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

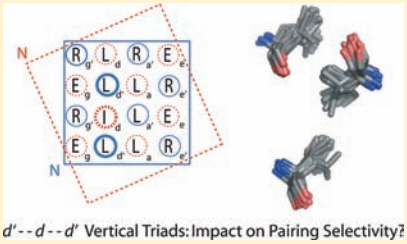
The *d′*--*d*--*d′* Vertical Triad Is Less Discriminating Than the *a′*--*a*--*a′* Vertical Triad in the Antiparallel Coiled-Coil Dimer Motif

Jay D. Steinkruger,[†] Gail J. Bartlett,[§] Erik B. Hadley,[†] Lindsay Fay,[†] Derek N. Woolfson,*[,§,‡] and Samuel H. Gellman*[,†]

[†]Department of Chemistry, University of Wisconsin, Madison, Wisconsin 53706, United States
[§]School of Chemistry, University of Bristol, Bristol BS8 1TS, U.K.
[‡]School of Biochemistry, University of Bristol, Bristol BS8 1TD, U.K.

Supporting Information

ABSTRACT: Elucidating relationships between the amino-acid sequences of proteins and their three-dimensional structures, and uncovering non-covalent interactions that underlie polypeptide folding, are major goals in protein science. One approach toward these goals is to study interactions between selected residues, or among constellations of residues, in small folding motifs. The α-helical coiled coil has served as a platform for such studies because this folding unit is relatively simple in terms of both sequence and structure. Amino acid side chains at the helix–helix interface of a coiled coil participate in so-called "knobs-into-holes" (KIH) packing whereby a side chain (the knob) on one helix inserts into a space (the hole) generated by four side chains on a partner helix. The vast majority of sequence–stability studies on coiled-coil dimers have focused on lateral interactions within these KIH arrangements, for example, between an *a* position on one helix and an *a′* position of the partner in a parallel coiled-coil dimer, or between *a*--*d′* pairs in an antiparallel dimer. More recently, it has been shown that vertical triads (specifically, *a′*--*a*--*a′* triads) in antiparallel dimers exert a significant impact on pairing preferences. This observation provides impetus for analysis of other complex networks of side-chain interactions at the helix–helix interface. Here, we describe a combination of experimental and bioinformatics studies that show that *d′*--*d*--*d′* triads have much less impact on pairing preference than do *a′*--*a*--*a′* triads in a small, designed antiparallel coiled-coil dimer. However, the influence of the *d′*--*d*--*d′* triad depends on the lateral *a′*--*d* interaction. Taken together, these results strengthen the emerging understanding that simple pairwise interactions are not sufficient to describe side-chain interactions and overall stability in antiparallel coiled-coil dimers; higher-order interactions must be considered as well.



INTRODUCTION

The coiled coil is a very common association mode for α-helices, both within and between polypeptides.[1] Intensive study has elucidated some of the "rules" that govern coiled-coil assembly, which has heightened interest in this motif in terms of predicting structure from primary sequence,[2] optimizing computational tools,[3] underpinning rational protein design,[1b] and developing new approaches to biomaterials science[4] and synthetic biology.[5] Most sequences that engage in coiled-coil interactions feature a heptad repeat pattern in which the positions are denoted *abcdefg*.[6] The first and fourth positions (*a* and *d*) are typically occupied by hydrophobic amino acid residues, while the remaining sites are variable, with many *b*, *c*, *e*, *f*, and *g* positions occupied by polar residues. This sequence pattern leads to amphipathic α-helix formation, with the *a* and *d* positions aligning and thereby creating a hydrophobic "stripe" along one side of the helix. In aqueous solution the hydrophobic effect drives the association of two or more α-helices into coiled-coil tertiary or quaternary structures. Different associated states have been observed in nature, including dimeric, trimeric, tetrameric, and higher-order oligomers, parallel and antiparallel arrangements of helices, and homo- and heteromeric assemblies. A common feature of all these structures is that the interhelical interfaces are characterized by the interdigitation of side chains at *a* and *d* and flanking *e* and *g* sites, a packing mode that Crick described as "knobs-into-holes" (KIH).[7] Here we explore multi-side-chain KIH interactions in a minimal, designed antiparallel heterodimeric system.

Experimental exploration of sequence–stability relationships among coiled coils and bioinformatics analysis of the protein structure database (PDB) has identified a number of principles that underlie coiled-coil dimer stability and partner pairing preferences. First, the hydrophobic core is formed mostly from side chains contributed by *a* and *d* positions, with a strong preference for the aliphatic side chains of Leu, Ile, and Val.[8] Second, occasional polar side chains can be accommodated in the hydrophobic core, particularly if they are paired in the coiled-coil state. The necessity for pairing, to satisfy H-bonding or other polar interaction potential, can provide a potent determinant of partner preference and orientation, as well as

Received: September 20, 2011
Published: January 31, 2012

   dx.doi.org/10.1021/ja208855x | J. Am. Chem.Soc. 2012, 134, 2626–2633

helix stoichiometry.[9,10] Third, interhelical Coulombic interactions among acidic and basic side chains at *e* and *g* positions, which flank the hydrophobic core, can strongly influence pairing selectivity.[11]

Despite this accumulation of insight, it remains challenging to predict coiled-coil association preferences and stabilities, especially in antiparallel systems, for sequences featuring the heptad repeat pattern. In addition, it is difficult to design sequences that display robust heteropairing specificities.[12] These difficulties arise at least in part from inadequate understanding of side-chain packing at coiled-coil dimer interfaces. For any given *a* or *d* side chain located at such an interface (the knob), KIH packing results in intimate contacts with four side chains from the partner (the hole).

The identities of these four side chains depend on helix orientation. For parallel coiled coils, an *a* knob is flanked by one *a*′, one *g*′, and two *d*′ side chains (Figure 1A). The *a*--*a*′ interaction may be described as "lateral", because a line connecting these two side chains is roughly perpendicular to the long axis of the coiled coil; the two *a*--*d*′ interactions may be described as "vertical". A *d* knob in a parallel coiled coil is flanked by a lateral *d*′ partner, two vertical *a*′ partners and an *e*′ partner (Figure 1B). The KIH juxtapositions are altered in antiparallel coiled-coil dimers. Now an *a* knob has a lateral *d*′ partner, two vertical *a*′ partners, and an *e*′ partner (Figure 1C), while a *d* knob has a lateral *a*′ partner, two vertical *d*′ partners, and a *g*′ partner (Figure 1D).

Lateral *a*--*a*′ and *d*--*d*′ side-chain homopairings have been evaluated for all 20 proteinogenic residues by Hodges et al. in a designed parallel homodimeric coiled coil; these studies necessarily involve mutation of two residues at once (one from each partner).[13] Vinson et al. have used a *heterodimeric* parallel coiled coil to examine lateral pairings; this versatile system allows mutation of a single residue, in contrast to the homodimeric system of Hodges et al., which enables evaluation of more subtle variations in interfacial side-chain packing. Lateral *a*--*a*′ pairings for 10 of the 20 proteinogenic residues have been evaluated, as have a small number of *d*--*d*′ pairings.[14] We have developed an *antiparallel* coiled-coil model system and examined lateral *a*--*d*′ pairings for five proteinogenic residues.[15] This brief summary shows that the data set for lateral side-chain pairings in coiled-coil dimers remains incomplete; nevertheless, the available information shows that lateral neighbor identity at a coiled-coil interface can exert a substantial impact on dimer stability.

Recently, we reported the first study of *vertical* interactions among side chains at a coiled-coil interface, focusing on *a*′--*a*--*a*′ vertical triads in antiparallel dimers (Figure 1C).[16] For a very limited comparison of triads comprising only Leu and/or Ile, we found clear evidence for energetically significant preferences in both discrete experimental model systems and, via bioinformatics analysis, in natural protein structures themselves. Moreover, we observed an interplay between vertical and lateral interactions, for a given knob side chain, in terms of influencing helix partner preferences. Here we report a complementary study focused on *d*′--*d*--*d*′ vertical triads in an antiparallel coiled-coil dimer (Figure 1D). To our surprise, the results indicate that this type of packing interaction exerts substantially less impact on partner preference than do vertical interactions among *a* side chains.

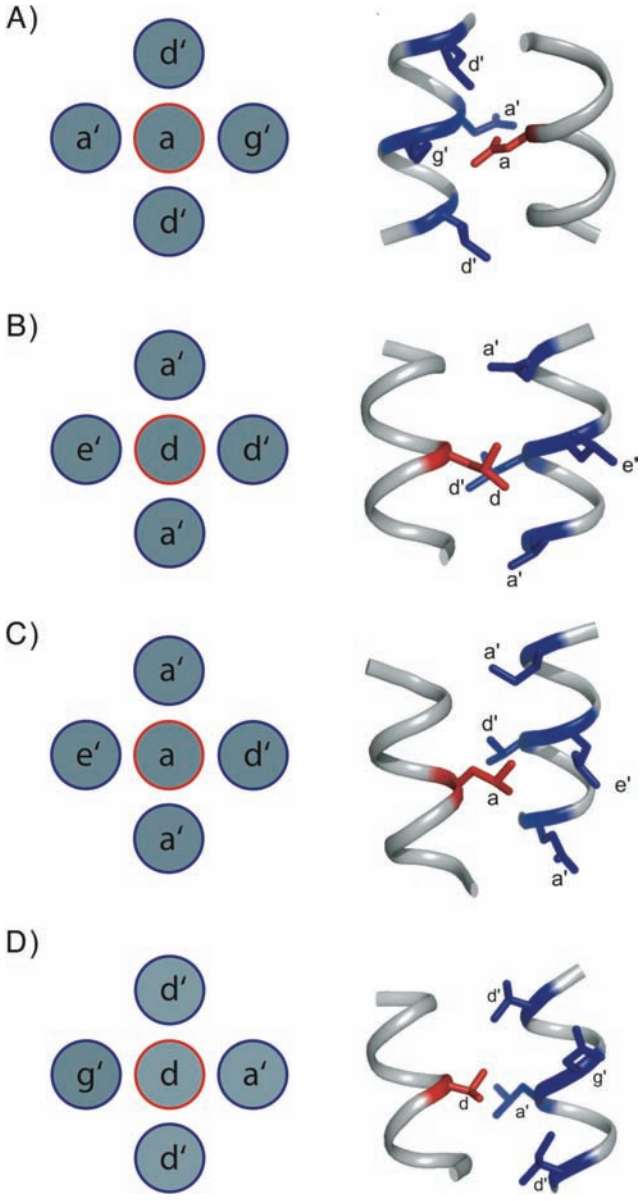


Figure 1. The relationship between the *a* or *d* heptad position (knob) and the four interhelical amino acid side-chain contacts (the hole). Cartoons on the left are projections of the structures on the right viewed from behind (i.e., from the *f* position) of the helix labeled with blue side chains. Knob-into-hole (KIH) interactions in A and B represent a parallel dimer (structures at right from PDB code 2ZTA), while C and D represent an antiparallel dimer (structures at right from PDB code 3GPV).

RESULTS AND DISCUSSION

Experimental Design. We used the thioester exchange (TE) method to assess sequence-dependent variations in antiparallel coiled-coil stability. TE has previously been shown to facilitate rapid evaluation of sequence–stability relationships in small protein folding motifs.[10a,15−17] This technique has been used to explore a variety of tertiary structures, including a designed helix–loop–helix unit in which the tertiary contacts arise from intramolecular antiparallel coiled-coil formation (Figure 2). This system, which forms the basis of the studies reported here, consists of two α-helix-prone segments connected by a flexible linker ($\mathbf{N_T}$-**C**, Figure 2A). The linker contains a thioester bond, which allows $\mathbf{N_T}$-**C** to participate in

dx.doi.org/10.1021/ja208855x | J. Am. Chem.Soc. 2012, 134, 2626−2633

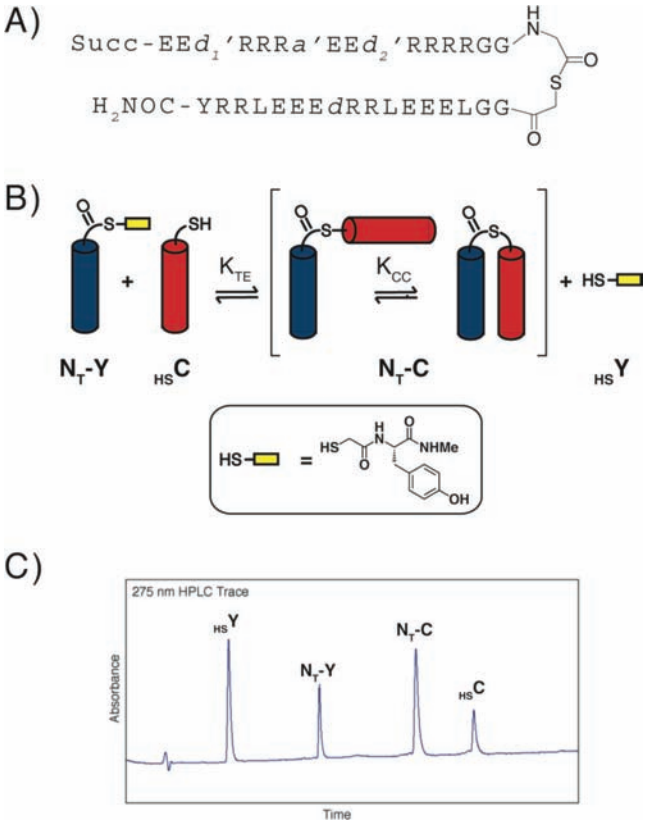


Figure 2. (A) Design and sequence of $\mathbf{N_T}$-**C**; Succ = N-terminal succinyl group. Residues d_1', a', d_2', and d correspond to mutation sites. (B) Thioester exchange process for $\mathbf{N_T}$-**C**. The thioester-thiol pair on the left comprises N- (blue) and C-terminal (red) segments, whereas the pair on the right contains the full-length coiled coil and a small thiol fragment. (C) Representative HPLC chromatogram of a TE assay.

reversible thiol-thioester exchange reactions that achieve equilibrium rapidly in aqueous buffer at pH 7 (Figure 2B). The equilibrium constant for the thiol-thioester exchange (K_{TE}) can be determined by analytical HPLC (Figure 2C) and provides insight on the degree to which the isolated α-helix-prone segments engage one another to form an antiparallel coiled coil.

Analysis of *d*′--*d*--*d*′ Vertical Triads. Four positions in the $\mathbf{N_T}$-**C** sequence were varied for the present studies; these sites are designated a', d, d_1' and d_2' in Figure 2A. Our previous studies of the $\mathbf{N_T}$-**C** system included physical characterization of an analogue, designated $\mathbf{N_A}$-**C**, with $a' = d = d_1' = d_2'$ = Leu, in which the thioester bond was replaced with an amide bond (i.e., the thioglycolic acid residue was replaced with a glycine residue). This replacement was necessary to confer sufficient chemical stability for analytical ultracentrifugation (AUC) studies, which allowed us to determine whether or not the peptide self-associates. (The thioester bond of $\mathbf{N_T}$-**C** itself partially hydrolyzes during the several days required for sedimentation equilibrium studies via AUC.) AUC analysis of $\mathbf{N_A}$-**C** and several mutants indicated that members of this series do not self-associate under conditions similar to those used for TE assays, which allowed us to interpret TE results strictly in terms of intramolecular interactions within the full-length thioesters.[15,16] Circular dichroism (CD) studies of $\mathbf{N_A}$-**C** and mutants confirmed that these molecules are highly folded in α-helical conformations.[15,16] We have now extended these control studies by evaluating the mutant of $\mathbf{N_A}$-**C** with $a' = d$ = Leu and $d_1' = d_2'$ = Ile. Far-UV CD data revealed extensive α-helicity.[17] Sedimentation equilibrium AUC measurements conducted at 50 and 150 μM indicated that this version of $\mathbf{N_A}$-**C** does not self-associate under TE assay conditions.[18]

To evaluate the impact of *d*′--*d*--*d*′ vertical contacts on antiparallel coiled-coil dimer stability, we determined K_{TE} for the 25 versions of $\mathbf{N_T}$-**C** in which $d_1' = d_2'$ = Ile and a' and d are independently varied among the five residues Leu, Ile, Val, Asn, and Ala. For each specific sequence, the previously established relationship $K_{CC} = K_{TE} - 1$ (where K_{CC} is the equilibrium constant for intramolecular antiparallel coiled-coil formation; Figure 2B) allowed us to determine the free energy for intramolecular coiled-coil formation, ΔG_{CC}. The resulting 25 ΔG_{CC} values are shown in Table 1. Also shown in Table 1 are

Table 1. Thermodynamic Data (ΔG_{CC}[a]) Obtained from Thioester Exchange of $\mathbf{N_T}$-C Mutants

	d				
a′	Leu	Ile	Val	Asn	Ala
			$d_1' = d_2'$ = Ile		
Leu	−1.3	−0.9	−0.7	0.0	−0.5
Ile	−1.8	−0.9	−0.6	−0.1	−0.8
Val	−1.1	−0.7	−0.4	0.2	−0.4
Asn	0.3	0.4	0.4	0.5	0.5
Ala	−0.5	−0.5	−0.1	0.1	0.1
			$d_1' = d_2'$ = Leu		
Leu	−1.4	−1.3	−0.9	0.2	−0.4
Ile	−1.7	−1.0	−0.8	−0.1	−0.9
Val	−1.4	−0.8	−0.6	0.0	−0.9
Asn	0.3	0.2	0.4	0.5	0.8
Ala	−0.7	−0.7	−0.5	0.4	0.0

[a]Values are reported in kcal/mol; uncertainty ~±0.1 kcal/mol.

25 ΔG_{CC} values previous determined for the mutants in which $d_1' = d_2'$ = Leu;[15] comparing the two data sets allowed the impact of changing the vertical partners for position *d*, i.e., Leu-*d*-Leu triads vs Ile-*d*-Ile triads, to be assessed. In addition, the 50 ΔG_{CC} values in Table 1 provide insight on interplay between the vertical *d*′ partners and the lateral *a*′ partner in determining antiparallel coiled-coil stability as *d* is varied; a', d_1', and d_2' represent three of the four residues that define the hole surrounding knob *d* (Figure 1D). (It should be noted that in the previous reports on this system the identities of the "primed" and "unprimed" helical segments were reversed; thus, the position we designate here as *d* was previously designated *d*′, and *a*′ here was previously *a*.)

Inspection of Table 1 suggests that for many *a*′--*d* pairings ΔG_{CC} is relatively insensitive to the vertical context in which *d* finds itself (Ile-*d*-Ile vs Leu-*d*-Leu). A different trend was previously observed for vertical *a* triads: many lateral pairs involving the central *a* position had significantly different energetic consequences for antiparallel coiled-coil stability depending on whether the vertical partners were Leu or Ile.[16] Detailed interpretation of the ΔG_{CC} data set in Table 1, however, is complicated by the fact that changing d_1' and d_2' from Ile to Leu leads to multiple changes at the coiled-coil interface, because several interwoven KIH motifs are altered simultaneously.

Previously, we defined a quantity termed the "discrimination energy" as a way of isolating the impact of vertical context on

dx.doi.org/10.1021/ja208855x | J. Am. Chem.Soc. 2012, 134, 2626−2633

antiparallel coiled-coil pairing preferences.[16] This isolation is achieved by considering ΔG_{CC} values for a set of four versions of N_T-C, involving two different residues at position *d* and a constant residue at *a*′. This approach is illustrated in Figure 3

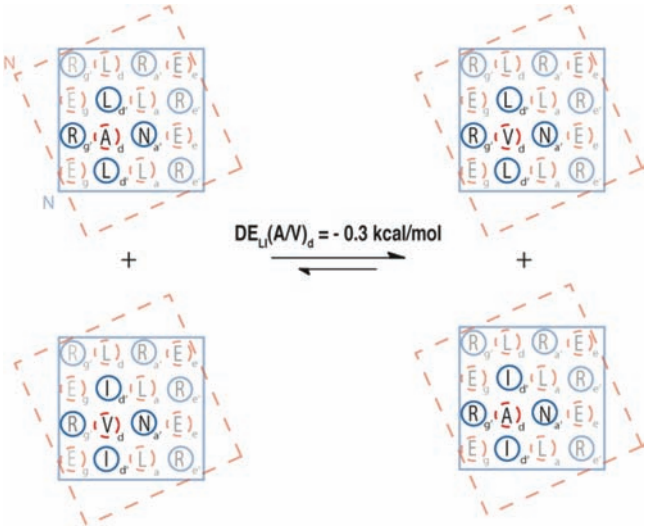


Figure 3. Partial helical net diagrams representing four mutants of N_T-C used to calculate a discrimination energy (DE). The reported DE value was derived from the thermodynamic data in Table 1.

for *d* = Val or Ala and *a*′ = Asn. We use the four relevant ΔG_{CC} values from Table 1 to calculate the Gibbs free energy for the equilibrium between the two different pairs of coiled coils indicated in Figure 3.

The pair of coiled coils on the left have *d* = Ala with vertical Leu partners (i.e., a Leu-Ala-Leu vertical triad) and *d* = Val with vertical Ile partners (Ile-Val-Ile vertical triad). The pair of coiled coils on the right have *d* = Ala with vertical Ile partners (Ile-Ala-Ile vertical triad) and *d* = Val with vertical Leu partners (Leu-Val-Leu vertical triad). Inspection of the helical-net diagrams in Figure 3 shows that this hypothetical equilibrium should depend on only the energetic difference between the two sets of vertical triads centered on the *d* position: Leu-Ala-Leu + Ile-Val-Ile on the left vs Leu-Val-Leu + Ile-Ala-Ile on the right. We refer to the resulting Gibbs free energy value as a discrimination energy (DE) because this quantity indicates the extent to which a preference for one of the two vertical triad pairs could contribute to partner selectivity among antiparallel coiled-coil sequences. We specify this particular discrimination energy with the designation $DE_{LI}(A/V)_d$ to indicate the following: (1) we are using Leu and Ile as the alternative vertical partners; (2) on the left Ala is placed between Leu vertical partners while Val is placed between Ile vertical partners; and (3) the vertical triad involves *d* positions in the heptad sequence repeat.

Table 2 provides the 50 $DE_{LI}(X/Y)_d$ values that can be calculated on the basis of the ΔG_{CC} data in Table 1. (Values are shown only in the upper right portion of each box because $DE_{LI}(X/Y)_d = -DE_{LI}(Y/X)_d$.) The uncertainty in each ΔG_{CC} is estimated to be 0.1 kcal/mol, and we consider that $DE_{LI}(X/Y)_d$ values are significant only if the absolute value is ≥0.4 kcal/mol. In other words, only $DE_{LI}(X/Y)_d$ values above this threshold represent a significant contribution to a preference for one of two possible antiparallel coiled-coil pairings. Only 9 of the 50 $DE_{LI}(X/Y)_d$ values (18%) in Table 2 are above this threshold for significance. This result differs markedly from our earlier

Table 2. Discrimination Energy Values ($DE_{LI}(X/Y)_d$, kcal/mol) Obtained from Thioester Exchange Data of N_T-C Mutants

a′ = Leu	X: Leu	Ile	Val	Asn	Ala
Y: Leu		0.3	0.1	-0.3	-0.2
Ile			-0.2	-0.6	-0.5
Val				-0.4	-0.3
Asn					0.1
Ala					

a′ = Ile	X: Leu	Ile	Val	Asn	Ala
Y: Leu		0.2	0.3	0.1	0.2
Ile			0.1	-0.1	0.0
Val				-0.2	-0.1
Asn					0.1
Ala					

a′ = Val	X: Leu	Ile	Val	Asn	Ala
Y: Leu		-0.2	-0.1	-0.1	0.2
Ile			0.1	0.1	0.4
Val				0.0	0.3
Asn					0.3
Ala					

a′ = Asn	X: Leu	Ile	Val	Asn	Ala
Y: Leu		0.2	0.0	0.0	-0.3
Ile			-0.2	-0.2	-0.5
Val				0.0	-0.3
Asn					-0.3
Ala					

a′ = Ala	X: Leu	Ile	Val	Asn	Ala
Y: Leu		0.0	0.2	-0.5	-0.1
Ile			0.2	-0.5	-0.1
Val				-0.7	-0.3
Asn					0.4
Ala					

finding with *a*′--*a*--*a*′ vertical trials, where 21 of 50 $DE_{LI}(X/Y)_a$ values (42%) were significant by the same metric.[16] This distinction indicates that for antiparallel coiled coils the vertical component of KIH packing at *d* positions is less important in terms of pairing preferences than the vertical component of KIH packing at *a* positions.

Our previous analysis of *a*′--*a*--*a*′ vertical triads revealed that changing the lateral partner of the central *a* residue could alter $DE_{LI}(X/Y)_a$. The new data show an analogous interplay among vertical and lateral partners for a given KIH unit, despite the generally smaller impact of *d* vertical triads relative to *a* vertical triads in terms of antiparallel coiled-coil partner selectivity. For example, $DE_{LI}(N/V)_d$ is significant when the lateral partner (*a*′) for the central *d* position is Leu (−0.4 kcal/mol) or Ala (−0.7 kcal/mol), but not when the lateral partner is Ile, Val, or Asn. Comparable variations are observed for $DE_{LI}(N/I)_d$. For $DE_{LI}(A/I)_d$, significant values are observed when *a*′ = Leu, Val, or Asn, but not Ile or Ala. Thus, the new results strengthen the view that the energetics of interactions between a given knob side chain at a coiled-coil interface and individual side chains within the set that defines the surrounding hole cannot be viewed in a simple pairwise fashion. Instead, we must recognize that the side chains comprising a given hole influence one another in the way that they interact with the knob side chain. We note that $DE_{LI}(X/Y)_d$ values calculated with low stability mutants ($\Delta G_{CC} > 0$) should be regarded with some caution. These DE values are derived from K_{TE} values in the range of 1−2, and in this range small variation in K_{TE} can result in significant variation in ΔG_{CC}.

dx.doi.org/10.1021/ja208855x | J. Am. Chem.Soc. 2012, 134, 2626−2633

We undertook a bioinformatics analysis of antiparallel dimeric coiled coils in the CC+ database (http://coiledcoils.chm.bris.ac.uk)[19] in an effort to determine whether trends detected in our TE studies would be reflected among proteins themselves. Unfortunately, the number of relevant structures turned out to be too small to provide conclusive answers, but some trends were apparent. Table 3 shows results for all *d′--d--d′* vertical triads with both *d′* = Ile, or both *d′* = Leu, and *d* = Leu, Ile, Val, Asn, or Ala.

Table 3. Numbers of Specific *d′--d--d′* Combinations in Natural Antiparallel Coiled Coil Structures[a]

	d′ = d′ = Leu				d′ = d′ = Ile		
d	obsd	o/e	rmsd	d	obsd	o/e	rmsd
Leu	129	1.30	0.98	Leu	18	2.4	0.66
Ile	39	1.36	0.74	Ile	1	0.45	n.d.
Val	45	2.07	0.67	Val	1	0.62	n.d.
Asn	3	0.52	n.d.	Asn	0	n.d.	n.d.
Ala	21	0.61	0.6	Ala	0	n.d.	n.d.

[a]o/e = observed/expected, n.d. = not determined. See Experimental Section for calculation of expected values and rmsd.

These data show that triads in which both *d′* = Ile are far more rare than triads in which both *d′* = Leu; in contrast, our previous bioinformatics analysis of *a′--a--a′* vertical triads revealed substantial numbers for triads in which both *a′* = Ile and for triads in which both *a′* = Leu.[16] For the new data set, when both *d′* = Ile, there appears to be a preference for Leu at the *d* position, but the number of triads in which both *d′* = Ile is so small and the general preference for Leu at *d* positions is so high that it is difficult to interpret this observation. With both *d′* = Leu, there seems to be a very slight preference for Ile at the *d* position, as the ratio of observed triads over expected number of triads is slightly greater for Ile than for Leu at the *d* position (1.36 vs 1.30). However, the difference between *d* = Ile and *d* = Leu was not statistically significant (the two-tailed *p*-value = 0.85 obtained from the χ^2 test). Because the bioinformatics analysis of residue identities was inconclusive, we turned to a structural comparison among selected *d′--d--d′* vertical triads. When all hetero-vertical triads with sequences Leu-Ile-Leu, Ile-Leu-Ile, Leu-Val-Leu, or Leu-Ala-Leu are superposed (average rmsd = 0.66 Å, SD = 0.06 Å), the structures display greater similarity to one another than when all Leu-Leu-Leu homo-vertical triads are superposed (rmsd = 0.98 Å).[18] This structural comparison suggests that side chains in hetero-vertical triads might pack more efficiently than do side chains in homo-vertical triads, as we have noted elsewhere for *a′--a--a′* triads.[16]

We considered hypotheses that would rationalize the remarkable contrast we have observed between *a′--a--a′* triads and *d′--d--d′* triads in the antiparallel coiled-coil dimer. Recently, Grigoryan and DeGrado described a study that probed designable space in coiled-coil structures using the Crick parameterization.[20] In this work, all coiled-coil structures greater than 11 residues in length were culled from the CC+ database and analyzed using the automated program coiled coil Crick parameterization (CCCP). This analysis led the authors to conclude that KIH packing in antiparallel coiled coils is asymmetric; for example, in an *a′--a--a′* triad, *a* is closer to one *a′* than the other. Grigoryan and DeGrado rationalized this observation by noting that the C_α–C_β vectors for the longer *a--a′* separation within a given KIH motif point directly toward one another, and therefore the side chains of these two residues can potentially sterically repel one another. Figure 4A illustrates this geometric consideration by showing the C_α–C_β vectors of a selected *a--a′* pairing (two orthogonal views). The vectors point toward one another, and appear to be on a trajectory to cross. We considered whether C_α–C_β vectors of an analogous *d--d′* pairing would project in a similar geometric arrangement (Figure 4B). In this case the C_α–C_β vectors are nearly parallel but offset so that their trajectories do not appear to cross. This type of observation led us to analyze the distribution of crossing angles between the C_α–C_β vectors for *a--a′* and *d--d′* pairs in all dimeric antiparallel coiled-coil structures in the CC+ database. If the C_α–C_β bond lengths and the distances between the C_α–C_β vectors are roughly constant, which would be expected in a coiled-coil structure, then a more acute angle would indicate that the side chains are more likely to cross, and therefore experience steric repulsion. The results of this analysis are shown as a histogram in Figure 4C.

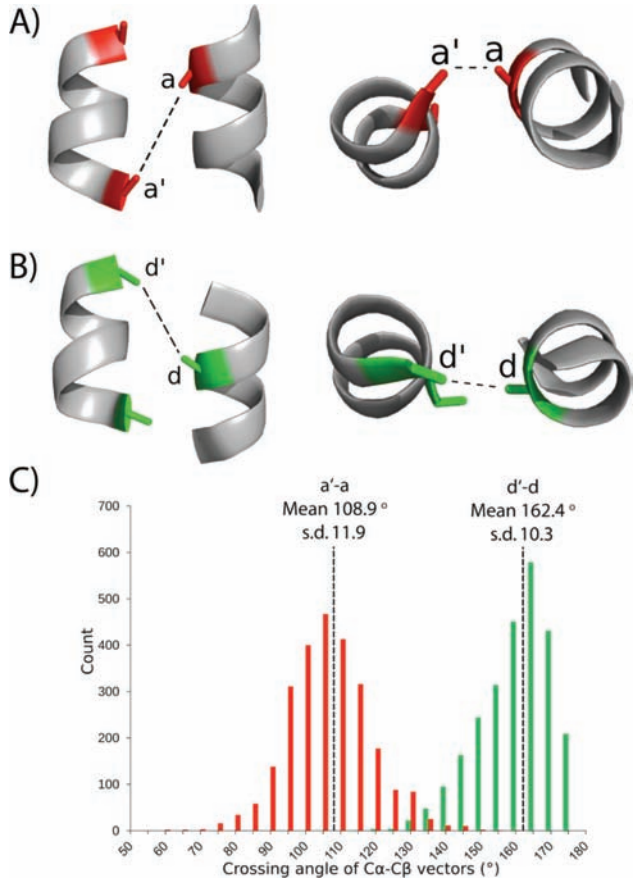


Figure 4. Structural representations of potential steric repulsions (black dotted line) between core residues on opposite chains in the antiparallel helix orientation: (A) *a--a′* (B) *d--d′* (two orthogonal views in each case). Images were generated from PDB entry 2NOV. (C) Histogram of crossing angles for *a--a′* (red) and *d--d′* (green) side-chain pairings for antiparallel vertical interactions, as calculated from the dot product of the C_α–C_β vectors. Mean values are shown for each distribution by a dotted line. The difference in the means of the two distributions is statistically significant (Student's *t* test, $p < 0.01$).

For *a--a′* side-chain pairs, the average angle was 108.9° ± 11.9°, indicating that the vectors cross one another; i.e., steric repulsion between the two residues should be significant. This result indicates changing the side chain at an *a* position in the heptad repeat should have a significant effect on the

dx.doi.org/10.1021/ja208855x | J. Am. Chem.Soc. 2012, 134, 2626–2633

discrimination energy of *a′--a--a′* triads, which is consistent with our previous studies.[16] However, for *d--d′* side-chain pairs, the average angle is significantly larger (162.4° ± 10.3°). This finding indicates that the vectors are almost parallel, and, as mentioned above, these vectors tend to point past one another. This observation suggests that changing the side chain at a *d* position of the heptad repeat should, on average, have a smaller effect on the discrimination energy of *d′--d--d′* triads in an antiparallel coiled coil, relative to the effect of changes at *a* on the discrimination energy of *a′--a--a′* triads. Thus, these structural observations are consistent with the trends revealed in our TE analysis.

Figure 5 provides a three-dimensional impression of the side-chain C_α–C_β vector relationships discussed in the preceding paragraph. The images are based on the C_α–C_β vectors for a set of 10 *a--a′* and 10 *d--d′* pairs; the *knob* C_α–C_β vectors are overlaid for all pairs. These examples were chosen because the angles between the C_α–C_β vectors are close to the mean angle values for the entire sample set (*a--a′* or *d--d′*, as appropriate). The orthogonal views of the vector overlay show that most *a--a′* vectors will cross one another, while the *d--d′* vectors tend to run past one another. Overall, this analysis supports the major conclusion of our study of *a* and *d* positions in the antiparallel coiled-coil dimer motif: the *a′--a--a′* vertical triad is significantly more discriminating than the *d′--d--d′* triad.

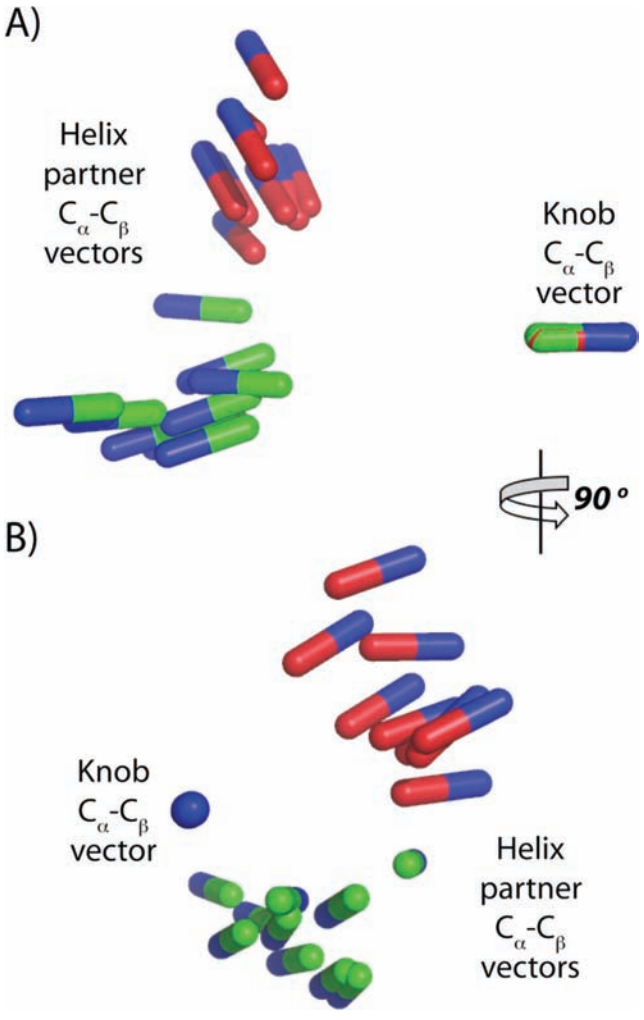


Figure 5. Orthogonal views of the three-dimensional relationship among C_α–C_β vectors in selected *a--a′* and *d--d′* pairs across antiparallel coiled-coil dimer interfaces. C_α atoms are blue; C_β atoms are coded by color (*a--a′* red, *d--d′* green). See text for details.

CONCLUSIONS

We have explored the significance of vertical KIH interactions among side chains at *d* positions in antiparallel coiled-coil dimers. This study complements a previous analysis of *a′--a--a′* vertical triads.[16] Collectively, these studies indicate that vertical interactions can exert a substantial influence on coiled-coil stability and pairing preferences, but that vertical interactions involving *d* sites in the heptad repeat are less sensitive to sequence variations than are vertical interactions involving *a* sites. The data show that for a single KIH motif, comprising a knob side chain from one helix and four hole-defining side chains from the other helix, energetically significant interplay can occur between the vertical and lateral side chains; such effects are detected at both *a* and *d* sites. The existence of this interplay suggests that efforts to decompose helix–helix interaction energies into contributions from pairwise side-chain contacts may yield incomplete understanding of sequence–stability relationships among coiled coils.

EXPERIMENTAL SECTION

General. Peptides were synthesized on solid phase using Fmoc chemistry. Amino acids were activated by 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) and *N*-hydroxybenzotriazole (HOBt). Fmoc-protected α-amino acids with acid-labile side-chain protecting groups, NovaPEG Rink Amide resin and H-Gly-2-Cl-Trt resin, were purchased from Novabiochem. HOBt, *N*,*N*-dimethylformamide (DMF), and *N*,*N*-diisopropylethylamine (DIEA) were purchased from Sigma-Aldrich. HBTU was purchased from Anaspec.

Synthesis of peptide thiols. Peptide thiols were synthesized using microwave irradiation on NovaPEG Rink Amide resin (50 μmol) in Alltech filter tubes. Coupling reactions used 3 equiv of Fmoc-protected amino acid, 3 equiv of HBTU, 3 equiv of HOBt, and 6 equiv of DIEA, with DMF as the solvent. Coupling reactions were heated to 70 °C in a MARS V multimode microwave instrument (2 min ramp to 70 °C, 4 min hold 70 °C) with stirring. Fmoc deprotection reactions used 20% piperidine in DMF. Deprotection reactions were heated to 80 °C in the microwave instrument (2 min ramp to 80 °C, 2 min hold 80 °C) with stirring. After each coupling/deprotection cycle the resin was washed three times with DMF. Upon completion of the synthesis, the peptide was cleaved from the resin, deprotected, and then purified by HPLC. The sequences of peptide thiols are shown in Figure S1.

Synthesis of peptide thioesters. Peptide thioesters were synthesized on H-Gly-2-Cl-Trt resin (50 μmol) on a Symphony automated synthesizer (Protein Technologies, Inc.) at the UW-Madison Biotech Center peptide synthesis facility. Coupling reactions (30–60 min reaction times) used 5 equiv of Fmoc-protected amino acid, 5 equiv of HCTU, and 20 equiv of *N*-methylmorpholine (NMM), with DMF as the solvent. Deprotection steps used 20% piperidine in DMF for 1 × 5 min + 1 × 15 min. Upon completion of the synthesis the protected peptide was cleaved from the resin as a C-terminal acid by reacting the resin with 8:1:1 (dichloromethane:trifluoroethanol:acetic acid) for 1.5 h before precipitating with cold diethyl ether (ca. 50 mL). The solvent was removed under rotary evaporation using a room temperature water bath. Residual acetic acid was removed *in vacuo*. The remaining solid residue was dissolved by the addition of DMF (1.67 mL). The C-terminal acid was converted into the corresponding phenyl thioester by adding DIEA (13.7 μL, 80 μmol), benzene thiol (8.4 μL, 81 μmol), and PyBop (41.6 mg, 80 μmol) to the dissolved peptide and stirring for 1.5 h. The DMF was then removed *in vacuo*, and the resulting protected phenyl thioester peptide was globally deprotected and purified by HPLC. The purified phenyl thioester peptides were then combined with thiol $_{HS}$**Y** (ca. 10

dx.doi.org/10.1021/ja208855x | *J. Am. Chem.Soc.* 2012, 134, 2626–2633

mg) in aqueous solution (pH 7, 50 mM phosphate buffer, 2 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP)), which allowed for TE. After 2 h of TE at room temperature, the exchanged peptide thioester was collected by HPLC. The sequences of peptide thioesters used for this study are shown Figure S2.

Synthesis of N_A-C (d_1' d_2' = Ile). Peptide N_A-C (d_1' d_2' = Ile) was synthesized on NovaPEG Rink Amide resin (50 μmol) on a Symphony automated synthesizer (Protein Technologies, Inc.) at the UW-Madison Biotech Center peptide synthesis facility. Coupling reactions (30−60 min reaction times) used 5 equiv of Fmoc-protected amino acid, 5 equiv of HCTU, and 20 equiv of NMM, with DMF as the solvent. Deprotection steps used 20% piperidine in DMF for 1 × 5 min +1 × 15 min. The sequence for this peptide is shown in Figure S3.

Cleavage, Deprotection, HPLC Purification, and Characterization. Peptides were globally deprotected and cleaved from the resin by stirring the resin in 95% trifluoroacetic acid (TFA), 2.5% water, and 2.5% triisopropylsilane for 2−4 h (for peptide thiols, 2.5% ethanedithiol was added to the cleavage cocktail). Following the reaction period, peptides were precipitated from the TFA solution by addition of cold ether. The precipitated peptide was collected by centrifugation. Ether was decanted, and the peptide pellet was dried under nitrogen. All of the peptides described were purified by reverse-phase HPLC. Identity was confirmed using matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) spectrometry. Purity was determined by analytical HPLC.

Thioester Exchange Assays. TE assays were initiated by mixing approximately equal portions of peptide thiol and peptide thioester in buffer composed of 50 mM sodium phosphate (diluted from a 250 mM, pH 7.0 stock solution) and 2 mM TCEP (diluted from a 20 mM stock solution; included to prevent disulfide formation during the assay). The assay mixtures were allowed to equilibrate for 90 min, after which an aliquot was injected onto a C18 analytical HPLC column (4.6 × 250 mm). Assay components were eluted at a flow rate of 1.0 mL/min using a gradient of acetonitrile in water with 0.1% TFA. Gradients used for separation were optimized in each case to achieve resolution of the four equilibrating species. Equilibrating species were identified by HPLC retention time, and (relatively) quantified by integration of the corresponding HPLC peaks as detected by absorbance at 275 nm. Each species in the TE assay should have the same extinction coefficient at 275 nm, because each contains a single tyrosine residue chromophore. As a result, HPLC peak areas for each species (as detected at 275 nm) are directly proportional to peptide concentration, and can be used to calculate the TE equilibrium constant, K_{TE}. Equilibrium was assumed when the measured K_{TE} did not change as a function of time. HPLC chromatograms of TE assays can be found in the Supporting Information.

Circular Dichroism. CD spectra were recorded on an Aviv Model 420 spectropolarimeter (Aviv Biomedical) using 1 nm bandwidth in 1 mm quartz cells. Samples were prepared by dissolving ca. 1 mg of peptide in H_2O. Concentrations were determined by UV spectrophotometry in a 1 cm quartz cell by diluting the stock solution 20-fold into an 8 M guanidine hydrochloride solution and measuring the absorbance at 275 nm. The absorbance and a calculated extinction coefficient for the peptide were used to determine the concentration of the prepared stock solution.[21] Wavelength scans were taken with 1 nm step sizes. Data were collected for 5 s and averaged at each individual point. Molar ellipticities (θ) were calculated using the equation $\theta = \theta_{obsd}/(10lc)$, where θ_{obsd} is the measured ellipticity (mdeg), l is the length of the cell (cm), and c is the concentration of the peptide. Spectra were corrected for baseline molar ellipticity at 260 nm.

Analytical Ultracentrifugation. Sedimentation equilibrium studies were conducted with a Beckman Optima XLA analytical ultracentrifuge at 25 °C. Solutions of 50 and 150 μM concentrations (~100 μL) of peptide in 50 mM pH 7.0 phosphate buffer were each loaded into a sector of a 12 mm double-sector charcoal-filled Epon centerpiece; ~110 μL of buffer was added to the reference sector. The concentration gradients were monitored at 238 nm. Data were collected at rotor speeds of 18, 27, 36, 45, and 53 krpm. For each speed, equilibrium was assumed when gradients collected 2 h apart were superimposible. For a single, ideal, homogeneous species, a plot of $\ln(c)$ vs r^2 should be linear with a slope proportional to the reduced molecular weight (M_R) of the species. If this condition was met, data from all speeds and concentrations were fit using nonlinear regression analysis to $c_r = BOD + c_o \exp[M(1-\nu\rho)\omega^2(r^2 - r_o^2)/2RT]$, where c_r is the concentration (in absorbance units) at radial position r, c_o is the concentration at an arbitrary reference position r_o (49.5 cm^2 here), ν is the partial specific volume in mL/g, ρ is the solvent density in g/mL, ω is the rotor speed in rad/s, R is the gas constant, T is the temperature in K, and BOD is a baseline absorbance correction to account for non-sedimenting species. Fitting to $c_r = BOD + c_o \exp[M(1-\nu\rho)\omega^2(r^2 - r_o^2)/2RT]$ was done with the BOD as a fitting parameter. $M(1-\nu\rho)$ was fit as a single global constant, from which the weight-average molecular weight was calculated. Global fits to the data at all speeds and concentrations were judged to be adequate by randomness of residuals. Partial specific volumes were calculated on the basis of amino acid composition.[22] A solvent density of 0.9982 g/mL was used.

Bioinformatic Analysis Using CC+. A data set of antiparallel dimeric coiled-coil structures was obtained from the September 2009 release of CC+ (http://coiledcoils.chm.bris.ac.uk/).[19] Proteins were allowed to share no more than 70% sequence identity at the coiled-coil level. As the coiled-coil data are quite sparse, this cutoff is a balance between giving sufficient numbers to perform a meaningful analysis while avoiding bias toward structures with similar sequences, and is in line with our previous work.[16] Coiled-coil length had to be greater than 11 amino acid residues. As structures with two identical chains could report equivalent between-chain interactions, one chain was manually removed from each homodimer to ensure an unbiased data set. For each sequence, knobs-into-holes interactions were identified from the MySQL interface to CC+ using a perl script. Where the vertical d' and d' "hole" residues were both isoleucine or leucine, the numbers of Ile, Leu, Val, Ala, and Asn residues at the d "knob" position were recorded. Expected values were calculated as follows. First, for each "knob" residue type (I, L, V, A, N), the proportion of that residue type found at a d position in a non-redundant coiled-coil data set (two-helix, antiparallel coiled coils culled from CC+ at ≤70% sequence identity) was identified. This defines the probability of finding that residue type at a d position in an antiparallel two-helix coiled coil (I = 0.094, L = 0.325, V = 0.071, A = 0.113, N = 0.019). Second, this probability was multiplied by the number of observed vertical triads (d' = d' = Leu (305) or d' = d' = Ile (23)) to give an expected number of times we would find that residue type in a particular vertical triad, given the numver of such triads identified across the entire coiled-coil data set. For example, for d' = d' = Leu, the expected number of d = Ile is 0.094 × 305 = 28.6. The observed number of residues was then divided by the expected number to give a ratio (e.g., for Leu-Ile-Leu this ratio is 39/28.6 = 1.36). If this ratio is <1, then there are fewer examples observed than would be expected; if this ratio is >1, then there are more examples observed than would be expected. For each type of d'--d--d' interaction, coordinates of the residues participating in each individual interaction of that d'--d--d' type were extracted from the appropriate PDB file and superposed using ProFit (http://www.bioinf.org.uk/software/profit/), which uses an implementation of the McLachlan algorithm.[23] Rmsd values were calculated over side-chain atoms only. For calculation of projected crossing angles for a--a' and d--d' C_α−C_β vector pairs, the vertical interactions for each entry in the coiled-coil data set described above were extracted, regardless of residue identity. Crossing angles were calculated for 2559 a--a' and 2565 d--d' side-chain pairs from the dot product and magnitude of the C_α−C_β vectors using standard routines implemented in PyRosetta.[24]

■ ASSOCIATED CONTENT

Supporting Information

Peptide sequences, peptide characterization, CD and AUC data, TE data, HPLC chromatograms of TE assays, pdb codes with residue assignments, and structural superposition of structures uncovered using CC+. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

d.n.woolfson@bristol.ac.uk; gellman@chem.wisc.edu

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was supported by the NIH (GM061238 to S.H.G.) and the BBSRC of the U.K. (BB/G008833/1 to D.N.W.). We thank Dr. Tom Vincent for maintaining the CC+ database, and Dr. Darrell R. McCaslin for assistance with AUC experiments.

REFERENCES

(1) (a) Grigoryan, G.; Keating, A. E. *Curr. Opin. Struct. Biol.* **2008**, *18*, 477−483. (b) Woolfson, D. N. *Adv. Protein Chem.* **2005**, *70*, 79−112. (c) Mason, J. M.; Arndt, K. M. *ChemBioChem* **2004**, *5*, 170−176. (d) Parry, D. A. D.; Bruce Fraser, R. D.; Squire, J. M. *J. Struct. Biol.* **2008**, *163*, 258−269. (e) Lupas, A. N.; Gruber, M. *Adv. Protein Chem.* **2005**, *70*, 37−38. (f) Lupas, A. *Trends Biochem. Sci.* **1996**, *21*, 375−382. (g) Cohen, C.; Parry, D. A. D. *Proteins* **1990**, *7*, 1−15.

(2) Keating, A. E.; Malashkevich, V. N.; Tidor, B; Kim, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14825−14830.

(3) (a) Grigoryan, G.; Reinke, A. W.; Keating, A. E. *Nature* **2009**, *458*, 859−864. (b) Grigoryan, G.; Keating, A. E. *J. Mol. Biol.* **2006**, *355*, 1125−1142. (c) Havranek, J. J.; Harbury, P. B. *Nat. Struct. Biol.* **2003**, *10*, 45−52. (d) Barth, P.; Schoeffler, A.; Alber, T. *J. Am. Chem. Soc.* **2008**, *130*, 12038−12044. (e) Mendes, J.; Guerois, R.; Serrano, L. *Curr. Opin. Struct. Biol.* **2002**, *12*, 441−446.

(4) (a) Boyle, A. L.; Woolfson, D. N. *Chem. Soc. Rev.* **2011**, *40*, 4295−4306. Apostolovic, B.; Danial, M.; Klok, H. A. *Chem. Soc. Rev.* **2010**, *39*, 3541−3575.

(5) (a) Marsden, H. R.; Kros, A. *Angew. Chem., Int. Ed.* **2010**, *49*, 2988−3005. (b) Bromley, E. H.; Channon, K.; Moutevelis, E.; Woolfson, D. N. *ACS Chem. Biol.* **2008**, *3*, 38−50.

(6) (a) Hodges, R. S.; Sodak, J.; Smillie, L. B.; Jurasek, L. *Cold Spring Harbor Symp. Quant. Biol.* **1972**, *37*, 299−310. McLachlan, A. D.; Stewart, M. *J. Mol. Biol.* **1975**, *98*, 293−304.

(7) Crick, F. H. S. *Acta Crystallogr.* **1953**, *6*, 689−697.

(8) (a) Harbury, P. B.; Zhang, T.; Kim, P. S. *Science* **1993**, *262*, 1401−1407. (b) Woolfson, D. N.; Alber, T. *Protein Sci.* **1995**, *4*, 1596−1607.

(9) (a) O'Shea, E. K.; Klemm, J. D.; Kim, P. S.; Alber, T. *Science* **1991**, *254*, 539−544. (b) Lumb, K. J.; Kim, P. S. *Biochemistry* **1995**, *34*, 8642−8648.

(10) (a) Steinkruger, J. D.; Woolfson, D. N.; Gellman, S. H. *J. Am. Chem. Soc.* **2010**, *132*, 7586−7588. (b) McClain, D. L.; Gurnon, D. G.; Oakley, M. G. *J. Mol. Biol.* **2002**, *324*, 257−270. (c) Campbell, K. M.; Lumb, K. J. *Biochemistry* **2002**, *41*, 4866−4871. (d) Gonzalez, L.; Woolfson, D. N.; Alber, T. *Nat. Struct. Biol.* **1996**, *3*, 1011−1018. (e) Diss, M. L.; Kennan, A. J. *J. Am. Chem. Soc.* **2008**, *130*, 1321−1327.

(11) (a) O'Shea, E. K.; Rutkowski, R.; Stafford, W. F. III; Kim, P. S. *Science* **1989**, *245*, 646−648. (b) O'Shea, E. K.; Rutkowski, R.; Kim, P. S. *Cell* **1992**, *68*, 699−708. (c) O'Shea, E. K.; Lumb, K. J.; Kim, P. S. *Curr. Biol.* **1993**, *3*, 658−667. (d) Graddis, T. J.; Myszka, D. G.; Chaiken, I. M. *Biochemistry* **1993**, *32*, 12664−12671. (d) Vinson, C. R.; Hai, T.; Boyd, S. M. *Genes Dev.* **1993**, *7*, 1047−1058. (e) Krylov, D.; Mikhailenko, I.; Vinson, C. *EMBO J.* **1994**, *13*, 2849−2861. (f) Krylov, D.; Barchi, J.; Vinson, C. *J. Mol. Biol.* **1998**, *279*, 959−972. (g) Monera, O. D.; Kay, C. M.; Hodges, R. S. *Biochemistry* **1994**, *33*, 3862−3871. (h) Litowski, J. R.; Hodges, R. S. *J. Biol. Chem.* **2002**, *277*, 37272−37279. (i) Ryan, S. J.; Kennan, A. J. *J. Am. Chem. Soc.* **2007**, *129*, 10255−10260.

(12) Reinke, A. W.; Grant, R. A.; Keating, A. E. *J. Am. Chem. Soc.* **2010**, *132*, 6025−6031. Chen, T. S.; Reinke, A. W.; Keating, A. E. *J. Mol. Biol.* **2011**, *408*, 304−320.

(13) (a) Tripet, B.; Wagschal, K; Lavigne, P.; Mant, C. T.; Hodges, R. S. *J. Mol. Biol.* **2000**, *300*, 377−402. (b) Wagschal, K.; Tripet, B.; Lavigne, P.; Mant, C.; Hodges, R. S. *Protein Sci.* **1999**, *8*, 2312−2329.

(14) (a) Acharya, A.; Rishi, V.; Vinson, C. *Biochemistry* **2006**, *45*, 11324−11332. (b) Acharya, A.; Ruvinov, S. B.; Gal, J.; Moll, J. R.; Vinson, C. *Biochemistry* **2002**, *41*, 14122−14131. (b) Moitra, J.; Szilak, L.; Krylov, D.; Vinson, C. *Biochemistry* **1997**, *36*, 12567−12573.

(15) Hadley, E. B.; Gellman, S. H. *J. Am. Chem. Soc.* **2006**, *128*, 16444−16445.

(16) Hadley, E. B.; Testa, O. D.; Woolfson, D. N.; Gellman, S. H. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 530−535.

(17) (a) Woll, M. G.; Gellman, S. H. *J. Am. Chem. Soc.* **2004**, *126*, 11172−11174. (b) Woll, M. G.; Hadley, E. B.; Mecozzi, S.; Gellman, S. H. *J. Am. Chem. Soc.* **2006**, *128*, 15932−15933. (c) Hadley, E. B.; Witek, A. M.; Freire, F.; Peoples, A. J.; Gellman, S. H. *Angew. Chem., Int. Ed.* **2007**, *46*, 7056−7059. (d) Pomerantz, W. C.; Hadley, E. B.; Fry, C. G.; Gellman, S. H. *Chembiochem* **2009**, *10*, 2177−2181. (e) Price, J. L.; Hadley, E. B.; Steinkruger, J. D.; Gellman, S. H. *Angew. Chem., Int. Ed.* **2010**, *49*, 368−371.

(18) See Supporting Information.

(19) Testa, O.; Moutevelis, E.; Woolfson, D. *Nucleic Acids Res.* **2009**, *37*, D315−D322.

(20) Grigoryan, G.; DeGrado, W. F. *J. Mol. Biol.* **2011**, *405*, 1079−1100.

(21) Edelhoch, H. *Biochemistry* **1967**, *6*, 1948−1954.

(22) Reynolds, J. A.; McCaslin, D. R. *Methods Enzymol.* **1985**, *117*, 41−53.

(23) McLachlan, A. *Acta Crystallogr. A* **1982**, *38*, 871−873.

(24) Chaudhury, S.; Lyskov, S.; Gray, J. J. *Bioinformatics* **2010**, *26*, 689−691.