Dramatic Structural and Thermodynamic Consequences of Repacking a Protein's Hydrophobic Core

**Mark A. Willis,* Barney Bishop,* Lynne Regan,*†§
and Axel T. Brunger*‡§∥**
*Department of Molecular Biophysics and Biochemistry
†Department of Chemistry
‡The Howard Hughes Medical Institute
Yale University
New Haven, Connecticut 06520

Summary

Background: Rop is an RNA binding, dimeric, four-helix bundle protein with a well-defined, regular hydrophobic core ideally suited for redesign studies. A family of Rop variants in which the hydrophobic core was systematically redesigned has previously been created and characterized.

Results: We present a structural and thermodynamic analysis of Ala_2Ile_2-6, a variant of Rop with an extensively redesigned hydrophobic core. The structure of Ala_2Ile_2-6 reveals a completely new fold formed by a conformational "flip" of the two protomers around the dimeric interface. The free-energy profile of Ala_2Ile_2-6 is also very different from that of wild-type Rop. Ala_2Ile_2-6 has a higher melting temperature than Rop, but undergoes a slightly smaller free-energy change on unfolding.

Conclusions: The structure of Ala_2Ile_2-6, along with molecular modeling results, demonstrate the importance of tight packing of core residues and the adoption of favorable core side chain rotamer values in determining helix-helix interactions in the four-helix bundle fold. Structural disorder at the N and C termini of Ala_2Ile_2-6 provides a basis for the large differences in the enthalpy and entropy of Ala_2Ile_2-6 folding compared with wild-type Rop.

Introduction

A major goal of protein design and engineering is to construct stable protein frameworks on which to engineer functionality. An understanding of the factors that contribute to protein stability and define structure is therefore essential for successful protein design. It has long been recognized that a key feature of natural proteins, and the major driving force in their folding, is the existence of a solvent-inaccessible hydrophobic core. These cores are characterized both by the hydrophobic nature of the residues and by their appropriate placement to ensure an optimal combination of favorable side chain stereochemistry and high packing density [1]. Nonpolar residues in the core contribute to protein stability in a number of ways: Enthalpic contributions from van der Waals interactions between closely packed residues, and both enthalpic and entropic contribu-

tions arising from the desolvation of these residues upon folding [2].

Because of the importance of the hydrophobic core to the stability of natural proteins, it has been the focus of a number of studies. Several groups have examined the structural and thermodynamic consequences of mutations that repack, overpack, or underpack the hydrophobic core [3–11]. Mutations that slightly overpack the core can sometimes be tolerated, while cavity formation due to underpacking generally results in decreased stability. It has also been demonstrated that the identity of core residues in parallel coiled-coils can specify different oligomerization states [12, 13]. Others have applied principles of hydrophobic versus hydrophilic residue patterning to the de novo design of simple proteins [13–16]. In addition, both experimental selection methods [17] and computational algorithms [18–22] have been developed to identify combinations of residues that will repack a hydrophobic core given a desired backbone conformation.

To investigate the structural and thermodynamic consequences of systematically repacking the entire hydrophobic core of a simple protein, we have focused on the four-helix bundle protein Rop (also known as ROM). Rop is an RNA binding protein that is involved in regulation of the copy number of ColE1 plasmids in *Escherichia coli*. Rop is an attractive model system for a number of reasons: It is a small, soluble protein devoid of propeptide sequence, disulfide bonds, proline residues, and cofactors. It is also a homodimer of two helix-turn-helix protomers with a well-defined, regular, hydrophobic core comprised of eight stacked layers each of which is formed by four hydrophobic residues at the "a" and "d" positions of a heptad repeat (Figure 1a). Both the X-ray crystal structure [23] and the nuclear magnetic resonance structure [24] of wild-type Rop have been reported, and there is a simple in vitro assay for Rop's RNA binding activity [25].

A family of Rop variants in which the hydrophobic core has been systematically redesigned has been created and characterized [9, 10]. They are named for the identity of residues at the "a" and "d" positions in a layer and the number of repacked layers. These variants display a wide range of structural and thermodynamic properties. Among those with a conserved core volume are variants with wild-type RNA binding affinity (suggesting a structure close to that of Rop) that have significantly enhanced thermal stability and nativelike thermodynamic properties (Ala_2Leu_2-6 and Ala_2Leu_2-8); variants with the above thermodynamic properties that do not bind RNA (Ala_2Ile_2-6 and Leu_2Ala_2-6); and variants that maintain nativelike properties, but have reduced thermal stability (Ala_2Met_2-8).

Variants with significantly altered core volumes and hydrophobicity include proteins with overpacked cores that display extreme resistance to thermal and chemical denaturation and non-nativelike thermodynamic properties (Leu_4-8), and proteins in which the hydrophobic surface area and core volume are so low that the helices fail to fold and associate (Ala_4-8, Ala_2Val_2-8).

Here, we present a high-resolution structural characterization and detailed thermodynamic analysis of a protein that has

§To whom correspondence should be addressed (e-mail: lynne@csb.yale.edu [L. R.] and axel.brunger@stanford.edu [A. T. B.]).
∥Present address: Howard Hughes Medical Institute and Departments of Molecular and Cellular Physiology, Neurology and Neurological Sciences, and Stanford Synchrotron Radiation Laboratory, Stanford University, Stanford, California 94305.

(a)

(b)
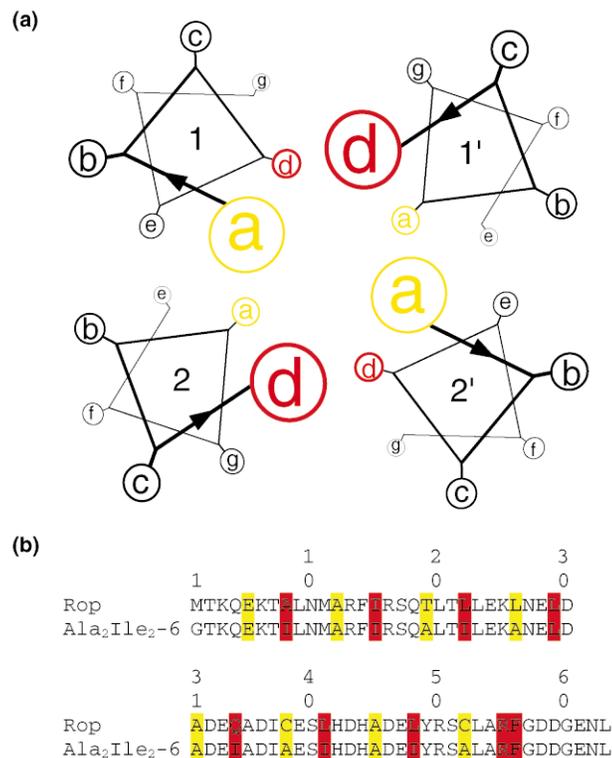
Figure 1. Design of Ala$_2$Ile$_2$-6

(a) An α-helical wheel diagram (looking down the long axis) of the heptad repeat of Rop. The "a" (yellow) and "d" (red) residues form the hydrophobic core and are the residues mutated in the repacked protein Ala$_2$Ile$_2$-6. Two layers of the core are shown. Helices from protomer A are designated 1 and 2, and the protomer B helices are labeled 1′ and 2′. Arrows indicate the direction of the polypeptide chain from the N terminus to the C terminus.
(b) Sequence alignment of Rop and Ala$_2$Ile$_2$-6 with the residue number placed above every tenth residue. The "a" and "d" residues are colored to match the diagram. To create Ala$_2$Ile$_2$-6, residues in the "a" and "d" positions of Rop were changed to alanine and isoleucine, respectively. The outermost layer at each end of the four-helix bundle, consisting of residues 5, 29, 31, and 56, were not changed. Residue 56, in the "e" position of the heptad repeat, acts as a "d" residue by packing its side chain into the appropriate core position.

lost RNA binding activity, but which has enhanced, nativelike thermal stability.

Results and Discussion

Design and Initial Characterization
Rop variants were generated by replacing the core residues (Figure 1a) of the heptad repeat with a regular pattern of hydrophobic amino acids. One of the first and most conservative mutants created was Ala$_2$Leu$_2$-6 [9], in which the middle six layers of the hydrophobic core incorporate alanines in the "a" positions and leucines in the "d" positions. Ala$_2$Leu$_2$-6 has nativelike structural and thermodynamic properties, binds RNA with wild-type affinity, and has a significantly higher melting temperature than that of wild-type Rop. To investigate the importance of side chain geometry in packing the core of Rop, we created Ala$_2$Ile$_2$-6 with alanine in the "a" positions and iso-leucine in the "d" positions of the middle six layers of the core (Figure 1b). Because isoleucine and leucine share similar side

(a)

(b)
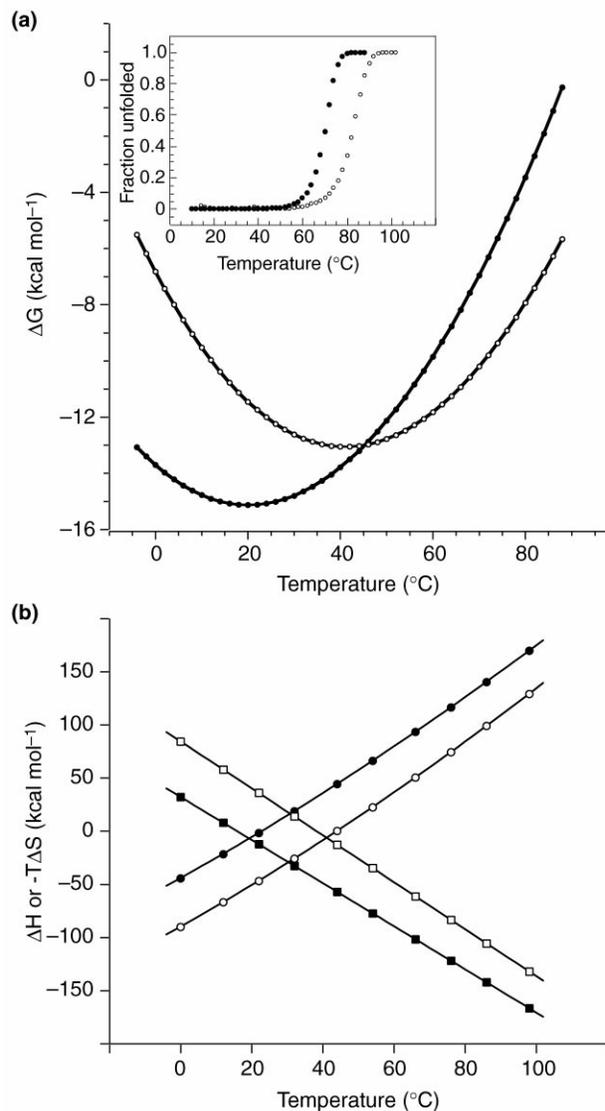
Figure 2. Thermodynamic Comparison of Rop and Ala$_2$Ile$_2$-6

(a) Thermal stability profile (ΔG versus T) and representative thermal denaturations (inset) of Rop (solid circles) and Ala$_2$Ile$_2$-6 (open circles).
(b) Calculated values of ΔH (squares) and -TΔS (circles) as a function of temperature for Rop (solid) and Ala$_2$Ile$_2$-6 (open).

chain volume and hydrophobicity, any structural and thermodynamic differences between Ala$_2$Leu$_2$-6 and Ala$_2$Ile$_2$-6 are, therefore, primarily a consequence of the different side chain stereochemistries of leucine and isoleucine.

The initial characterization of Ala$_2$Ile$_2$-6 demonstrated that it is highly helical with a circular dichroism (CD) spectrum similar to that of wild-type Rop. Furthermore, Ala$_2$Ile$_2$-6 maintained nativelike thermodynamic properties illustrated, for example, by a cooperative and reversible thermal unfolding transition (Figure 2a, inset) accompanied by a large change in the heat capacity (Table 1). Ala$_2$Ile$_2$-6 was shown to be a dimer by both sedimentation equilibrium centrifugation and multiangle laser light scattering measurements (data not shown). However, the electromobility shift assay for protein-RNA interaction [25] demonstrated that Ala$_2$Ile$_2$-6 had completely lost the ability to

Table 1. Thermodynamic Results

	Rop	Ala$_2$Ile$_2$-6
T$_{m,app}$	69.9 $\pm$ 0.3	82.6 $\pm$ 0.1
T$_{max}$	20 $\pm$ 8	42 $\pm$ 4
ΔH$_m$	$-$109.3 $\pm$ 2.3	$-$97.7 $\pm$ 2.2
ΔS$_m$	$-$0.298 $\pm$ 0.007	$-$0.254 $\pm$ 0.006
ΔG$_{max}$	$-$15.1 $\pm$ 1.2	$-$13.1 $\pm$ 0.7
ΔC$_p$	$-$2.02 $\pm$ 0.41	$-$2.21 $\pm$ 0.35
ΔH$_{293K}$	$-$8.33	40.3
$-$TΔS$_{293K}$	$-$5.52	$-$50.7
ΔG$_{293K}$	$-$13.8	$-$11.2
ΔASA$_{polar}$	4,856	4,917
ΔASA$_{nonpolar}$	8,015	8,001
ΔC$_{p,calc}$	$-$1.67	$-$1.65
ΔH$_{293K,calc}$	$-$19.4	$-$21.7

ΔH$_m$ and ΔS$_m$ are calculated at the apparent T$_m$: T$_{m,app}$. ΔG$_{max}$ is calculated at the temperature of maximum stability, T$_{max}$. For comparison, the values of ΔH$_{293K}$, $-$TΔS$_{293K}$, and ΔG$_{293K}$ have been calculated at 293 K. The changes in the accessible polar (ΔASA$_{polar}$) and nonpolar (ΔASA$_{nonpolar}$) surface area upon folding have been calculated using ACCESS [58] and are used in the structure-based calculations [31] of ΔC$_p$ (calc.) and ΔH$_{293K}$ (calc.). Errors are based on triplicate measurements. Units are °C for T$_{m,app}$ and T$_{max}$, and Å^2 for both ΔASA calculations. All ΔG, and $-$TΔS values are in kcal mol^{-1}, and all ΔS and ΔC$_p$ values are in kcal mol^{-1} K^{-1}.

bind RNA (data not shown). To obtain detailed structural insight into how the protein adapts to the changes in core packing (and to determine the structural basis for the loss of RNA binding activity) we performed a crystallographic structure determination of Ala$_2$Ile$_2$-6.

Structure of Ala$_2$Ile$_2$-6

The structure of Ala$_2$Ile$_2$-6, containing a single dimer in the asymmetric unit, was solved in the C2 crystal form to 1.9 Å resolution using a combination of MAD and SIRAS phasing methods (Tables 2–5). This structure was then used as a molecular replacement model to solve the 2.25 Å structure of three molecules of Ala$_2$Ile$_2$-6 in the P3$_2$ space group (Tables 2, 3, 5).

The structure of Ala$_2$Ile$_2$-6 reveals a dramatic change in topology from that of wild-type Rop (Figure 3) and from a point mutant of Rop that adopts a different fold [26]. Ala$_2$Ile$_2$-6 is a dimer of two helix-turn-helix protomers that form an antiparallel four-helix bundle; the dimeric interface is transformed, however, by a 180° flip of one protomer around an axis normal to this interface. This reorientation changes the fold of the protein, placing both turns of Ala$_2$Ile$_2$-6 at one end of the four-helix bundle and the two N and two C termini at the other. This structural flip also explains the loss of RNA binding activity.

Table 2. Crystal Forms

Space Group	Crystal #	a (Å)	b (Å)	c (Å)	β (°)	Dimers/a.s.u.
C2 Native	1	82.62	36.49	47.73	121.21	1
C2 K$_2$PtCl$_4$	2	82.89	35.52	48.21	120.65	1
P3$_2$ Native	3	73.09	73.09	65.92	120	3

The RNA binding surface of Rop is formed primarily by residues along the helix 1/helix 1′ face [27]. The opposite face of the protein (the 2/2′ face) is a highly negatively charged surface that may aid in the orientation of Rop to the RNA complex [27]. In Ala$_2$Ile$_2$-6, these two faces are both disrupted and replaced by two new faces (1/2′ and 1′/2). The four-helix bundle is still all antiparallel, however, because the rotation that swaps the positions of helices 1′ and 2′ of protomer B also switches their orientation with respect to protomer A.

The large change in the overall fold of Ala$_2$Ile$_2$-6 is accompanied by a significant change in protomer structure. Helices 1 and 2 in the Rop protomer are aligned such that a plane intersecting the α-carbons of the four "a" and "d" residues in a given layer is roughly normal to the axis of the four-helix bundle. In Ala$_2$Ile$_2$-6, however, helix 2 is translated down the four-helix bundle with respect to helix 1 by approximately half a turn (Figure 3b). This changes the packing environment of the whole core, replacing Rop's eight planar layers with a uniform pattern of staggered layers (Figure 4a). This creates a more regular pattern of knobs and holes along the dimeric interface and allows for a slightly more compact four-helix bundle (Figure 4b). The slight bend in helix 1 at Phe-14 (Figure 3a) is more pronounced in Ala$_2$Ile$_2$-6 than in Rop, and facilitates the close packing along the length of the four-helix bundle.

Differences in the alignment of close-packed antiparallel helices have been analyzed previously [28]. Gernert et al. attribute the preference of the "aligned" or "offset" helical register to the placement of small residues in the "d" or "a" positions, respectively, of the heptad repeat. After accounting for a difference in the classification of residues in the heptad repeat of Rop (their "d" position is our "a" position), it is clear that Ala$_2$Ile$_2$-6 does not conform to the observed helical packing patterns [28]. Rop and Ala$_2$Ile$_2$-6 differ in having aligned versus offset helical packing even though both proteins have small residues at the "a" positions. The reason for the switch in packing arrangements appears to be due to the switch from leucines to isoleucines at the "d" position. The core isoleucines point directly across the dimeric interface to interact with the isoleucines on the diagonally apposed helix (Figure 5)—there

Table 3. Crystallographic Data[a]

Crystal #	1	2	2	2	3
Data set	C2-native	C2-Pt1	C2-Pt2	C2-Pt3	P3$_2$-native
Wavelength (Å)	1.54128	1.54128	1.07157	1.08095	1.54128
D$_{min}$ (Å)	1.9	2.0	1.9	1.9	2.25
Unique reflections (#)	9,460	15,892	18,612	17,749	18,055
Redundancy	6.7	3.4	3.3	1.4	5.0
Completeness (%)	96.9 (84.0)	99.6 (99.0)	99.6 (99.9)	95.0 (95.2)	93.2 (63.5)
Average I/σ	12.6	18.3	11.3	12.1	8.1
R$_{sym}$[b] (%)	5.5 (23.8)	4.0 (20.3)	5.5 (29.9)	3.8 (35.8)	8.5 (27.2)
R$_{iso}$[c] (%)		33.8	31.2	32.7	

[a] Values in parentheses are for the high-resolution bin. Data sets for crystals 1 and 3 were processed nonanomalously, whereas those for crystal 2 were processed anomalously.

[b] R$_{sym}$ = $\Sigma_h\Sigma_i|I_i(h) - \langle I(h)\rangle|/\Sigma_h\Sigma_i|I_i(h)|$ where I$_i$(h) is the ith measurement and $\langle I(h)\rangle$ is the mean of all measurements of I(h) for Miller indices h.

[c] R$_{iso}$ = $\Sigma_h||F_{PH}| - F_P|/\Sigma_h|F_P|$, where |F$_{PH}$| and |F$_P$| are the measured structure factor amplitudes of the derivative and native structures.

Table 4. Phasing Statistics

	Observed Diffraction Ratios[a]			Phasing Power (PP)[b]	
MAD Data Set	C2-Pt1	C2-Pt2	C2-Pt3	+Friedel Mate	−Friedel Mate
C2-Pt1	0.0711	0.1261	0.1703	(Reference)	1.28
C2-Pt2		0.0899	0.0945	2.35	2.65
C2-Pt3			0.0832	2.03	2.10

Overall FOM = 0.715 (0.459); SIRAS PP_{iso} = 1.29 (0.82); PP_{ano} = 1.31 (0.58); Overall FOM = 0.510 (0.398).

[a] Values are $<(\Delta|F|^2>^{1/2}/<|F|^2>^{1/2}$, where $\Delta|F|$ is the dispersive (off-diagonal elements) or Bijvoet difference (diagonal elements), computed between 42 and 1.9 Å resolution.

[b] MAD phasing power is defined as $[<|F_D| - |F_N||^2>/\int_\phi P(\phi)(||F_N|e^{i\phi} + \Delta F_H| - |F_D|)^2 d_\phi]^{1/2}$. $P_{(\phi)}$ is the experimental phase probability distribution, F_N represents the structure factors at the reference wavelength (C2-Pt1), F_D represents the structure factors wavelength i (C2-Pti) or its Friedel mate, and ΔF_H is the difference in heavy atom structure factors between the two wavelengths. SIRAS phasing power is the same except that the differences are either isomorphous or anomalous (PP_{iso} and PP_{ano}, respectively).

Table 5. Crystallographic Model Refinement Statistics

Crystal Form	C2	P3$_2$
Resolution range (Å)	41–1.9	46–2.25
Unique reflections (#)	18,445	34,854
R_{work}[a]	20.4	23.3
R_{free}[b]	23.7	28.2
Luzzati coordinate error (Å)	0.23	0.31
Cross-validated Luzzati coordinate error (Å)	0.27	0.38
Refined non-H, nonsolvent atoms (#)	867	2,617
Solvent molecules (#)	98	111
Bond-length deviation (Å)	0.010	0.006
Bond-angle deviation (°)	1.19	0.94
Improper-angle deviation (°)	0.81	0.56
Dihedrals (°)	15.7	15.2
Average B factor (Å^2)	23.7	39.1
Minimum B factor (Å^2)	3.9[c]/7.5	5.0
Maximum B factor (Å^2)	66.4	90.3
Residues in core ϕ–ψ region (%)	96.2	96.2
Residues in disallowed regions (%)	0.0	0.0

[a] $R_{work} = \Sigma(|F_{obs}| - k|F_{calc}|)/\Sigma|F_{obs}|$ for the set of reflections remaining after the test set (10%) has been removed.

[b] R_{free} is the R value obtained for a test set of reflections, consisting of a randomly selected 10% subset of the diffraction data, not used during refinement.

[c] The 3.9 Å^2 is for a calcium ion on a special position. The lowest B factor for a protein atom is 7.5.

is little interaction with the helices on either side. Thus the protomer structure of Ala$_2$Ile$_2$-6 is specified more by the dimerization of the protein than the close packing of adjacent helices.

By using an offset packing arrangement, Ala$_2$Ile$_2$-6 alternates the knobs of isoleucine interactions between helices 1 and 1′ with the isoleucine interactions of helices 2 and 2′. This arrangement results in an optimal packing environment of knobs and holes along the dimeric interface (Figure 6). By shifting helix 2 a half-turn toward the N terminus of helix 1, the isoleucine knobs on both helices merge just enough to create a low ridge between them. The ridges of the resulting "dog bone-shaped" protrusions on protomer A lie across the ridges connecting the matching protrusions on protomer B, such that the knobs fit neatly into the holes between the bone-shaped knobs on the protomer across the dimeric interface. This packing arrangement creates a more regular and densely packed core than the wild-type protein.

The Two Crystal Forms of Ala$_2$Ile$_2$-6

The four dimers of Ala$_2$Ile$_2$-6 in the two crystal forms provide a means for us to evaluate the conformational flexibility of

Ala$_2$Ile$_2$-6 and the effects of crystal packing. Overall, the dimers are very similar. In the P3$_2$ crystal structure the three dimers related by noncrystallographic symmetry can be superimposed with a backbone root mean square (RMS) deviation of less than 0.33 Å. Each dimer is located on a 3-fold screw axis, and the symmetry creates three propellerlike stacks that interact via residues near the turns of each helical bundle. There are a few disordered residues at some of the helix termini, but the dimers overlay very well (Figure 7a).

The structure of Ala$_2$Ile$_2$-6 in the C2 space group is very similar to the P3$_2$ dimers (0.71 Å backbone RMS deviation) despite some differences at the helical termini due to crystal packing contacts (Figure 7a). Calcium, essential for formation of both crystal forms of Ala$_2$Ile$_2$-6, is involved in completely different interactions. A binuclear calcium site at the C terminus
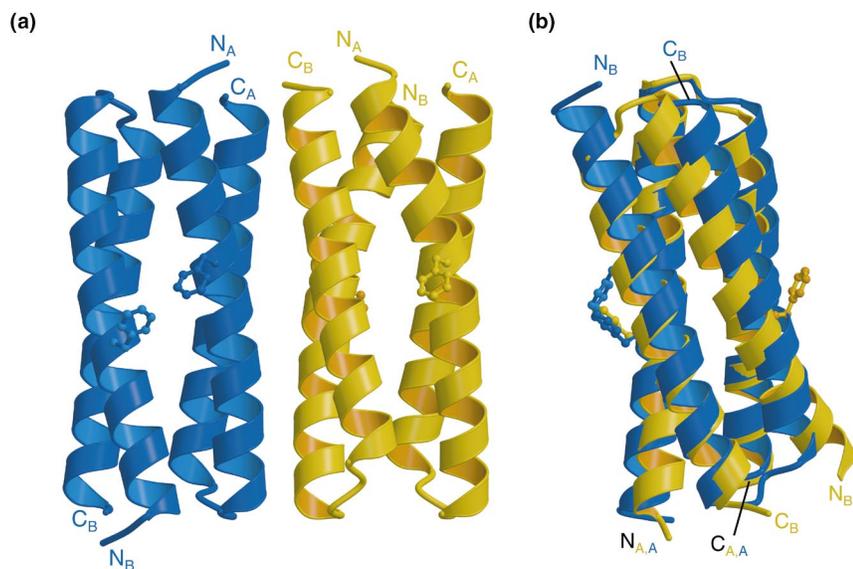
Figure 3. A Comparison of the Rop and Ala$_2$Ile$_2$-6 Structures

Rop is shown in blue and Ala$_2$Ile$_2$-6 is shown in yellow. The labels for the polypeptide termini use a subscript to denote either protomer A or B. The key phenylalanine at position 14, at the center of the RNA binding interface, is shown in ball-and-sticks on each protomer to aid viewer orientation.

(a) Side-by-side view showing the loops of Ala$_2$Ile$_2$-6 at the same end of the four-helix bundle and the splitting of the Helix1-Helix1′ binding face (pointing out of the page) of wild-type Rop.

(b) Overlay view with protomer A (in the back) of Ala$_2$Ile$_2$-6 best fit to the backbone of protomer A of Rop. Helix 2 of the Ala$_2$Ile$_2$-6 protomers are shifted by approximately half a turn toward the C terminus with respect to Rop.

(a)

(b)

8.2 Å

9.5 Å

7.4 Å

8.5 Å
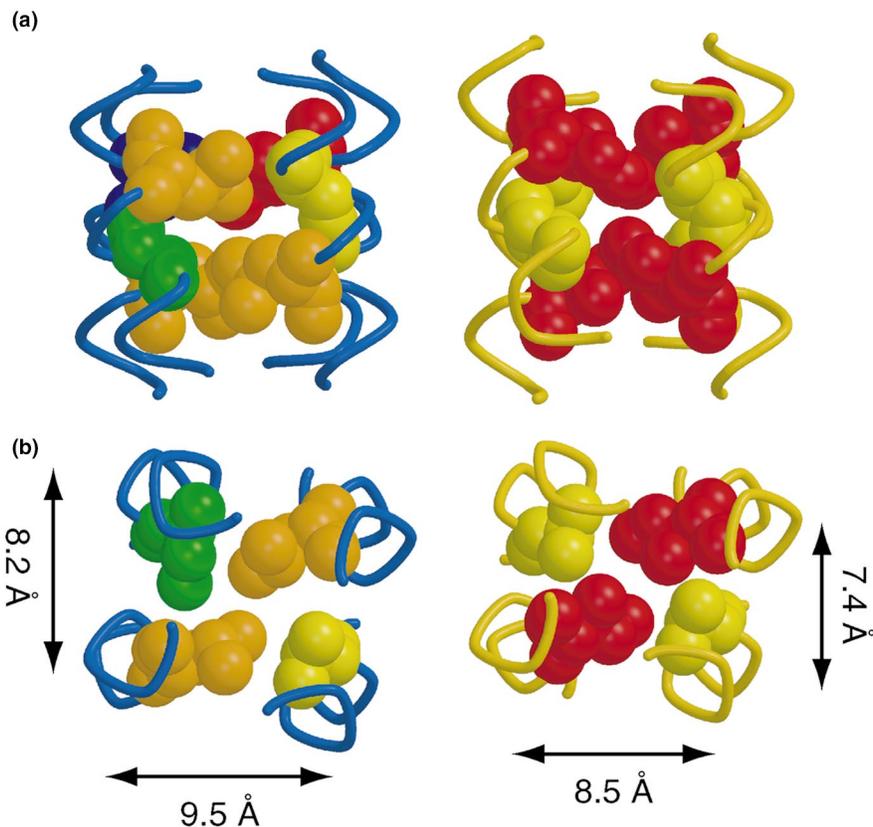
Figure 4. A Comparison of the Hydrophobic Core Layers in Rop and Ala$_2$Ile$_2$-6

The helical backbone, truncated and displayed as coils, are colored as in Figure 3. Atoms of core "a" and "d" residues are displayed as space-filling models with radii of 1.5 Å for clarity, and are colored by residue: Ala is yellow, Leu is orange, Ile is red, Cys is green, and Thr is blue.

(a) A side view showing the interaction of the two layers. Rop's layers are fairly planar, while those of Ala$_2$Ile$_2$-6 are less regular.

(b) A view down the four-helix bundle axis. The closer packing of the helices in Ala$_2$Ile$_2$-6 is shown (distances were calculated using Promotif [57]).

of helix 2 of protomer A in the C2 crystal form is not present in the P3$_2$ structure. These C2 calcium sites are coordinated by residues Ala-54, Phe-56, Asp-58, three water molecules, and residues from two symmetry-related molecules. This creates an extra bulge at the end of the helical bundle that prevents helix 1 of protomer A from fully extending to the N terminus. Thus helix 1 breaks at Thr-7 and residues 1–4 are unobserved. The other three helices pack slightly closer together (Figure 7a). This is further accommodated by the removal from the core of the Phe-56 side chain of protomer A. The other two calcium ions in the C2 structure and the two calcium sites per dimer in the P3$_2$ structure are all involved in crystal contacts. However, these sites are more remote from the termini of the dimer and thus are not involved in inducing order or disorder at the termini.

These results indicate that it is extremely unlikely that crystal packing plays any role in the formation of Ala$_2$Ile$_2$-6's new fold, and that the crystal structure of Ala$_2$Ile$_2$-6 represents a true structural rearrangement of the protomers relative to wild-type Rop.

Modeling
Why is there such a large conformational flip in Ala$_2$Ile$_2$-6? We investigated possible reasons for such structural changes by generating models with the sequence of Ala$_2$Ile$_2$-6 imposed on the wild-type structure (and vice versa), and compared these models with the crystal structures. There are two types of structural differences between Ala$_2$Ile$_2$-6 and Rop: (1) the structure of individual protomers, and (2) the association of the protomers to form the overall four-helical bundle. "Protomer structure" refers to the alignment between the helices of a single protomer, and "fold" designates the orientation of one protomer with respect to the other. Eight models for analysis arise from all possible combinations of the two sequences, two
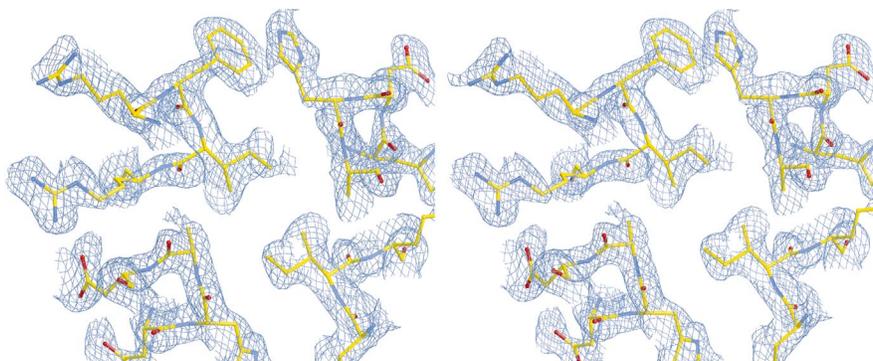
Figure 5. Ala$_2$Ile$_2$-6 Model and Electron Density

A view, in stereo, down the four-helical bundle axis of a layer of the Ala$_2$Ile$_2$-6 dimer. The C2 crystal-form dimer, with the same atomic coloring scheme used in Figure 4, is covered with the solvent-flattened experimental electron density contoured at one σ. The interaction of the Ile residues across the dimer is typical of all core Ile-Ile interactions. Also shown is a PheB-14-HisA-44 interaction that is created by the flip in Ala$_2$Ile$_2$-6. Because of this thick section, a mask was implemented to remove from this illustration any electron density beyond 1.4 Å from the model.

(a)

(b)
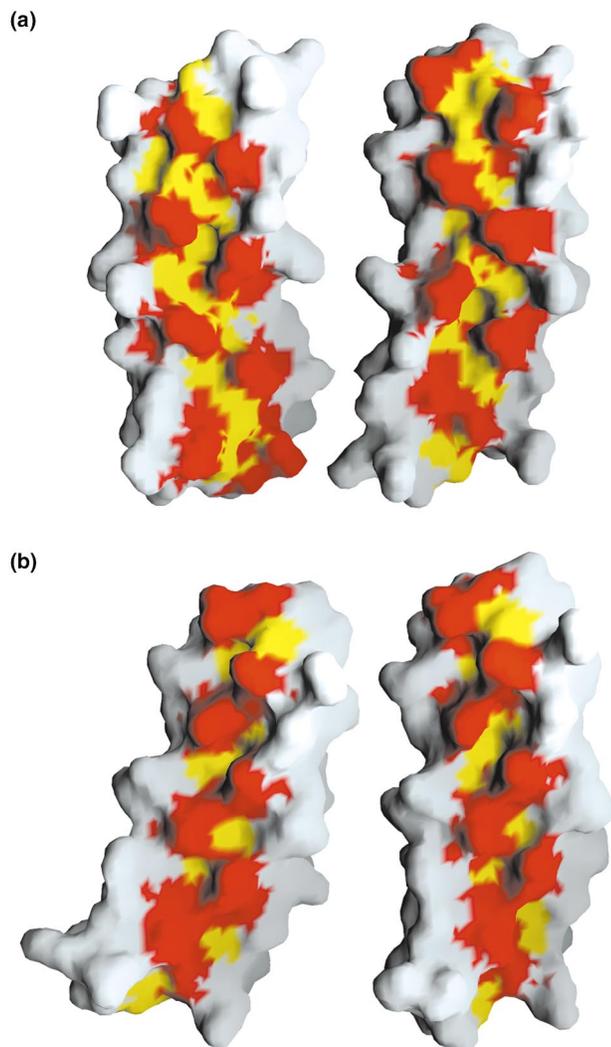
Figure 6. Dimeric Interface of Rop and Ala$_2$Ile$_2$-6

The molecular surface of the dimeric interface of Rop (a) and Ala$_2$Ile$_2$-6 (b) are shown. To display the core residues of the dimeric interface, the two protomers have been separated and displayed in the same manner one would open and display the pages of a book. The surface of "d" and "a" residues are colored red and yellow, respectively. In Ala$_2$Ile$_2$-6 the knobs of the "d" residues on one protomer pack nicely over the ridges connecting the knobs on the other protomer and into the holes created by the smaller "a" residues.

protomer structures, and two folds (Rop and Ala$_2$Ile$_2$-6 for each).

Each model is identified using the following naming format: XYZ, where X represents the amino acid sequence, Y the protomer structure, and Z the fold used to generate that model. For example, WRW indicates a model generated using the wild-type sequence (W) on the Ala$_2$Ile$_2$-6 (repacked or R) protomer structure and the wild-type (W) fold. The names of the eight models generated are WWW, WWR, WRW, WRR, RWW, RWR, RRW, and RRR. WWW and RRR use the wild-type and Ala$_2$Ile$_2$-6 crystal structures, respectively, as their starting models for simulated annealing and minimization. The other models were generated and selected as described in the Experimental Procedures section. We evaluated each model by examining cavity formation and core packing, as well as rotamer preferences. The modeling results are summarized in Figure 8.

We must point out that these models, some of which look quite good, represent a worst-case scenario; for consistency, restraints were placed on the annealing and minimization procedure, which bias the models. To maintain the integrity of the models when imposing a different sequence onto the starting structures, we employed harmonic point restraints on the backbone atoms. However, as mentioned earlier, Ala$_2$Ile$_2$-6 attains tighter overall packing in the hydrophobic core by moving the helices closer together. Thus we would expect cavity formation and reduced core packing when forcing the Ala$_2$Ile$_2$-6 sequence onto the Rop fold. Conversely, forcing the Rop sequence onto the Ala$_2$Ile$_2$-6 fold would presumably result in steric clash and poor values for side chain rotamers. The results in Figure 8 do, in fact, show the poorest packing for the Ala$_2$Ile$_2$-6 sequence on the Rop fold. However, this coincides with poor choices for the core side chain rotamer values. In spite of creating more room for the side chains to adopt favorable rotamer values, this combination of sequence and fold appears to be unfavorable for steric reasons. The fact that the Ala$_2$Ile$_2$-6 sequence results in unfavorable rotamer values for all models except the Ala$_2$Ile$_2$-6 crystal structure is most likely a consequence of the limited rotamer options for isoleucine in an α-helix.

The results of forcing the Rop sequence onto the more compact Ala$_2$Ile$_2$-6 fold are also unexpected. The cavities in the core are smaller, favoring the Ala$_2$Ile$_2$-6 fold, but the side chain rotamer values of these models (WWR and WRR) are equally favorable to those observed in the wild-type crystal structure and the WWW model. Why, then, does the Rop sequence prefer the Rop fold to that of Ala$_2$Ile$_2$-6?

Polar side chains can play an important role in determining specificity in protein-protein interactions and oligomerization [29]. The only polar side chains removed in the generation of Ala$_2$Ile$_2$-6 are Thr-19 and Gln-34. Thr-19 adopts a reasonable conformation in both the WRR and WWR models so we would not expect this residue to play a significant role in determining the Rop fold. In contrast, each Gln-34 side chain in the wild-type structure adopts a slightly disfavored rotamer value in placing the polar terminus out of the hydrophobic core and into a hydrogen bonding network involving two water molecules and the side chains of Ser-50 and Arg-155. When the Rop sequence is forced onto the Ala$_2$Ile$_2$-6 fold, however, these favorable hydrogen-bonding networks are replaced by unfavorable nonpolar interactions with Leu-29 and the nonpolar portion of Lys-25. Thus it is possible that these polar interactions may play some role in the selection of the Rop fold over that of an Ala$_2$Ile$_2$-6-type fold, but are likely not the dominant cause.

Thermodynamic Analysis

To further analyze the consequences of repacking the core of Rop, we have compared the thermodynamic properties of Ala$_2$Ile$_2$-6 and Rop. Both proteins exist in solution as folded four-helix bundles and demonstrate cooperative reversible thermal denaturation transitions (Figure 2a, inset). However, as a consequence of large changes in both the enthalpy (ΔH) and entropy (ΔS) of folding relative to wild-type Rop (Figure 2b, Table 1), there is a shift of the free energy profile of Ala$_2$Ile$_2$-6 to higher temperatures (Figure 2a, Table 1) and the appearance of cold denaturation in the presence of guanidine hydrochloride (GuHCl) (data not shown).

The change in the heat capacity (ΔC_p) upon protein folding for Rop and Ala$_2$Ile$_2$-6 are the same within experimental error. This is in agreement with the view that the ΔC_p of folding reflects the polar and especially nonpolar surface area [30]

(a)

(b)
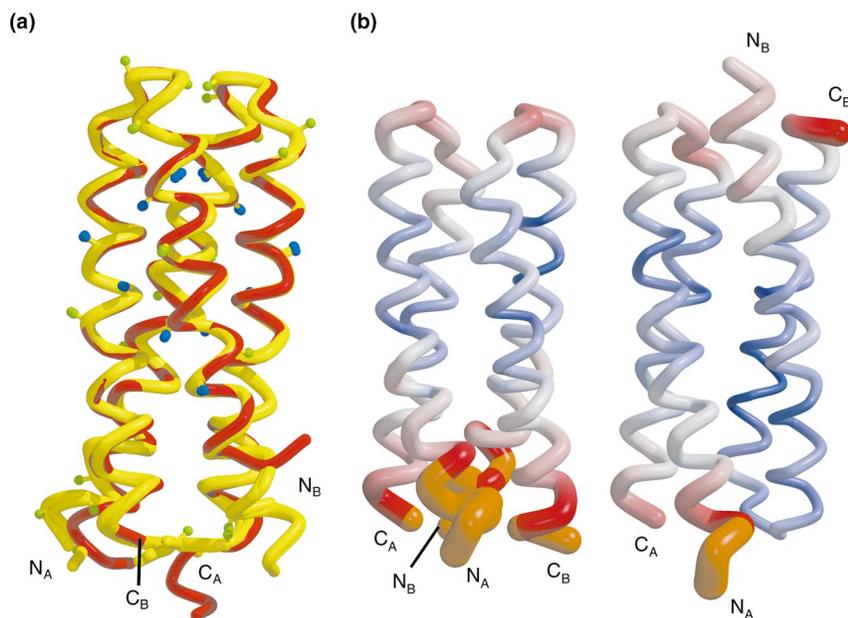
Figure 7. Crystal Contacts and B-Factors

Helix termini are labeled as in Figure 3.
(a) A comparison of the two crystal forms of Ala$_2$Ile$_2$-6 along with nonsolvent crystal contacts. The dimers from the P3$_2$ crystal form are all shown in yellow and are fit to the backbone of the C2 crystal structure, in red. Crystal contacts for the P3$_2$ dimers are indicated by C$_\beta$ atoms displayed as colored balls. Blue indicates crystal contacts seen on all six protomers. Crystal contacts not seen on all six protomers are colored green.
(b) B-factor representation of Ala$_2$Ile$_2$-6 (left) and the D30G mutant [36] of Rop (right). The D30G mutant shown is the one with the highest B factors (out of two D30G dimers and the wild-type protein). The coils are colored from blue (B factor of 12 Å^2) to red (B factor of 60 Å^2). Residues with B factors greater than 60 Å^2 are colored orange. The coil radius is also determined by the B factor: the radius of the tube, from 0.2 to 2.0 Å, represents the B factor range from 12 to 90 Å^2.

that is buried: The calculated difference in buried surface area between Rop and Ala$_2$Ile$_2$-6 is very small (Table 1). This similarity of buried surface area also explains why the structure-based calculations [31] of ΔC_p and ΔH for both Rop and Ala$_2$Ile$_2$-6 are essentially identical (Table 1). However, the experimental values for ΔH and ΔS of Ala$_2$Ile$_2$-6 are very different from those of Rop. At room temperature (293 K) the difference in ΔH between these two proteins ($\Delta\Delta H$) is 48 kcal mol^{-1}. The value for T($\Delta\Delta S$) is similarly high: 45 kcal mol^{-1}. These large differences are surprising given the small changes to the overall helical composition and size of Ala$_2$Ile$_2$-6.

To better understand the molecular basis for these large changes in the enthalpy and entropy of folding, it is helpful to further breakdown these thermodynamic parameters [2]:

$$\Delta S = \Delta S^{conf} + (\Delta S_{pol} + \Delta S_{npl})^{hyd}$$
$$\Delta H = (\Delta H^{HB} + \Delta H^{hyd})_{pol} + (\Delta H^{vdW} + \Delta H^{hyd})_{npl}$$

The entropy of hydration has both polar (ΔS_{pol}) and nonpolar (ΔS_{npl}) terms while the configurational entropy is designated ΔS^{conf}. The enthalpy of folding has contributions from both polar and nonpolar atoms due to the desolvation effects (ΔH^{hyd}), as well as hydrogen bonding (ΔH^{HB}) and van der Waals (ΔH^{vdW}) terms. If we ignore the differences between the four dimers of Ala$_2$Ile$_2$-6 in the two crystal forms by focusing on the most complete dimer as the representative structure, it is clear that we cannot account for the large differences in ΔH and T(ΔS) between Rop and Ala$_2$Ile$_2$-6. Based on the nearly identical values for both the polar and nonpolar buried surface area of Rop and Ala$_2$Ile$_2$-6, the hydration terms in the above equations contribute little toward the $\Delta\Delta H$ and T($\Delta\Delta S$). Rop has only three more hydrogen bonds [32] than the Ala$_2$Ile$_2$-6 dimer, and the core residues of Rop are less densely packed [33]. Thus hydrogen bonding and van der Waals effects probably cannot account for the 48 kcal mol^{-1} difference in ΔH. Similarly, changes in side chain configurational entropy can only account for approximately 10 of the 45 kcal mol^{-1} difference in T(ΔS) observed

Structure		Sequence		Parameter
Protomer	Fold	Rop (**W**)	Ala$_2$Ile$_2$-6 (**R**)	
W	**W**	3 (89 Å^3)	3 (252 Å^3)	Cavities
		Good	Low	Packing
		Good	3 Bad, 2 OK	χ_1 Rotamer
		2 OK	1 Bad	χ_2 Rotamer
W	**R**	1 (18 Å^3)	2 (127 Å^3)	Cavities
		Good	Low	Packing
		2 OK	2 OK	χ_1 Rotamer
		2 OK	2 Bad	χ_2 Rotamer
R	**W**	2 (35 Å^3)	0-1 (0-18 Å^3)	Cavities
		Good	Good	Packing
		2 Bad	2 Bad	χ_1 Rotamer
		1 Bad, 2 OK	3-4 Bad	χ_2 Rotamer
R	**R**	2-3 (47-83 Å^3)	2 (41 Å^3)	Cavities
		Good	Good	Packing
		Good	Good	χ_1 Rotamer
		2 OK	Good	χ_2 Rotamer

Figure 8. Summary of Parameters Used to Evaluate the Models

The letters in bold indicate either wild-type (W) or repacked (R) sequence, protomer structure, or fold. "Good" values are in green, "OK" values are in orange, and "poor" values are in red. Cavities, calculated in GRASP [52], are designated "good" for total cavity volumes of less than 50 Å^3; "OK" for cavity volumes between 50 and 100 Å^3; and "poor" for total cavity volumes of more than 100 Å^3. The number of cavities is listed along with the total cavity volume for these cavities (in parentheses). The packing [33] is considered "good" unless the values are significantly and consistently worse than the packing values of the crystal structures. Rotamer probabilities are arbitrarily labeled "good" for probabilities greater than 10%. Probabilities of 5%–10% are considered "OK", and values less than 5% are considered "poor."

between Rop and Ala$_2$Ile$_2$-6 [34]. Because the hydration terms of $\Delta\Delta$S are also small, we expect that main chain configurational entropy plays a role.

If we consider the disorder present in some of the dimers of Ala$_2$Ile$_2$-6, the large values of $\Delta\Delta$H and T($\Delta\Delta$S) are less enigmatic. Based on theoretical estimates of 3.3–4.3 kcal mol^{-1} configurational entropy per residue [34, 35], this would suggest approximately two to three disordered residues per helix to account for the rest of the $\Delta\Delta$S. Thus, the disordered residues may account for most of the large entropic difference in folding between Rop and Ala$_2$Ile$_2$-6. The same argument helps to rationalize the large change in the enthalpy of folding. If the helix termini of Ala$_2$Ile$_2$-6 are partially disordered in solution, then the number and strength of the hydrogen bonds in this region would be reduced along with favorable van der Waals interactions. To bolster this argument we note the following: (1) the crystal contacts in these structures are primarily along the faces of the proteins and at the turns (Figure 7a). Where there is strong electron density all the way to the N terminus, it is always associated with crystal contacts in that region; (2) the B factors near the N and C termini of Ala$_2$Ile$_2$-6 show considerably more inflation relative to the rest of the protein than those of Rop [23] and the D30G point mutant of Rop solved in our laboratories [36] (Figure 7b); and (3) the MRE of Ala$_2$Ile$_2$-6 is approximately 15% less than Rop. This agrees well with the estimated decrease in configurational entropy noted above.

The structure of Ala$_2$Ile$_2$-6 may explain the observed disorder in this region of the protein. The reorientation of the two protomers in Ala$_2$Ile$_2$-6 places two Phe-56, two Glu-5 and two Arg-55 residues at the same end of the core of Ala$_2$Ile$_2$-6. This combination of large and charged residues in a compact, hydrophobic core may be unfavorable. We anticipate that the eight-layer repacked version of this mutant (Ala$_2$Ile$_2$-8) will have less disorder in this region because of the introduction of smaller, more hydrophobic amino acids at these positions.

Implications for Protein Engineering

We have described how it is possible, by appropriate selection of hydrophobic core residues, to induce a dramatic conformational "flip" in a four-helix bundle protein. Although the antiparallel packing of the helices is conserved, the relative orientation of the two helix-loop-helix protomers is rotated by 180°. The thermodynamic properties of Ala$_2$Ile$_2$-6 also differ from those of wild-type Rop. The temperature of maximum stability and melting temperature of Ala$_2$Ile$_2$-6 are significantly higher than those of Rop, and the balance of entropic and enthalpic contributions to the free energy of folding is markedly different.

These results illustrate the great structural, functional, and stability differences that can be achieved with the apparently simple four-helix bundle and illustrate its utility as a framework for protein design and regulation.

Biological Implications

The hydrophobic cores of proteins play a crucial role in the folding of most proteins. Consequently, they have been the focus of considerable attention to aid protein design and the understanding of protein folding. Here, we describe the dramatic changes in both the structure and thermodynamic properties of a protein with a systematically redesigned hydrophobic core. Structurally, this protein (Ala$_2$Ile$_2$-6) retains the dimeric nature of wild-type Rop, but adopts a new fold that destroys the RNA binding function of Rop. The structural

changes are accompanied by an increase in the thermostability of Ala$_2$Ile$_2$-6 relative to wild-type protein. These results demonstrate how protein structure, function, and stability can be altered for practical requirements in pharmaceutical and biotechnological applications.

Experimental Procedures

Cloning and Purification

Ala$_2$Ile$_2$-6 was cloned (in pMR101) [37], expressed, and purified in a similar manner to previous repacked versions of Rop [9] with the following changes: The Sequenase enzyme (USB) was used for the initial primer extension, and it was essential to include a number of additional protein purification steps to remove nucleic acid for protein crystallization. After two passes over a diethylaminoethyl sephadex column [9], the protein was precipitated by adding 1 M Na acetate (pH 4.6) drop-wise until the solution became very cloudy, and the pH had dropped to 4.6. The precipitant was pelleted at 5000 rpm for 10 min in a Beckman J20 centrifuge and resuspended in 15 ml of 25 mM Tris (pH 8.0), 100 mM NaCl. The protein was then dialyzed against 2.0 liter of 25 mM Tris (pH 8.0), 100 mM NaCl before being loaded onto a POROS HQ102 or HQ20 column. The protein was eluted over a ten column-volume NaCl gradient from 0.2 to 1.5 M NaCl at 25 mM Tris (pH 8.0). Pooled fractions of Ala$_2$Ile$_2$-6 were then concentrated using an Amicon Concentrator with a 3000 molecular weight cutoff filter and the protein finally dialyzed versus 10 mM Tris (pH 8.0), 25 mM NaCl. Protein was either used immediately or lyophilized for later use. Rehydrated protein crystallized in the same manner as freshly purified protein. Protein identity was verified by DNA sequencing of the gene, amino acid analysis, and MALDI mass spectroscopy. Mass spectroscopy was also used to verify that both crystal forms contained full-length Ala$_2$Ile$_2$-6.

Thermodynamic Measurements

Protein thermal denaturation was followed by CD using an AVIV 62DS spectrometer (Aviv Instruments) and quartz cuvettes with a 2-mm pathlength (Wilmad Glass). The ellipticity at 222 nm was monitored over a temperature range of 10°C to 88°C for Rop and 14°C to 102°C for Ala$_2$Ile$_2$-6. In each case, the temperature was raised in 2°C increments and allowed to equilibrate 1 min upon reaching the target temperature before taking a measurement averaged over a 30 s period.

For both proteins, samples were prepared in triplicate at eight different concentrations of GuHCl (0.0 M, 0.3 M, 0.6 M, 1.0 M, 1.3 M, 1.6 M, 2.0 M, and 2.3 M for Rop; 0.0 M, 0.25 M, 0.5 M, 0.75 M, 1.0 M, 1.25 M, 1.5 M, and 1.75 M for Ala$_2$Ile$_2$-6). Buffered (100 mM Na phosphate [pH 7], 200 mM NaCl) protein solutions of Rop and Ala$_2$Ile$_2$-6 (19 μM, dimer) were prepared and allowed to equilibrate at room temperature for 48 and 24 hr, respectively, prior to study. The buffer solution for Rop contained 1 mM dithiothreitol to generate a reducing environment.

Both wild-type protein and Ala$_2$Ile$_2$-6 are homodimeric proteins. The equilibrium between folded dimer, F$_2$, and unfolded monomer, U, can be expressed: F$_2 \leftrightarrow$ 2U.

In dimeric proteins, the disassociation equilibrium constant, K, can be expressed:

$$K = \frac{4Cd\alpha^2}{1 - \alpha}$$

where, C$_d$ is the total molar concentration of dimeric protein and α corresponds to the unfolded fraction of total protein. For each sample, the thermal denaturation traces were baseline corrected, and their respective T$_{m,app}$ values determined as the midpoint in the corrected transition curve. Van't Hoff analysis of the data near the T$_{m,app}$ was used to calculate ΔH$_m$ for each sample, using the program ThermoDynaCD [27].

The ΔH$_m$ and T$_{m,app}$ values were used to determine ΔC$_p$ for both Rop and Ala$_2$Ile$_2$-6 by the method of Scholtz [38] using the program Kaleidagraph (Abelbeck Software). The ΔC$_p$ in this method is determined by the slope of ΔH$_m$ versus T$_m$ obtained from thermal melts at different GuHCl concentrations. The assumption is made that ΔC$_p$ is temperature invariant and that both ΔC$_p$ and ΔH are invariant with GuHCl concentration at a given temperature. The Gibbs-Helmholtz equation, in the form given below and the determined ΔC$_p$, ΔH$_m$, and T$_{m,app}$ at 0M GuHCl, were used to determine the free energy of unfolding for each protein as a function of temperature, ΔG(T).

$$\Delta G(T) = \Delta H_m \left[1 - \frac{T}{T_{m,app}} \right] + \Delta C_p \left[T - T_{m,app} + T \left(\ln \frac{T_{m,app}}{T} \right) \right] - RT_{m,app} \left[\ln(2Cd) \right]$$

The entropy at $T_{m,app}$ was then calculated using the free energy at this $T_{m,app}$:

$$\Delta S_m = \frac{-[\Delta G_{Tm,app} - \Delta H_m]}{T_{m,app}}$$

Crystallization

Three different crystal forms of Ala$_2$Ile$_2$-6 were obtained by hanging-drop vapor diffusion experiments. The crystal conditions were determined from fine-grid screens that optimized initial results from screens using the Crystal Screen I and II kits (Hampton Research) and incomplete factorial methods [39]. The C2 and P3$_2$ crystals (described in this article) grew in the same wells. Rod-like crystals belonging to the space group P3$_2$ appeared within days, whereas the C2 plates appeared to grow off the P3$_2$ crystals after 4 weeks or more. Crystals used for data collection were grown in drops containing equal volumes of 18 mg/ml protein and well solution (18%–22% MPD, 50–100 mM CaCl$_2$, 100 mM sodium HEPES[pH 7.5]).

Crystals were flash frozen in liquid propane at near-liquid nitrogen temperatures. These crystal "popsicles" were then stored in liquid nitrogen. For heavy-atom derivatives, the crystals were soaked in a mother liquor containing 20% MPD, 50 mM CaCl$_2$, 50 mM sodium HEPES (pH 7.5) and 6.0 mM K$_2$PtCl$_4$ for 3 days before the crystals were frozen.

The native data for both the C2 and P3$_2$ crystal forms were collected on R-Axis IV area detectors with Cu K$_\alpha$ radiation and Yale mirrors (Molecular Structure Corporation). The platinum data sets were all collected in 10° wedges using inverse beam geometry on a single C2 crystal. In addition to the Cu K$_\alpha$ platinum data set, two wavelengths were collected at the platinum LIII edge on an R-Axis IV area detector at the NSLS X4A beamline. Data were processed and scaled using DENZO and SCALEPACK in the HKL suite [40].

Structure Determination, Model Building, and Refinement

All calculations were performed in CNS [41]. The two platinum sites were located via heavy-atom translation searches [42] against the anomalous and isomorphous differences in the C2 platinum data sets. A maximum-likelihood target function [43] was used to refine the heavy atom parameters including the atomic f′ and f″ values that were constrained to be equal for both sites. Density modification using solvent flipping [44] and histogram matching [45] were essential for calculating traceable electron density maps from the MAD phases. The initial C2 crystal form model of Ala$_2$Ile$_2$-6 was manually built in O [46] using σ_A–weighted 2F$_o$ – F$_c$ maps [47]. Model refinement proceeded by torsion angle dynamics simulated annealing [48] with an MLHL target function [49] followed by model rebuilding in O. Unmodified SIRAS phases were used for the final stages of coordinate refinement. This consisted of alternating rounds of torsion angle molecular dynamics simulated annealing [48], restrained atomic B factor refinement [50], and model rebuilding in O. Model rebuilding included adding appropriate water molecules, calcium ions, and an MPD molecule. Residues A1-A56 and B7-B54 of the C2 crystal structure were then used as a molecular replacement model for solving the Ala$_2$Ile$_2$-6 structure in the P3$_2$ crystal form. The refinement of the three dimers in the P3$_2$ unit cell proceeded as above with the following exceptions: Model phases from the C2 structure were used with the MLF target function and 3-fold noncrystallographic symmetry constraints were employed for the initial rounds of refinement. Dropping these constraints in the final rounds of refinement resulted in significant drops in the R factors (more than 1% in R$_{free}$, and 3% in R). Full crystallographic data and phasing statistics are included as supplementary material.

Modeling Methods

For modeling purposes, we considered the following three factors: the amino acid sequence (either wild-type or Ala$_2$Ile$_2$-6), the protomer structure (with helices 1 and 2 of a given protomer either aligned or offset by half a turn), and the overall fold (the wild-type or flipped orientation of the protomers within a dimer). Thus we created $2 \times 2 \times 2 = 8$ initial models based on the two possible states for each of these three factors. We used the most complete dimer in the P3$_2$ crystal structure as a starting point for the Ala$_2$Ile$_2$-6 protomer structure and fold. The naming convention for the models is described in the Modeling section of the Discussion. To generate RWW, the Rop crystal structure was mutated to the Ala$_2$Ile$_2$-6 sequence using the program O [46]. Likewise, WRR was generated by simply mutating the Ala$_2$Ile$_2$-6 crystal structure to the wild-type sequence. RRW was generated by performing an independent least-squares backbone fit of each Ala$_2$Ile$_2$-6

protomer to the appropriate protomer of Rop. In the same manner WWR was generated by performing an independent least-squares backbone fit of each Rop protomer to the appropriate protomer of Ala$_2$Ile$_2$-6. Finally, to produce WRW and RWR, the sequences of RRW and WWR generated above were changed back to the wild-type and Ala$_2$Ile$_2$-6 sequences, respectively.

Each of these eight starting models then underwent ten rounds of simulated annealing followed by energy minimization to create a family of ten structures for each model. Harmonic restraints on backbone atoms were used during both simulated annealing and energy minimization. The simulated annealing used a slow-cooling Cartesian molecular dynamics protocol with a starting temperature of 2500 K, using velocity scaling temperature control, and a drop of 25 K per cycle. The time course for the simulation proceeded over 10,000 steps at 0.0005 picoseconds per step. The energy minimization consisted of 200 steps of conjugate gradient minimization. Both the simulated annealing and energy minimization utilized a nonbonded cutoff of 13 Å and a dielectric constant of one. Core positions ("a" and "d") for each of these populations were analyzed to determine the most populated rotamers. If one of the ten structures in a given family did not have every core residue in the most populated rotamer conformation, the side chain was manually rotated to the preferred rotamer position and the model reminimized as above. WRR and RRW have two and three model structures, respectively, because for these models there was one core position each with two or three equally populated rotamer positions.

The programs GRASP [51, 52] and OS [33] were used for cavity calculations and the packing analysis, respectively. In addition to calculating the packing values for the structure in each population deemed most representative, the packing calculations were done on each of the ten minimized structures in a population in order to derive a mean packing value and standard deviation for the core residues of that population. The rotamer analysis was performed using an updated (August 1999) backbone-dependent library [53].

Generation of Figures

Figures 3, 4, and 7a were generated using MOLSCRIPT [54]. Figures 5 and 7b were made using BOBSCRIPT [55]. Raster3D [56] was used to render Figures 3, 4, and 7. GRASP [52] was used to create Figure 6.

References

1. Richards, F.M. (1977). Areas, volumes, packing and protein structure. Annu. Rev. Biophys. Bioeng. *6*, 151–176.
2. Makhatadze, G.I., and Privalov, P.L. (1995). Energetics of protein structure. Adv. Protein Chem. *47*, 307–425.
3. Sandberg, W., and Terwilliger, T. (1989). Influence of interior packing and hydrophobicity on the stability of a protein. Science *245*, 54–57.
4. Lim, W., and Sauer, R. (1991). The role of internal packing interactions in determining the structure and stability of a protein. J. Mol. Biol. *219*, 359–376.
5. Eriksson, A.E., et al. (1992). Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. Science *255*, 178–183.
6. Hellinga, H.W., Wynn, R., and Richards, F.M. (1992). The hydrophobic core of *Escherichia coli* thioredoxin shows a high tolerance to nonconservative single amino acid substitutions. Biochemistry *31*, 11203–11209.
7. Jackson, S., Moracci, M., elMasry, N., Johnson, C., and Fersht, A. (1993). Effect of cavity-creating mutations in the hydrophobic core of chymotrypsin inhibitor 2. Biochemistry *32*, 11259–11269.
8. Lim, W.A., Hodel, A., Sauer, R.T., and Richards, F.M. (1994). The crystal

structure of a mutant protein with altered but improved hydrophobic core packing. Proc. Natl. Acad. Sci. USA *91*, 423–427.

9. Munson, M., O'Brien, R., Sturtevant, J.M., and Regan, L. (1994). Redesigning the hydrophobic core of a four-helix-bundle protein. Protein Sci. *3*, 2015–2022.

10. Munson, M., et al., and Regan, L. (1996). What makes a protein a protein? Hydrophobic core designs that specify stability and structural properties. Protein Sci. *5*, 1584–1593.

11. Xu, J., Baase, W.A., Baldwin, E., and Matthews, B.W. (1998). The response of T4 lysozymes to large to small substitutions within the core and its relation to the hydrophobic effect. Protein Sci. *7*, 158–177.

12. Harbury, P.B., Zhang, T., Kim, P.S., and Alber, T. (1993). A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. Science *262*, 1401–1407.

13. Wagschal, K., Tripet, B., and Hodges, R.S. (1999). De novo design of a model peptide sequence to examine the effects of single amino acid substitutions in the hydrophobic core on both stability and oligomerization state of coiled-coils. J. Mol. Biol. *285*, 785–803.

14. Regan, L., and DeGrado, W.F. (1988). Characterization of a helical protein designed from first principles. Science *241*, 976–978.

15. Kamtekar, S., Schiffer, J.M., Xiong, H., Babik, J.M., and Hecht, M.H. (1993). Protein design by binary patterning of polar and nonpolar amino acids. Science *262*, 1680–1685.

16. Davidson, A.R., Lumb, K.J., and Sauer, R.T. (1995). Cooperatively folded proteins in random sequence libraries. Nat. Struct. Biol. *2*, 856–864.

17. Finucane, M.D., Tuna, M., Lees, J.H., and Woolfson, D.N. (1999). Core-directed protein design. I. An experimental method for selecting stable proteins from combinatorial libraries. Biochemistry *38*, 11604–11612.

18. Ponder, J.W., and Richards, F.M. (1987). Tertiary templates for proteins: use of packing criteria in the enumeration of allowed sequences for different structural classes. J. Mol. Biol. *193*, 775–791.

19. Desjarlais, J.R., and Handel, T.M. (1995). De novo design of the hydrophobic cores of proteins. Protein Sci. *4*, 2006–2018.

20. Dahiyat, B.I., and Mayo, S.L. (1996). Protein design automation. Protein Sci. *5*, 895–903.

21. Lee, C., and Subbiah, S. (1991). Prediction of protein side-chain conformation by packing optimization. J. Mol. Biol. *217*, 373–388.

22. Harbury, P.B., Tidor, B., and Kim, P.S. (1995). Repacking protein cores with backbone freedom: structure prediction for coiled coils. Proc. Natl. Acad. Sci. USA *92*, 8408–8412.

23. Banner, D., Kokkinidis, M., and Tsernoglou, D. (1987). Structure of the ColE1 Rop protein at 1.7 Å resolution. J. Mol. Biol. *196*, 657–675.

24. Eberle, W., Pastore, A., Sander, C., and Rosch, P. (1991). The structure of ColE1 rop in solution. J. Biomol. NMR *1*, 71–82.

25. Gregorian, R., Jr., and Crothers, D. (1995). Determinants of RNA hairpin loop-loop complex stability. J. Mol. Biol. *248*, 968–984.

26. Glykos, N.M., Cesareni, G., and Kokkinidis, M. (1999). Protein plasticity to the extreme: changing the topology of a 4-alpha-helical bundle with a single amino acid substitution. Structure *7*, 597–603.

27. Predki, P.F., Nayak, L.M., Gottlieb, M.B., and Regan, L. (1995). Dissecting RNA-protein interactions: RNA-RNA recognition by Rop. Cell *80*, 41–50.

28. Gernert, K.M., Surles, M.C., Labean, T.H., Richardson, J.S., and Richardson, D.C. (1995). The alacoil: a very tight, antiparallel coiled-coil of helices. Protein Sci. *4*, 2252–2260.

29. Gonzalez, L., Jr., Woolfson, D.N., and Alber, T. (1996). Buried polar residues and structural specificity in the GCN4 leucine zipper. Nat. Struct. Biol. *3*, 1011–1018.

30. Gomez, J., Hilser, V.J., Xie, D., and Freire, E. (1995). The heat capacity of proteins. Proteins *22*, 404–412.

31. Baker, B.M., and Murphy, K.P. (1998). Prediction of binding energetics from structure using empirical parameterization. Methods Enzymol. *295*, 294–315.

32. McDonald, I.K., and Thornton, J.M. (1994). Satisfying hydrogen bonding potential in proteins. J. Mol. Biol. *238*, 777–793.

33. Pattabiraman, N., Ward, K.B., and Fleming, P.J. (1995). Occluded molecular surface: analysis of protein packing. J. Mol. Recognit. *8*, 334–344.

34. Doig, A.J., and Sternberg, M.J. (1995). Side-chain conformational entropy in protein folding. Protein Sci. *4*, 2247–2251.

35. Makhatadze, G.I., and Privalov, P.L. (1996). On the entropy of protein folding. Protein Sci. *5*, 507–510.

36. Predki, P.F., Agrawal, V., Brunger, A.T., and Regan, L. (1996). Amino-acid substitutions in a surface turn modulate protein stability. Nat. Struct. Biol. *3*, 54–58.

37. Munson, M., Predki, P.F., and Regan, L. (1994). ColE1-compatible vectors for high-level expression of cloned DNAs from the T7 promoter. Gene *144*, 59–62.

38. Scholtz, J.M. (1995). Conformational stability of HPr: the histidine-containing phosphocarrier protein from *Bacillus subtilis*. Protein Sci. *4*, 35–43.

39. Audic, S., Lopez, F., Claverie, J.M., Poirot, O., and Abergel, C. (1997). SAmBA: an interactive software for optimizing the design of biological macromolecules crystallization experiments. Proteins *29*, 252–257.

40. Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

41. Brunger, A.T., et al., and Warren, G.L. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D *54*, 905–921.

42. Grosse-Kunstleve, R.W., and Brunger, A.T. (1999). A highly automated heavy-atom search procedure for macromolecular structures. Acta Crystallogr. D *55*, 1568–1577.

43. Burling, F.T., Weis, W.I., Flaherty, K.M., and Brunger, A.T. (1996). Direct observation of protein solvation and discrete disorder with experimental crystallographic phases. Science *271*, 72–77.

44. Abrahams, J.P., and Leslie, A.G.W. (1996). Methods used in the structure determination of bovine mitochondrial F-1 ATPase. Acta Crystallogr. D *52*, 30–42.

45. Zhang, K.Y.J., and Main, P. (1990). The use of sayre equation with solvent flattening and histogram matching for phase extension and refinement of protein structures. Acta Crystallogr. A *46*, 377–381.

46. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard. (1991). Improved methods for binding protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

47. Read, R.J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. Acta Crystallogr. A *42*, 140–149.

48. Rice, L.M., and Brunger, A.T. (1994). Torsion angle dynamics: reduced variable conformational sampling enhances crystallographic structure refinement. Proteins *19*, 277–290.

49. Pannu, N.S., Murshudov, G.N., Dodson, E.J., and Read, R.J. (1998). Incorporation of prior phase information strengthens maximum-likelihood structure refinement. Acta Crystallogr. D *54*, 1285–1294.

50. Hendrickson, W.A. (1985). Stereochemically restrained refinement of macromolecular structures. Methods Enzymol. *115*, 252–270.

51. Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281–296.

52. Nicholls, A., Bharadwaj, R., and Honig, B. (1993). Grasp—graphical representation and analysis of surface-properties. Biophys. J. *64*, A166–A166.

53. Dunbrack, R.L., and Karplus, M. (1993). Backbone-dependent rotamer library for proteins—application to side-chain prediction. J. Mol. Biol. *230*, 543–574.

54. Kraulis, P.J. (1991). Molscript—a program to produce both detailed and schematic plots of protein structures. J. Appl. Cryst. *24*, 946–950.

55. Esnouf, R.M. (1997). An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. J. Mol. Graph. Model. *15*, 132–134, 112–113.

56. Merritt, E.A., and Bacon, D.J. (1997). Raster3D: photorealistic molecular graphics. Methods Enzymol. *277*, 505–524.

57. Hutchinson, E.G., and Thornton, J.M. (1996). PROMOTIF—A program to identify and analyze structural motifs in proteins. Protein Sci. *5*, 212–220.

58. Lee, B., and Richards, F. (1971). The interpretation of protein structures: estimation of static accessibility. J. Mol. Biol. *55*, 397–400.

Accession Codes

The coordinates for the Ala$_2$Ile$_2$-6 structure in the C2 and P32 crystal forms have been submitted to the Protein Data Bank (accession codes 1F4M and 1F4N, respectively).