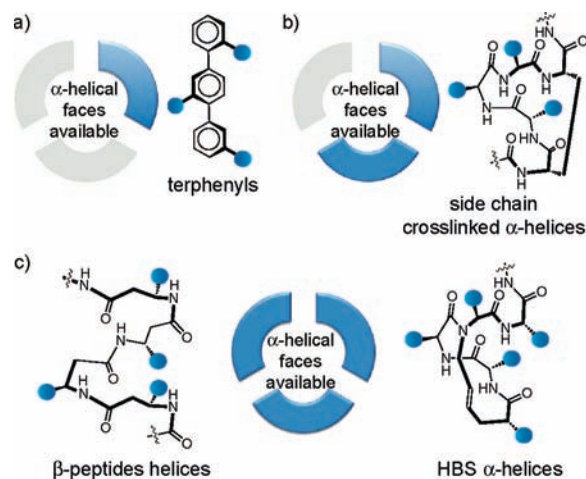

Figure 3. Potential of various helix mimetics to reproduce functionality of one, two, or all three faces of protein α-helices.

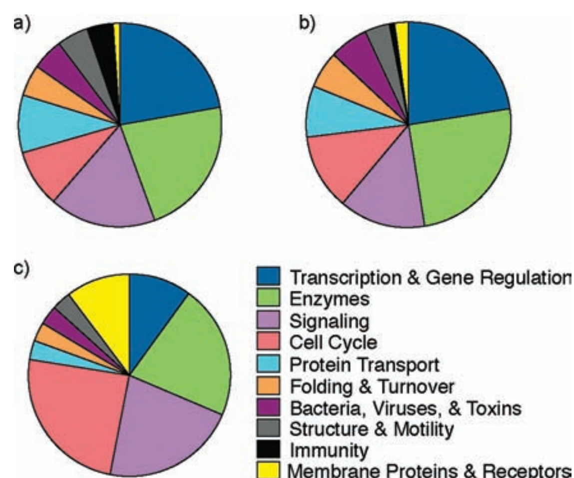

Figure 4. Functions associated with PPIs featuring hot spots on (a) one helical face, (b) two helical faces, and (c) three helical faces.

targeting potential of various helix mimetics. Terphenyls, the prototypical helix surface mimetics, imitate one helical face, and side-chain cross-linked helices can reproduce functionality of up to two faces, although the linker itself may interact with the protein pocket. HBS helices and β-peptide foldamers potentially afford complete replicas of functionality present on protein α-helices. We categorized the functions of PPIs featuring hot spots on different numbers of helical faces as defined in the PDB (Figure 4). Some interactions could fall into more than one function category. The four largest categories for each type are gene regulation, enzymatic function, cell cycle, and signaling.

The helical interfaces that form this data set allow a detailed analysis of basic interactions that underlie protein complex formation. Examination of these fundamental forces will inform the design of PPI inhibitors. We calculated the percentage of each helical residue that contributes strongly to binding. (Glycine and proline residues were exempted from alanine scanning because substitutions of proline or glycine by alanine may cause a conformational change in the protein backbone.) Leucine dominates the interface region (Figure 5a), which is not surprising because leucine is also the most prevalent residue in proteins in general. After normalization for natural abundance,[45] we found that aromatic residues and arginine, along with leucine, are overrepresented as hot spots at helical interfaces in comparison to polar residues (Figure 5c). These results correspond with those of previous studies of the types of amino acids appearing as hot spot residues in protein interfaces (Figure S2 in the SI),[5,9,10,46,47] although our data set is considerably larger than those previously examined. We expect these results to help guide the design of helix mimetic libraries.[40,43,44,48−50]

Hydrophobic and aromatic residues constitute a majority of hot spot residues; however, polar and charged residues are also significant contributors at interfaces (Figure 5b).[51] This analysis supports the common perception that PPIs are generally hydrophobic but feature key salt bridges and other polar interactions that appreciably influence the binding energy landscape.[8] This view is further supported by the evaluation of residues on the partner protein that are within 5 Å of the helical hotspot residue (Figure S3). Not surprisingly, a majority of residues that are within the specified radius of a hydrophobic residue are

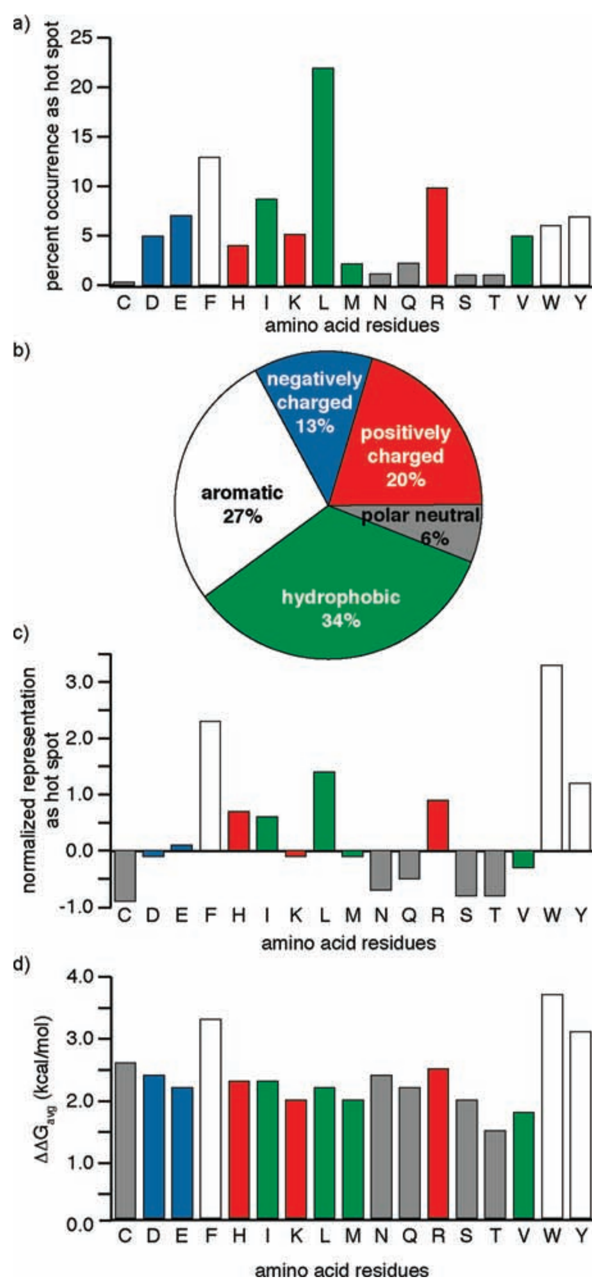

Figure 5. (a) Percent occurrence of hot spot amino acids in helix-mediated protein interfaces. (b) Percent occurrence of hot spot residues classified into similar groups. (c) Representation of hot spot amino acids normalized to the natural abundance of amino acids in proteins. (d) Average predicted decrease in binding energy of helical interfaces upon mutation of hot spot residues to alanine. Color code: aromatic (phenylalanine, tryptophan, and tyrosine), white; hydrophobic (isoleucine, leucine, and valine), green; negatively charged (aspartic acid and glutamic acid), blue; polar neutral (asparagine, cysteine, glutamine, serine, and threonine), gray; positively charged (arginine, histidine, and lysine), red.

themselves hydrophobic, which is consistent with the hypothesis that the burial of a hot spot in a hydrophobic environment is a major stabilizing influence.[5] In this respect, it is interesting to note that on average, mutations of aromatic residues to alanine are more destabilizing than substitution of other interfacial residues, with the effect being dependent on the size of the aromatic ring (Figure 5d).

Helical protein−protein interactions have to date been successfully targeted by a diverse array of mimetics.[12,14,16,18,21,23] Preliminary success in this field validates helix design concepts from multiple research groups and provides an impetus for designing inhibitors of interactions previously considered to be intractable to inhibition by synthetic ligands. A key motivation for our approach is to bridge the significant chasm between the elegant design of helix mimetics and their sporadic use in biology. This study provides a list of targets to be considered for different classes of helix mimetics based on the number of contact surfaces the target helix utilizes for interactions with partner proteins. We have successfully used this information to identify two new classes of PPIs amenable to disruption by helix mimetics,[13,14] supporting the basic hypotheses and results of these computational efforts.

■ ASSOCIATED CONTENT

S Supporting Information. Lists of helical PPIs with predicted occurrences of hot spot residues on different faces of the target helix and a summary of helix contact residues. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
arora@nyu.edu

■ ACKNOWLEDGMENT

This work was financially supported by the National Institutes of Health (GM073943) and the National Science Foundation (CHE 0848410). B.N.B. thanks the New York University for a Kramer Predoctoral Fellowship, and A.L.J. thanks NYU for the Dean's Dissertation Fellowship.

■ REFERENCES

(1) Ryan, D. P.; Matthews, J. M. *Curr. Opin. Struct. Biol.* **2005**, *15*, 441.
(2) Arkin, M. R.; Wells, J. A. *Nat. Rev. Drug Discovery* **2004**, *3*, 301.
(3) Hajduk, P. J.; Greer, J. *Nat. Rev. Drug Discovery* **2007**, *6*, 211.
(4) Lo Conte, L.; Chothia, C.; Janin, J. *J. Mol. Biol.* **1999**, *285*, 2177.
(5) Bogan, A. A.; Thorn, K. S. *J. Mol. Biol.* **1998**, *280*, 1.
(6) Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. *Proteins* **2007**, *68*, 803.
(7) Clackson, T.; Wells, J. A. *Science* **1995**, *267*, 383.
(8) Keskin, Z.; Gursoy, A.; Ma, B.; Nussinov, R. *Chem. Rev.* **2008**, *108*, 1225.
(9) Guharoy, M.; Chakrabarti, P. *Bioinformatics* **2007**, *23*, 1909.
(10) Jones, S.; Thornton, J. M. *Prog. Biophys. Mol. Biol.* **1995**, *63*, 31.
(11) Goodman, C. M.; Choi, S.; Shandler, S.; DeGrado, W. F. *Nat. Chem. Biol.* **2007**, *3*, 252.
(12) Horne, W. S.; Johnson, L. M.; Ketas, T. J.; Klasse, P. J.; Lu, M.; Moore, J. P.; Gellman, S. H. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 14751.
(13) Henchey, L. K.; Kushal, S.; Dubey, R.; Chapman, R. N.; Olenyuk, B. Z.; Arora, P. S. *J. Am. Chem. Soc.* **2010**, *132*, 941.
(14) Patgiri, A.; Yadav, K. K.; Arora, P. S.; Bar-Sagi, D. *Nat. Chem. Biol.* **2011**, *7*, 585.
(15) Wang, D.; Lu, M.; Arora, P. S. *Angew. Chem., Int. Ed.* **2008**, *47*, 1879.
(16) Moellering, R. E.; Cornejo, M.; Davis, T. N.; Del Bianco, C.; Aster, J. C.; Blacklow, S. C.; Kung, A. L.; Gilliland, D. G.; Verdine, G. L.; Bradner, J. E. *Nature* **2009**, *462*, 182.
(17) Walensky, L. D.; Kung, A. L.; Escher, I.; Malia, T. J.; Barbuto, S.; Wright, R. D.; Wagner, G.; Verdine, G. L.; Korsmeyer, S. J. *Science* **2004**, *305*, 1466.
(18) Harrison, R. S.; Shepherd, N. E.; Hoang, H. N.; Ruiz-Gomez, G.; Hill, T. A.; Driver, R. W.; Desai, V. S.; Young, P. R.; Abbenante, G.; Fairlie, D. P. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 11686.
(19) Smith, B. A.; Daniels, D. S.; Coplin, A. E.; Jordan, G. E.; McGregor, L. M.; Schepartz, A. *J. Am. Chem. Soc.* **2008**, *130*, 2948.
(20) Harker, E. A.; Schepartz, A. *ChemBioChem* **2009**, *10*, 990.
(21) Cummings, C. G.; Hamilton, A. D. *Curr. Opin. Chem. Biol.* **2010**, *14*, 341.
(22) Hammond, M. C.; Harris, B. Z.; Lim, W. A.; Bartlett, P. A. *Chem. Biol.* **2006**, *13*, 1247.
(23) Ko, E.; Liu, J.; Burgess, K. *Chem. Soc. Rev.* **2011**, *40*, 4411.
(24) Jochim, A. L.; Arora, P. S. *Mol. BioSyst.* **2009**, *5*, 924.
(25) Jochim, A. L.; Arora, P. S. *ACS Chem. Biol.* **2010**, *5*, 919.
(26) Kortemme, T.; Kim, D. E.; Baker, D. *Sci. STKE* **2004**, l2.
(27) Kortemme, T.; Baker, D. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 14116.
(28) Cunningham, B. C.; Wells, J. A. *Science* **1989**, *244*, 1081.
(29) Henchey, L. K.; Jochim, A. L.; Arora, P. S. *Curr. Opin. Chem. Biol.* **2008**, *12*, 692.
(30) Edwards, T. A.; Wilson, A. J. *Amino Acids* **2011**, *41*, 743.
(31) Schafmeister, C. E.; Po, J.; Verdine, G. L. *J. Am. Chem. Soc.* **2000**, *122*, 5891.
(32) Patgiri, A.; Jochim, A. L.; Arora, P. S. *Acc. Chem. Res.* **2008**, *41*, 1289.
(33) Gellman, S. H. *Acc. Chem. Res.* **1998**, *31*, 173.
(34) Cheng, R. P.; Gellman, S. H.; DeGrado, W. F. *Chem. Rev.* **2001**, *101*, 3219.
(35) Horne, W. S.; Gellman, S. H. *Acc. Chem. Res.* **2008**, *41*, 1399.
(36) Seebach, D.; Gardiner, J. *Acc. Chem. Res.* **2008**, *41*, 1366.
(37) Yoo, B.; Kirshenbaum, K. *Curr. Opin. Chem. Biol.* **2008**, *12*, 714.
(38) Marimganti, S.; Cheemala, M. N.; Ahn, J. M. *Org. Lett.* **2009**, *11*, 4418.
(39) Plante, J. P.; Burnley, T.; Malkova, B.; Webb, M. E.; Warriner, S. L.; Edwards, T. A.; Wilson, A. J. *Chem. Commun.* **2009**, 5091.
(40) Shaginian, A.; Whitby, L. R.; Hong, S.; Hwang, I.; Farooqi, B.; Searcey, M.; Chen, J.; Vogt, P. K.; Boger, D. L. *J. Am. Chem. Soc.* **2009**, *131*, 5564.
(41) Restorp, P.; Rebek, J., Jr. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5909.
(42) Tosovska, P.; Arora, P. S. *Org. Lett.* **2010**, *12*, 1588.
(43) Lee, J. H.; Zhang, Q.; Jo, S.; Chai, S. C.; Oh, M.; Im, W.; Lu, H.; Lim, H. S. *J. Am. Chem. Soc.* **2011**, *133*, 676.
(44) Buhrlage, S. J.; Bates, C. A.; Rowe, S. P.; Minter, A. R.; Brennan, B. B.; Majmudar, C. Y.; Wemmer, D. E.; Al-Hashimi, H.; Mapp, A. K. *ACS Chem. Biol.* **2009**, *4*, 335.
(45) Nelson, D. L.; Lehninger, A. L.; Cox, M. M. *Principles of Biochemistry*, 5th ed.; W.H. Freeman: New York, 2008.
(46) Kossiakoff, A. A.; Koide, S. *Curr. Opin. Struct. Biol.* **2008**, *18*, 499.
(47) Argos, P. *Protein Eng.* **1988**, *2*, 101.
(48) Ko, E.; Liu, J.; Perez, L. M.; Lu, G.; Schaefer, A.; Burgess, K. *J. Am. Chem. Soc.* **2011**, *133*, 462.
(49) Campbell, F.; Plante, J. P.; Edwards, T. A.; Warriner, S. L.; Wilson, A. J. *Org. Biomol. Chem.* **2010**, *8*, 2344.
(50) Murray, J. K.; Farooqi, B.; Sadowsky, J. D.; Scalf, M.; Freund, W. A.; Smith, L. M.; Chen, J. D.; Gellman, S. H. *J. Am. Chem. Soc.* **2005**, *127*, 13271.
(51) Sheinerman, F. B.; Norel, R.; Honig, B. *Curr. Opin. Struct. Biol.* **2000**, *10*, 153.