JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

Broad Distribution of Energetically Important Contacts across an Extended Protein Interface

Lisa M. Johnson, W. Seth Horne, and Samuel H. Gellman*

Department of Chemistry, University of Wisconsin, Madison, Wisconsin 53706, United States

Supporting Information

ABSTRACT: Infection of cells by HIV depends upon profound structural rearrangements within the trimeric viral protein gp41. Critical to this process is the formation of a six-helix bundle in which a set of three N-terminal heptad repeat (NHR) helices assemble to form a core displaying long grooves that provide docking sites for three C-terminal heptad repeat (CHR) helices. We report experiments designed to discriminate between two alternative hypotheses regarding the source of affinity between individual CHR helices and the complementary groove: (1) affinity is dominated by interactions of a small cluster of side chains at one end of the CHR helix; or (2) affinity depends upon interactions distributed across the long CHR helix. We have employed two complementary experimental designs, and results from both favor the latter hypothesis.

Associations between protein molecules play critical biological roles, including transmission of information, regulation of gene expression, and recognition of hosts by pathogens. This functional importance has inspired widespread interest in inhibitors of specific protein−protein interactions as therapeutic agents.[1] However, blocking or mimicking protein−protein interactions with small molecules, the traditionally favored source of drugs, has proven to be extremely challenging. A few systems have yielded to clever designs and determined effort,[2] but it remains an open question whether approaches based on small molecules will be broadly successful for inhibiting disease-related macromolecular associations. Interactions that involve extensive protein−protein contact may be especially resistant to inhibition via small molecules because of surface area limitations, although the occurrence of "hot spots" or cryptic binding sites on large protein surfaces can alleviate this problem in some cases.[3] Here we use a combination of traditional and nontraditional strategies to evaluate whether sources of affinity are focused or distributed across a large protein interface that forms within the trimeric form of HIV protein gp41.

The entry of HIV RNA and proteins into the target cell cytoplasm is orchestrated by gp41, which induces fusion of the viral envelope with the cell membrane.[4] This process requires large conformational changes within the gp41 trimer.[5] The AIDS drug enfuvirtide, a 36-mer peptide derived from gp41, is thought to block rearrangement from an extended to a compact state of the gp41 trimer.[6] Formation of the compact state is driven by assembly of a bundle of six α-helices, with each gp41 molecule contributing one N-terminal heptad repeat (NHR) segment and one C-terminal heptad repeat (CHR) segment. The crystal structure of the six-helix bundle formed by gp41-derived peptides designated N36 (from the NHR segment) and C34 (from the CHR segment) reveals an NHR trimeric core with three long grooves on its surface; the helical CHR segments pack into these grooves.[5] Deep clefts occur at one end of the trimeric N36 core, and each cleft is filled by a trio of hydrophobic side chains from C34 (Trp-Trp-Ile motif) that are aligned by α-helix formation. These clefts in the NHR trimer have been suggested as potential sites for binding of small molecules, which might interfere with formation of the gp41 six-helix bundle and thereby block HIV entry.[7]

There have been a number of attempts to develop ligands of low molecular weight that occupy the gp41 NHR clefts.[8] However, despite the creativity manifested in these efforts, the small molecules and short peptides reported to date are at least 3 orders of magnitude less potent than the best large peptides for inhibition of HIV infection. These results raise the possibility that the affinity of a CHR α-helix (∼10 helical turns for C34) for the NHR trimer groove is so broadly distributed that the efficacy of small inhibitors will be intrinsically limited.

We have now probed the distribution of binding affinity for a full-length CHR helix along the NHR trimer groove via a novel experimental design based on the recent development of α/β-peptide foldamers that mimic the CHR α-helix.[9] These molecules were generated from a potent CHR-derived α-peptide, T-2635,[10] by replacing a subset of α-amino acid residues with analogous β-amino acid residues, many of which are preorganized to promote helix formation. Placement of α→β substitution sites throughout the sequence discourages protease degradation. This previous effort led to α/β-peptide **1** (Figure 1), which functions as a potent inhibitor of HIV infection in cell-based assays.[9] Here we use comparisons among T-2635, **1**, and chimeric peptides **2**−**4** to determine how different portions of the CHR helix contribute to binding to the protein gp41-5.[8e]

Designed protein gp41-5 contains three NHR segments and two CHR segments and is intended to adopt a five-helix bundle tertiary structure that displays a binding groove for a single CHR segment. This protein provides the basis for a previously reported competition fluorescence polarization (FP) assay for evaluating ligands for the NHR binding groove.[8e,9] To gain insight on the binding of CHR-mimetic α/β-peptides to gp41-5, we solved the crystal structure of **1** bound to this protein (Figure 2a). The α/β-peptide backbone adopts a conformation

Received: April 12, 2011
Published: June 06, 2011

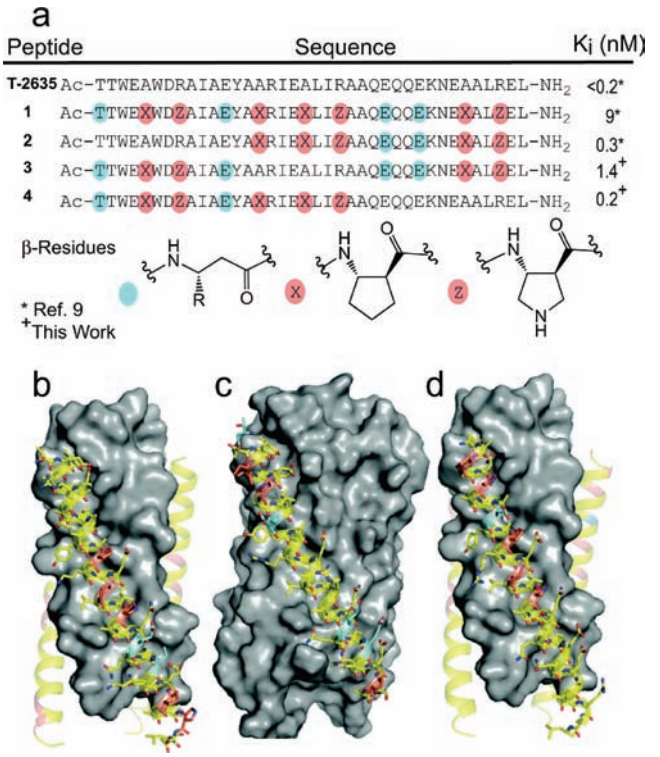

Figure 1. (a) Sequences of α/β-peptides derived from α-peptide T-2635, with K_i values for binding to designed protein gp41-5, as determined with a fluorescence polarization assay. (b) Six-helix bundle formed by three molecules of α-peptide N36 (gray surface) plus three molecules of **2** (atomic detail; carbon from α residues is yellow, carbon from acyclic β residues is light blue, carbon from cyclic β residues is pink, nitrogen is dark blue, and oxygen is red). Resolution = 2.8 Å; ref 9; PDB 3G7A. (c) Complex between gp41-5 (gray surface) and one molecule of **3** (atomic detail; colors as in part b). Resolution = 2.1 Å; PDB 3O40. (d) Six-helix bundle formed by three molecules of α-peptide N36 (gray surface) plus three molecules of **4** (atomic detail; colors as in part b). Resolution = 2.6 Å; PDB 3O3Z.

very similar to that of an α-helical α-peptide, despite the presence of an extra backbone carbon atom in each helical turn. Using PDBe PISA (Protein Interfaces, Surfaces and Assemblies),[11] we calculate that formation of this complex buries ∼1100 Å^2 of surface area from gp41-5 and ∼1230 Å^2 of surface area from α/β-peptide **1**.

Key side chains projecting from helical α/β-peptide **1** occupy the correct sites in the groove displayed by gp41-5, as illustrated in Figure 2b for the Trp-Trp-Ile motif. As previously reported,[9] **1** binds tightly to gp41-5 (K_i = 9 nM; FP assay). However, despite the structural similarity of α/β-peptide **1** to an α-helical CHR segment, **1** binds less tightly to gp41-5 than does α-peptide T-2635 itself (K_i < 0.2 nM). This difference is unlikely to arise directly from changes in CHR surface functionality arising from α→β substitutions, because all interfacial residues are identical in **1** and T-2635 (i.e., the interfacial residues in these two analogues are derived from the same set of α-amino acids; the β residues of **1** are oriented away from gp41-5). Therefore, it is possible that replacing the pure α backbone of T-2635 with the α/β backbone of **1** causes some degree of mismatch between the surfaces of the α/β-peptide and gp41-5, relative to the complex between T-2635 and gp41-5. Such a mismatch would have to be subtle, however, because comparison of crystallographic structures for

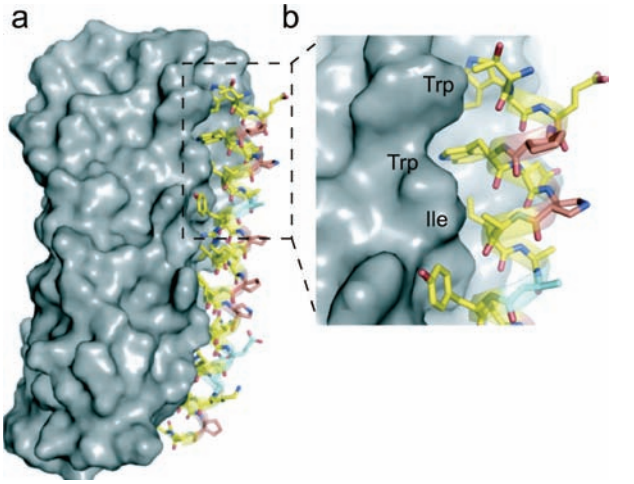

Figure 2. (a) Crystal structure of the complex between α/β-peptide **1** (atomic detail; colors as in Figure 1b) and designed protein gp41-5 (gray). Resolution = 2.8 Å; PDB 3O43. (b) Close-up of the interaction of the Trp-Trp-Ile motif from α/β-peptide **1** with the complementary cleft presented by gp41-5.

α/β-peptide **1** and α-peptide T-2635 reveals no glaring differences (this comparison is described in detail in the Supporting Information (SI)). Alternatively, the differences in affinity between T-2635 and α/β-peptide **1** might arise from a difference in the stabilites of the helical conformations adopted by these oligomers. In either case, this system provides a novel opportunity for assessing contributions to overall affinity from contacts along the entire interface between gp41-5 and a complementary 10-turn helix.

Compounds **2**−**4** constitute a set of α + α/β chimeric peptides[12] in which the β residues along one-third of the length of **1** have been replaced with the original α residues (Figure 1). The β→α reversions occur in the N-terminal region for chimera **2**, the middle region for chimera **3**, and the C-terminal region for chimera **4**. Figure 1 provides apparent dissociation constant (K_i) values derived from the competition FP assay. If the affinity of the CHR-derived α-helical peptide T-2635 is broadly distributed along the entire NHR groove, then each set of partial β→α replacements should lead to a significant increase in affinity, because complete β→α replacement (**1**→T-2635) causes a substantial decrease in K_i. On the other hand, if the Trp-Trp-Ile motif of the CHR segment makes a dominant contribution to six-helix bundle stability, then partial β→α replacement in the N-terminal portion (**2**) should have a more favorable effect on binding than does partial β→α replacement in the central or C-terminal portions (**3** or **4**). The K_i data in Figure 1 clearly show that substantial improvements in affinity result from each partial β→α replacement. It is particularly noteworthy that the effect of β→α replacement in the C-terminal portion (**4**) is comparable to that of replacement in the N-terminal portion (**2**). Because of FP assay sensitivity limits, we cannot determine whether chimeric molecules **2** and **4** match T-2635 in affinity; all three have K_i values in the sub-nanomolar range. Overall, these data suggest that affinity for a CHR helix is broadly distributed across the long groove formed by the NHR trimer core.

High-resolution structural data establish that chimeric peptides **2**−**4** are competent to participate in six-helix bundle formation, comparable in this regard to T-2635 and α/β-peptide **1**.

This behavior is essential for the validity of the conclusion reached in the preceding paragraph. Figure 1 shows the previously determined crystal structure of the six-helix bundle formed by **2** + N36,[9] along with new crystal structures for **3** + gp41-5 and for the six-helix bundle formed by **4** + N36. Although the stoichiometry of the complex differs across this series (3+3 assembly vs 1+1 assembly), the data show that each chimeric oligomer closely mimics an α-helical CHR segment. Alignment of the structure of T-2635 with α/β-peptides **1**−**4**, each involved in a six-helix bundle assembly, revealed excellent structural mimicry of the α-helical backbone and of the positions of side chains that make intermolecular contacts by the α/β-peptides (details in SI). Comparisons focused on the three regions, N-terminal, middle, and C-terminal, for each α/β-peptide vs T-2635 pairing indicated that the α-peptide segments of the chimeric molecules displayed the best alignment with T-2635, but segments containing β residues were nevertheless good structural mimics of the α-helical prototype. Overall, these data strengthen the conclusion that the α→β replacement strategy we employ provides oligomers that can recapitulate the structural and recognition properties of an α-helical surface, and that variations in affinity for gp41-5 among these CHR-mimetic peptides can be interpreted in terms of distribution of energetically important intermolecular contacts.

As a complement to the backbone-modification strategy described above, we undertook a more traditional side-chain-modification strategy (mutagenesis) to test the hypothesis that binding affinity for the NHR-defined groove of the gp41 six-helix bundle is broadly distributed across the contact surface on a CHR helix. The sequence of each CHR-derived oligomer can be viewed as a series of heptad segments, each corresponding to two helical turns, with positions designated *abcdefg*.[13] By convention, the "stripe" of side chains that constitutes the core of the interaction surface is assigned to positions *a* and *d* in the heptad pattern (Figure 3c); for example, the Trp-Trp-Ile motif in the N-terminal region corresponds to an *a*-*d*-*a* triplet. (In α/β-peptides **1**−**4**, the β-residues occupy positions *c* and *f*, which are oriented toward solvent in the six-helix bundle and therefore do not make contact with the NHR core.) Side chains from positions *e* and *g* flank the *a*+*d* stripe. Our structural data reveal numerous specific contacts between *e* or *g* side chains on CHR-derived oligomers (T-2635 or **1**−**4**) and the groove formed by NHR segments, which suggests that *e* and *g* side chains contribute to affinity for gp41-5. Alignment of each α/β-peptide among **1**−**4** vs T-2635 revealed strong similarities among the subset of *a* and *d* residues in each case (rmsd values 1.03−1.54 Å) and comparable similarities among the subset of *e* and *g* residues (rmsd values 1.12−1.65 Å; see SI for details).

Our mutagenesis approach was guided by the "regional" strategy outlined above; these experiments employed peptides composed entirely of α-amino acid residues (i.e., only side chains were varied relative to T-2635). Six α-peptides were prepared, and in each case multiple side chains were mutated to Ala (Figure 3a). In the N-terminal region, the three side chains at *a* and *d* positions (Trp-Trp-Ile motif) were mutated to Ala to generate the α-peptide designated **N-*ad***, and the two side chains of non-Ala residues at *e* and *g* positions were mutated to generate α-peptide **N-*eg***. Comparable mutations in the middle and C-terminal regions generated α-peptides **M-*ad***, **M-*eg***, **C-*ad***, and **C-*eg***. Although this multiple-mutation strategy does not reveal contributions of individual side chains (which would have required synthesis and evaluation of 17 long peptides), our approach enables us rapidly to elucidate the balance of contributions from

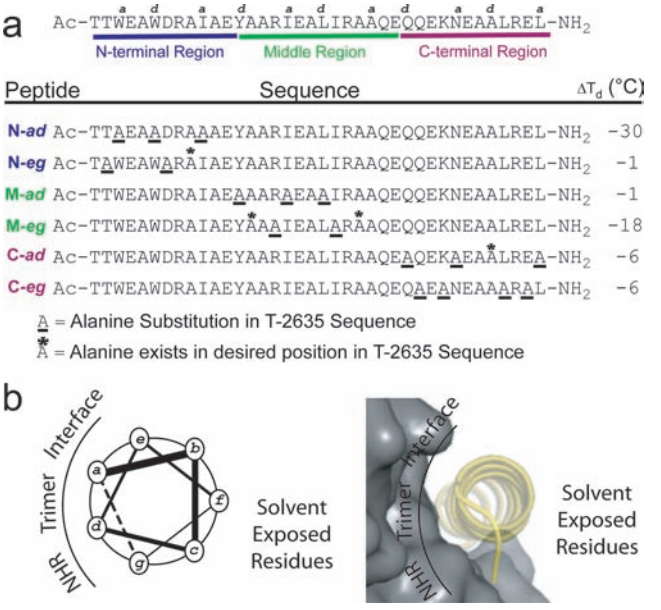

Figure 3. (a) Sequences of α-peptides derived from T-2635 via multiple mutations of original residues to alanine. T-2635 was divided into three regions, N-terminal, middle, and C-terminal, and alanine substitutions were made at all *a* and *d* positions or at all *e* and *g* positions within a region (see text for details). The alanines marked by an asterisk were unchanged from the T-2635 sequence. Shown at the right are changes in the thermal disruption temperature (T_d), relative to parent peptide T-2635, for the assembly formed by each of the alanine mutant peptides and N34. (b) Helical wheel diagram of T-2635, highlighting positions within the heptad repeat (*a*, *d*, *e*, and *g*) that make contact with the NHR trimer and positions within the heptad repeat (*b*, *f*, and *c*) that are solvent exposed. The *f* and *c* positions were the sites of β substitution.

a/*d* vs *e*/*g* side-chain sets to affinity and the way in which this balance varies along the sequence. This study is important because coiled-coil-type assemblies are typically thought to be driven largely by interactions involving side chains at *a* and *d* positions. Since it was necessary to compare these six mutant peptides to T-2635, which binds too tightly for K_i determination in the FP assay, we turned to another approach that has been widely used to assess the stability of CHR/NHR six-helix bundles: thermal disruption, as monitored by circular dichroism (CD).[7a,10] These measurements are typically not fully reversible, because some NHR peptide precipitates at high temperature; therefore, we refer to the midpoint of the transition as T_d (for "disruption"), rather than the more common T_m. Despite the lack of full reversibility, T_d values determined in this way are highly reproducible and therefore provide a basis for comparisons among T-2635 and mutants in terms of six-helix bundle formation with N36.

The T_d data summarized in Figure 3a support the hypothesis that six-helix bundle stability has significant contributions from side chains dispersed along the entire length of the CHR α-helix. The largest effect is observed for **N-*ad***, i.e., for mutation of the Trp-Trp-Ile motif to Ala-Ala-Ala, which is consistent with the widespread interest in this motif and the NHR-derived pocket into which it docks.[7,8] However, a substantial diminution of T_d is observed also for **M-*eg***, which shows that side chains in the flanking positions can contribute significantly to the stability of the six-helix assembly. The data suggest that the origin of key contacts varies along the length of the CHR helix, with *a* and *d* residues making dominant contributions at the N-terminus, *e* and *g*

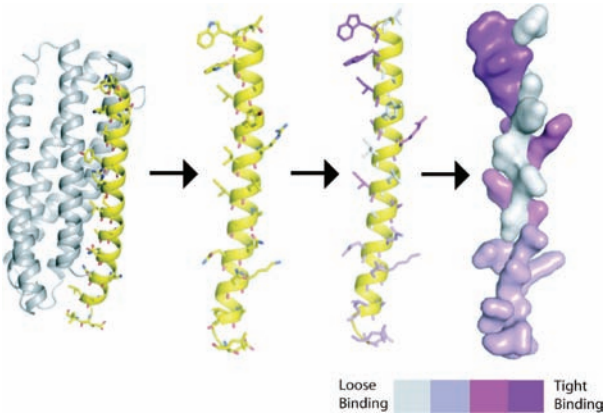

Figure 4. Qualitative summary of the distribution of affinity for the NHR groove across the surface of α-helical T-2635, based on T_d data determined with alanine mutant peptides.

residues dominant in the center, and comparable contributions from both sets at the C-terminus. Figure 4 shows qualitatively how the binding is distributed along the contact surface of 38-mer T-2635.

We have used two complementary approaches to explore the extent to which the sources of affinity are dispersed across a large protein interface. This interaction, within the trimeric HIV protein gp41, plays a crucial role in viral infection,[4,5] and disrupting this interaction leads to a clinically valuable outcome.[6] One of our experimental strategies is traditional: truncation of amino acid side chains to determine whether the omitted atoms play a stabilizing role. The other strategy is novel in that the polypeptide backbone is modified, while the set of side chains that make contact with the NHR-defined groove remains constant. (The latter effort has generated three new crystal structures that show at atomic resolution how α/β-peptides can mimic long α-helical α-peptides.) Both approaches suggest that energetically important contributions to overall binding affinity are scattered across the long contact surface of the CHR α-helix. These results may explain why it has so far proven difficult to identify gp41-directed small molecules or oligomers with high antiviral potency. The backbone modification strategy we have employed may prove useful for evaluating affinity distribution within other protein–protein interactions that rely on α-helices.[14]

■ ASSOCIATED CONTENT

Supporting Information. Complete refs 2b and 8l, experimental methods, and crystallographic information. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
gellman@chem.wisc.edu

■ ACKNOWLEDGMENT

This work was supported by the NIH (GM-056414). L.M.J. was supported in part by the Chemistry-Biology Interface Training Program (T32GM008505), and W.S.H. was supported in part by an NIH fellowship (CA119875). We are grateful to Katrina Forest and Kenneth Satyshur for technical advice and use of X-ray equipment. 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 no. DE-AC02-06CH11357

■ REFERENCES

(1) (a) Arkin, M. R.; Wells, J. A. *Nat. Rev. Drug Discovery* **2004**, *3*, 301. (b) Hajduk, P. J.; Greer, J. *Nat. Rev. Drug Discovery* **2007**, *6*, 211.

(2) (a) Vassilev, L. T.; Vu, B. T.; Graves, B.; Carvajal, D.; Podlaski, F.; Filipovic, Z.; Kong, N.; Kammlott, U.; Lukacs, C.; Klein, C.; Fotohui, N.; Liu, E. A. *Science* **2004**, *303*, 844–848. (b) Oltersdorf, T.; et al. *Nature* **2005**, *435*, 677–681. (c) Arkin, M. R.; Randal, M.; DeLano, W. L.; Hyde, J.; Luong, T. N.; Oslob, J. D.; Raphael, D. R.; Taylor, L.; Wang, J.; McDowell, R. S.; Wells, J. A.; Braisted, A. C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 1603. (d) 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. (e) Fry, D. C. *Biopolymers* **2006**, *84*, 535.

(3) (a) Clackson, T.; Wells, J. A. *Science* **2005**, *267*, 383. (b) Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. *Proteins* **2007**, *68*, 803.

(4) (a) Eckert, D. M.; Kim, P. S. *Annu. Rev. Biochem.* **2001**, *70*, 777. (b) Harrison, S. C. *Adv. Virus Res.* **2005**, *64*, 231.

(5) (a) Chan, D. C.; Fass, D.; Berger, J. M.; Kim, P. S. *Cell* **1997**, *89*, 263. (b) Tan, K.; Liu, J.; Wang, J.; Shen, S.; Lu, M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12303. (c) Weissenhorn, W.; Dessen, A.; Harrison, S. C.; Skehel, J. J.; Wiley, D. C. *Nature* **1997**, *387*, 426.

(6) Matthews, T.; Salgo, M.; Greenberg, M.; Chung, J.; DeMasi, R.; Bolognesi, D. *Nat. Rev. Drug Discovery* **2004**, *3*, 215.

(7) (a) Chan, D. C.; Chutkowski, C. T.; Kim, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 15613. (b) Sodroski, J. G. *Cell* **1999**, *99*, 243.

(8) (a) Eckert, D. M.; Malashkevich, V. N.; Hong, L. H.; Carr, P. A.; Kim, P. S. *Cell* **1999**, *99*, 103. (b) Ferrer, M.; Kapoor, T. M.; Strassmaier, T.; Weissenhorn, W.; Skehel, J. J.; Oprian, D.; Schreiber, S. L.; Wiley, D. C.; Harrison, S. C. *Nat. Struct. Biol.* **1999**, *6*, 953. (c) Ernst, J. T.; Kutzki, O.; Debnath, A. K.; Jiang, S.; Lu, H.; Hamilton, A. D. *Angew. Chem. Int. Ed.* **2002**, *41*, 278. (d) Stephens, O. M.; Kim, S.; Welch, B. D.; Hodsdon, M. E.; Kay, M. S.; Schepartz, A. *J. Am. Chem. Soc.* **2005**, *127*, 13126. (e) Frey, G.; Rits-Volloch, S.; Zhang, X. Q.; Schooley, R. T.; Chen, B.; Harrison, S. C. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13938. (f) Welch, B. D.; VanDemark, A. P.; Heroux, A.; Hill, C. P.; Kay, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 16828. (g) Liu, S.; Wu, S.; Jiang, S. *Curr. Pharm. Design* **2007**, *13*, 143. (h) Wang, D.; Lu, M.; Arora, P. S. *Angew. Chem. Int. Ed.* **2008**, *47*, 1879. (i) Liu, K.; Lu, H.; Qi, Z.; Teixeira, C.; Barbault, F.; Fan, B. T.; Liu, S.; Jiang, S.; Xie, L. *J. Med. Chem.* **2008**, *51*, 7843. (j) Gochin, M.; Cai, L. *J. Med. Chem.* **2009**, *52*, 4338. (k) Liu, B.; Joseph, R. W.; Dorwey, B. D.; Schiksnis, R. A.; Northrop, K.; Bukhtiyarova, M.; Springman, E. B. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 5693. (l) Stewart, K. D.; et al. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 612.

(9) 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.

(10) Dwyer, J. J.; Wilson, K. L.; Davison, D. K.; Freel, S. A.; Seedorff, J. E.; Wring, S. A.; Tvermoes, N. A.; Matthews, T. J.; Greenberg, M. L.; Delmedico, M. K. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 12772.

(11) Krissinel, E.; Henrick, K. *J. Mol. Biol.* **2007**, *372*, 774–797.PDBe PISA can be accessed via http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html.

(12) Sadowsky, J. D.; Schmitt, M. A.; Lee, H. S.; Umezawa, N.; Wang, S.; Tomita, Y.; Gellman, S. H. *J. Am. Chem. Soc.* **2005**, *127*, 11966.

(13) (a) Woolfson, D. N. *Adv. Protein Chem.* **2005**, *70*, 79. (b) Grigoryan, G.; Keating, A. E. *Curr. Opin. Struct. Biol.* **2008**, *18*, 477.

(14) (a) Jochim, A. L.; Arora, P. S. *Mol. BioSyst.* **2009**, *5*, 924. (b) Jochim, A. L.; Arora, P. S. *ACS Chem. Biol.* **2010**, *5*, 919.