

## Supporting Information

### Evaluation of a symmetry-based strategy for assembling protein complexes

Dustin P. Patterson<sup>a</sup>, Ankur M. Desai<sup>a,b</sup>, Mark M. Banaszak Holl<sup>a,b</sup> and E. Neil G.  
Marsh<sup>\*a,b,c</sup>

<sup>a</sup>Department of Chemistry, University of Michigan, Ann Arbor, MI 48109, USA

e-mail [nmarsh@umich.edu](mailto:nmarsh@umich.edu)

<sup>b</sup>Michigan Nanotechnology Institute for Medicine and Biological Sciences, University of  
Michigan, Ann Arbor, MI 48109, USA

<sup>c</sup>Department of Biological Chemistry, University of Michigan, Ann Arbor, MI 48109, USA

## Characterization of de-novo designed coiled-coil forming peptides Helix-(+) and Helix-(-)

The heterodimeric coiled-coil design and the protein sequences for Helix-(-) and Helix-(+) are shown in Figure S1. In addition to the coiled-coil forming region an N-terminal 6 residue histidine (6xHis) tag was incorporated to allow purification by Ni-affinity chromatography and a thrombin cleavage site was incorporated between the 6xHis tag and coiled coil peptide to allow removal of the tag if required. Additional amino acids at the C-termini were included as flexible spacer sequences in the event that the peptides would be appended to the N-terminus a protein.

### Construction of synthetic genes for Helix-(-) and Helix-(+)

Synthetic genes were designed and commercially synthesized (Picoscript, Houston, Texas) to express Helix-(-) and Helix-(+) with the codon usage optimized for expression in *E. coli*. The genes encoding Helix-(-) and Helix-(+) were subsequently subcloned into the expression vector pET28b(+). The DNA sequence encoding Helix-(+) as found in pET28b(+), with start and stop sequences underlined and unique restriction sites for genetic manipulation, is given below.

```

1   ATACCATGGGCTCCTCTCATCATCATCACCACCACAGCTCGGGCCTGGTGCCGCGCGGCACCCACGGCTCGAGCAAACAG   80
      NcoI                                     XhoI

81  TTGGAGAAAGAACTGGCGCAACTCAAAAAGAACTGCAGGCCATTAAAAAAAAGTGGCGCAATTGAAATGGAAGGCCCA   160
                                   PstI                               MfeI

161 GCGCTGAAAAAAAAGTGGCACAGCTGAAGAAAAAACTCCAAGGTGGCTCCGGAGCTAGCTAATAA   227
                        PvuII                BspEI      BmtI
                                           NheI
  
```

The DNA sequence encoding Helix-(-) in pET28b(+) as found in pET28b(+), with start and stop sequences underlined and unique restriction sites for genetic manipulation, is given below.

```

1  ATACCATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGGATCCAAACAG      80
    NcoI                                     NdeI
                                           BamHI
81  TTGGAAAAGAACTGGCCCAACTGGAAAAGGAAGTCAAGCGATTGAGAAAGAGCTCGCTCAACTGGAATGGGAGGCACA      160
                                           EcoICRI
                                           SacI
161 AGCTTTGGAAAAAACTGGCGCAGCTCGAAAAAGAGTTGCAGGGCGGCAGCGGCGGTACCTAATAA      227
    HindIII                                     Acc65I
                                           KpnI

```

## Expression and purification of Helix(-) and Helix(+)

Over-expression of Helix(-) and Helix(+) in *E. coli* and subsequent purification by affinity chromatography on Ni-NTA resin yielded pure peptides of expected size as observed by tris-tricine SDS-PAGE (Figure S2). Mass spectrometry analysis of samples desalted by reverse phase HPLC gave a mass of 7687.8 Da for Helix(+) consistent with the 7688.06 Da expected for removal of Met at the N-terminus and Ser at the C-terminus, and 7815 Da for Helix(-) consistent with the loss of Met from terminus, calculated to be 7813.7 Da.

## Secondary structure of Helix(-) and Helix(+)

Helix(-) displayed strong minima signals at 202 and 208 nm indicating that it adopts a predominantly alpha-helical structure (Figure S3). The C.D. spectrum for Helix(+) gave a signal consistent with a random coil structure. Mixing the two peptides in a 1:1 ratio gave a spectrum that was more helical than expected for the average for the C.D. spectra of the individual peptides. No change in CD of the 1:1 mixture was observed after incubation of the mixture for several days. Analysis of the spectra indicate that although Helix(-) predominantly self associates to form a homodimeric coiled-coil, 15 – 20 % of the peptide forms heterodimeric coiled-coils with Helix(+).

### **Sedimentation equilibrium AUC studies of Helix-(-) and Helix-(+)**

The peptides were analyzed by sedimentation equilibrium AUC. The traces for Helix-(-) fit well to a monomer-dimer equilibrium model with it existing primarily as a species of 15,180 Da, close to the expected molecular weight of a dimer (calculated dimer=15,634 Da) (Figure S3). This result is consistent with the C.D. spectrum of Helix-(-) which indicated that is structured in the absence of Helix-(+).

The sedimentation traces for Helix-(+) fit best to single component model with a molecular weight of 14,430 Da (Figure S4), a molecular weight much higher than that expected for a monomer (calculated  $M_r$  = 7,688 Da). The reason for this is unclear. If the peptide was forming a coiled coil dimer we would expect this to be reflected in a highly  $\alpha$ -helical C.D. spectrum for the peptide. Moreover fitting the data to a monomer—dimer equilibrium gave poor quality fits, suggesting that dimer formation is not the cause of the high apparent molecular weight. One possibility is that the high positive charge on the peptide results in non-ideality in its sedimentation properties; i.e. interactions with solvating anions which would tend to make the peptide appear larger than predicted.

### **Determination of the Stability of Helix-(-) homo-dimer**

Guanidinium hydrochloride denaturation studies were performed to determine the overall stability of the Helix-(-) dimer. Denaturation curves are consistent with cooperative two state folding by Helix-(-) (Figure S6). The Helix-(-) dimer is very stable with denaturation midpoints above 5 M for the two different concentrations (9.2  $\mu$ M and 21  $\mu$ M). Fitting the denaturation curves to a dimeric model gave  $\Delta G_{\text{unfold}}$  of  $24.5 \pm 2.6$  kcal/mol and  $26.8 \pm 3.9$  kcal/mol at concentrations of 9.2  $\mu$ M and 21  $\mu$ M, respectively.

Highly cooperative folding is evident from the  $m$  values of  $3.3 \pm 0$ . kcal mol<sup>-1</sup> M<sub>GuHCl</sub><sup>-1</sup> for 9.2 μM and  $3.5 \pm 0.65$  kcal mol<sup>-1</sup> M<sub>GuHCl</sub><sup>-1</sup> for 21 μM determined from the fits.

## Experimental Procedures

### Materials

DNA modifying enzymes and reagents were purchased from New England Biolabs. DNA primers were purchased from IDT DNA Technologies (Coralville, IA). PBS 10x stock solution was purchased from Invitrogen (Carlsbad, CA). *E. coli* BL21(λDE3) and expression vector pET-28b were purchased from Novagen (Madison, WI); Pfu turbo DNA polymerase and *E. coli* XL1-Blue were from Stratagene (Cedar Creek, TX). Nickel nitrilotriacetic acid (Ni-NTA) resin, QIAquick gel extraction kit, and QIAprep Spin Miniprep kit were purchased from Qiagen (Valencia, CA). All chemical reagents (IPTG, Sodium Chloride, etc.) were from Fisher (Pittsburgh, PA).

### Peptide Expression and Purification

*E. coli* strains harboring expression constructs were grown on 2xTY medium at 37°C in the presence of kanamycin (50 μg/mL) to maintain selection for the plasmids. Expression of the genes was induced by addition of isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM once the cells reached early log phase (OD<sub>600</sub>=0.8). Cultures were grown for 4 hours after addition of IPTG, then the cells were harvested by centrifugation and cell pellets stored at -20° C until needed.

Cell pellets were resuspended in ice-cold lysis buffer (100 mM Tris-HCl, 300 mM sodium chloride, 0.1% Triton-X 100, pH 8.0) containing 1 mM β-mercaptoethanol, 10 μM PMSF. Cells were lysed on ice by sonication with a microtip for 6 minutes total run

time, 30 second pulse, and 2 minute pause between bursts. Cell debris was removed by centrifugation at 24,000 g for 15 min at 4°C. The supernatant was then heated at 80°C in a water bath for 20 minutes to produce a white cloudy precipitate. Precipitated material was removed by centrifugation at 10,000 g for 15 min at 4°C. The supernatant was loaded onto a 5 mL bench top column of Ni-NTA superflow resin equilibrated in Buffer A (50 mM Tris-HCl, 300 mM NaCl, 10 mM imidazole, pH 7.0). Non-specifically bound proteins were eluted by washing the column with 10 column volumes of buffer A. Finally the 6xHis-tagged peptides were eluted with Buffer A containing 500 mM imidazole. Purified protein solutions were dialyzed twice overnight against 1 L PBS (10 mM sodium phosphate, 150 mM NaCl, pH 7.0) and stored at 4°C. Protein concentrations were determined by UV absorption measured at 280 nm assuming a molar extinction coefficient  $\epsilon_{280} = 5690 \text{ M}^{-1}\text{cm}^{-1}$ , peptides each contain one tryptophan for this purpose.

#### Tris-Tricine SDS-PAGE Analysis

Tris-tricine gels were made at 16.5%T/6.0%C for the separating and 4.0%T/3.0%C for stacking gel. Gels were subjected to electrophoresis overnight at 15 volts and then at 100 volts until the loading dye reached the bottom of the gel. Gels were subsequently stained with Coomassie, then destained and photographed.

#### HPLC Analysis

Chromatography was performed at room temperature with detection at 220 nm using an Alltech reverse phase C-4 column (150 mm x 4.6 mm, 5 $\mu$ ). The column was equilibrated with a solution of 0.15% TFA in deionized water and peptides eluted with an ascending gradient of Acetonitrile (95% Acetonitrile:5% Water and 0.13% TFA) using the time gradient program, shown in Table 1, with a flow rate of 1 mL/min. Peak

fractions were collected for analysis by electrospray ionization mass spectrometry (ESI-MS) to confirm identity of peptides.

#### Mass Spectrometry

Samples collected from the HPLC analysis were directly analyzed by time-of-flight electrospray mass spectrometry (LCT Micromass) under positive ion mode. ESI-MS samples were introduced into the spectrometer through an in-line HPLC pump in a carrier solvent of 90% methanol/10% water at 0.1 mL/min. The desolvation temperature was 150°C. The spectral data were analyzed using the MassLynx 4.0 software suite. Expected masses were calculated from sequences using ExPASy compute pI/Mw tool.

#### Circular Dichroism

Circular dichroism (CD) spectra were recorded on an Aviv model 202 circular dichroism spectrometer. Rectangular quartz cuvettes of 1 mm pathlength were used for all experiments. Experiments were run at 25°C. Typically, CD spectra were the average of three scans between 190 and 250 nm. Molar mean residue ellipticity ( $\theta$ ) was calculated from the equation:

$$\theta = (\theta_{\text{obs}} \times \text{MRW}) / (10 \times l \times C)$$

where  $\theta_{\text{obs}}$  is the ellipticity measured in millidegrees, MRW is the mean residue molecular weight (molecular weight of peptide divided by the number of amino acids),  $l$  is the optical pathlength in centimeters, and  $c$  is the concentration in milligrams per milliliter. Temperature denaturation was investigated by taking scans of a sample at 25°C and 90°C, and again at 25°C after cooling. Guanidine hydrochloride studies were carried out by monitoring the CD at 222 nm. Peptide samples were prepared from stock solutions of 168  $\mu\text{M}$  peptide in buffer (10 mM sodium phosphate, 100 mM sodium chloride, 1 mM

EDTA, pH 7.5) and 8 M Guanidine in buffer (same as peptide). Fraction unfolding was determined using the methods described by De Francesco et al. and Litowski et al.<sup>1-3</sup> Denaturation data was fit to a dimer association model to determine  $\Delta G_{\text{unfold}}$  and  $m$  ( $m$  is a constant of proportionality with dimensions of  $\text{kcal mol}^{-1} M_{\text{GuHCl}}^{-1}$ ) using the program Igor Pro.

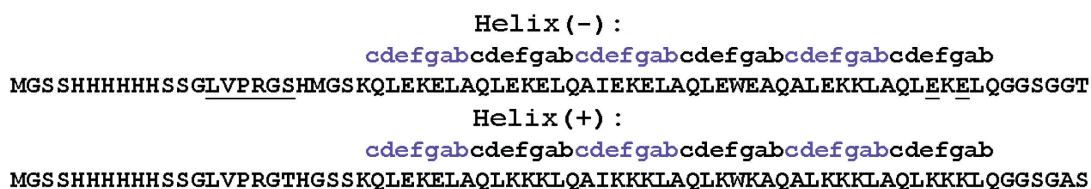
#### Sedimentation Equilibrium Analytical Ultracentrifugation

Sedimentation equilibrium analytical ultracentrifugation (SE AUC) experiments were performed on a Beckman XL-I analytical ultracentrifuge. Samples with  $A_{280}$  of 0.3, 0.5, and 0.7 were loaded into Epon 6 channel centrifuge cells and spun at four speeds (30 krpm, 35 krpm, 40krpm, and 45 krpm). Temperature was set at 20°C. Equilibrium was verified by observing overlap of scans taken several hours apart. Data for scans observed to be at equilibrium were analyzed using the software package Ultrascan using global fit analysis function. Partial specific volumes were calculated from peptide sequences using the program Sednterp.



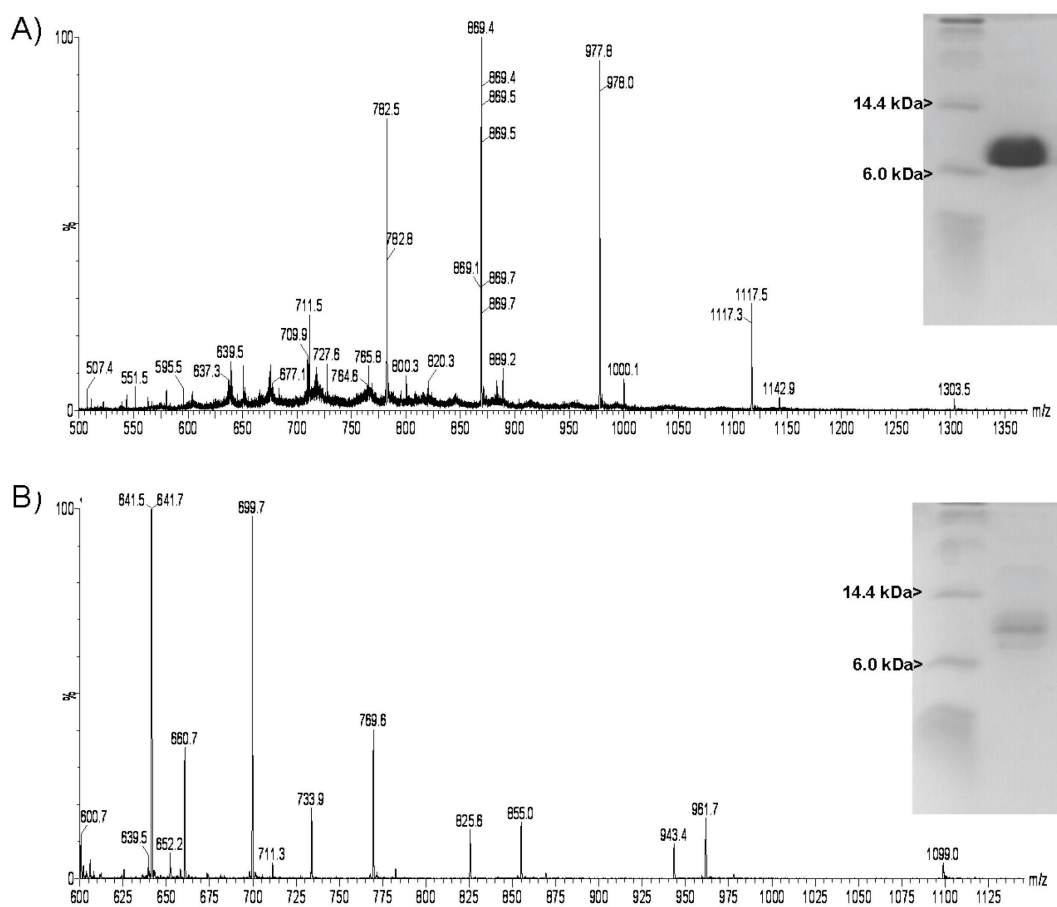
## References

- 1 De Francesco, R., Pastore, A., Vecchio, G. & Cortese, R. Circular dichroism study on the conformational stability of the dimerization domain of transcription factor LFB1. *Biochemistry* **30**, 143-147 (1991).
- 2 Litowski, J. R. & Hodges, R. S. Designing heterodimeric two-stranded alpha-helical coiled-coils: the effect of chain length on protein folding, stability and specificity. *J Pept Res* **58**, 477-492, doi:jpp10972 [pii] (2001).
- 3 Litowski, J. R. & Hodges, R. S. Designing heterodimeric two-stranded alpha-helical coiled-coils. Effects of hydrophobicity and alpha-helical propensity on protein folding, stability, and specificity. *J Biol Chem* **277**, 37272-37279, doi:10.1074/jbc.M204257200 [pii] (2002).



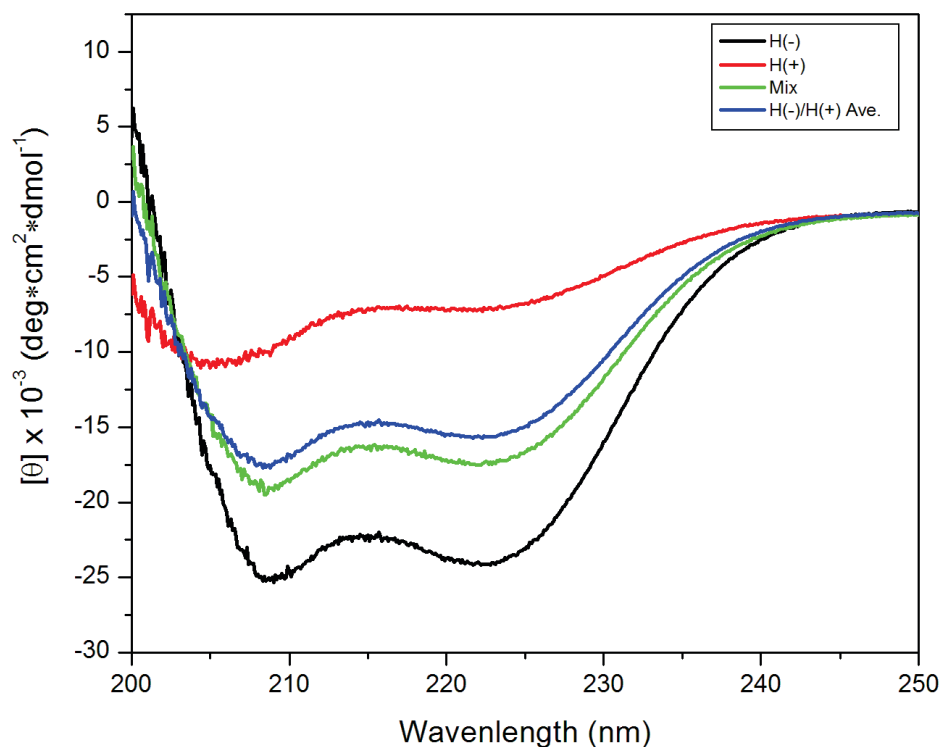
**Figure S1.**

Helical wheel representation and peptide sequences of the designed peptides Helix(-) and Helix(+). For simplicity, only amino acids encoding the coiled coil structure are shown in the helical wheel representation. Complementary electrostatic pairs incorporated to control orientation are outlined by boxes. Heptad repeats are denoted above their position in the peptide sequence. The 6xHis tag is found near the N-terminus and the thrombin cleavage site is underlined.



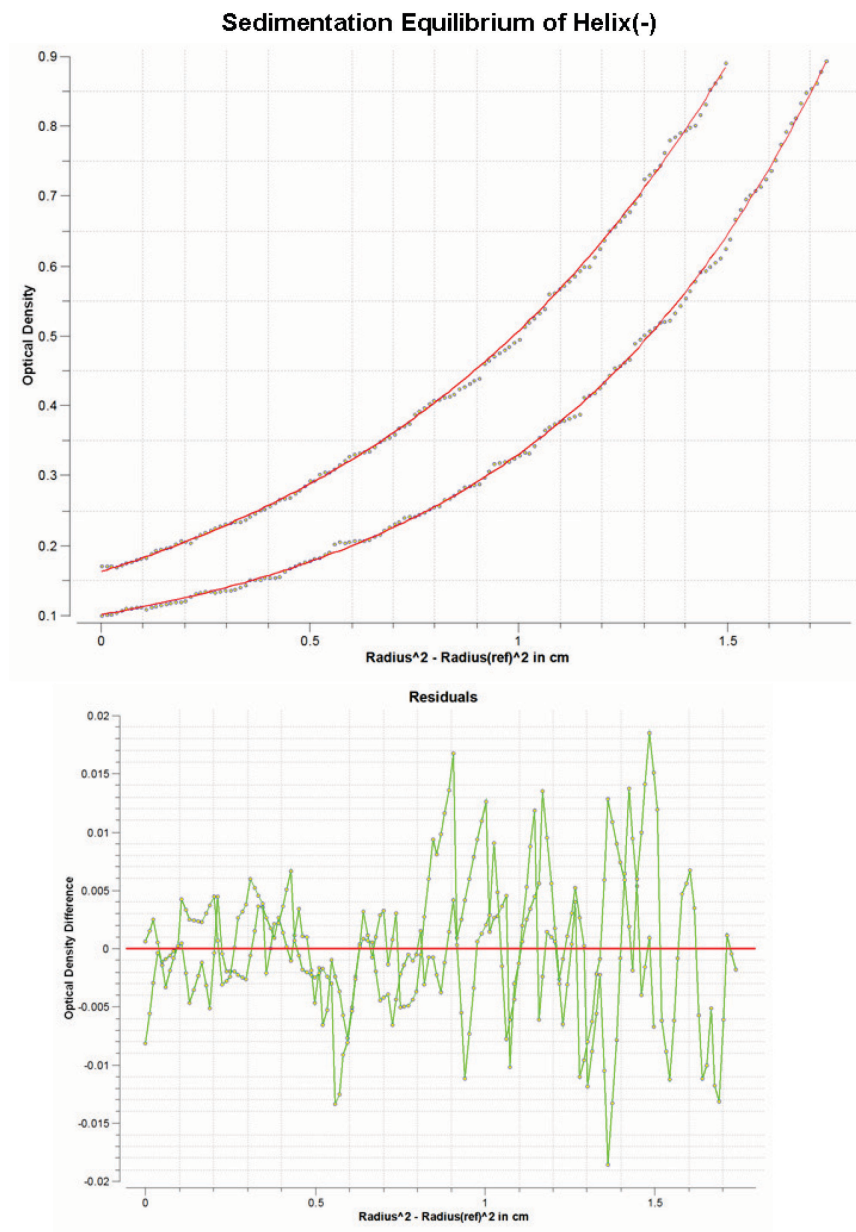
**Figure S2.**

SDS-PAGE and ESI-MS data analysis for Helix(-) and Helix(+). A) ESI-MS spectrum and tris-tricine gel of Helix(-) (inset) purified via Ni-affinity chromatography. B) ESI-MS and Tris-Tricine gel (inset) of Helix(+) purified via Ni-affinity chromatography. Deconvolution of the peaks gives a mass of 7815 Da for Helix(-) and 7687.8 for Helix(+) consistent with masses expected for removal of Met from the N-terminus for both and an additional removal of Ser from the C-terminus of Helix(+), 7813.7 Da and 7688.06 Da respectively.



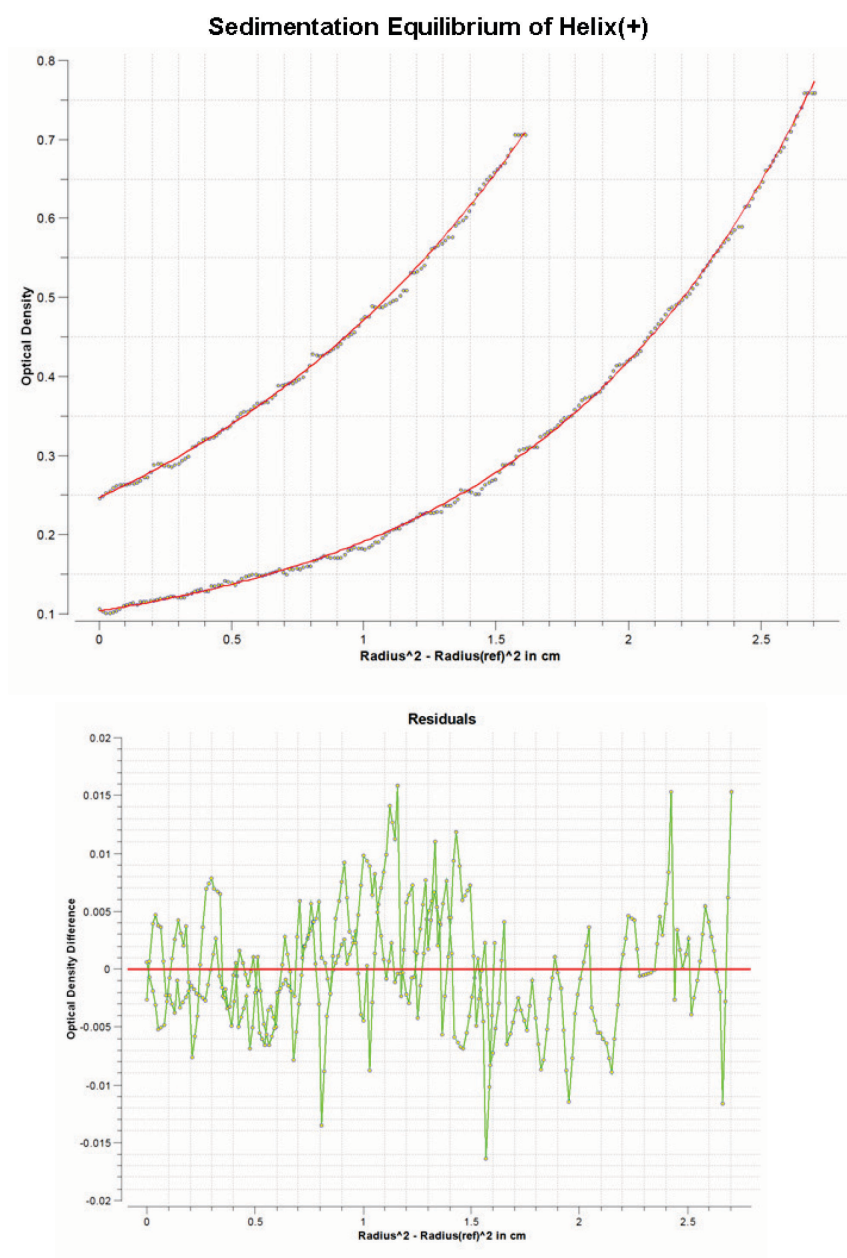
**Figure S3.**

C.D. spectra and analysis of the designed peptides Helix(-) and Helix(+). Helix(-) (black trace) spectra shows helical structure and Helix(+) (red trace) shows mostly random coil structure. The spectrum for the 1:1 mixture of Helix(-) and Helix(+) (green trace) is more helical than expected for the average of the spectra for the two individual peptides (blue trace). The difference in intensities indicate that 15 - 20 % of Helix(-) forms heterodimeric coiled coils with Helix(+), with the remainder Helix(-) forming homodimeric coiled-coils.



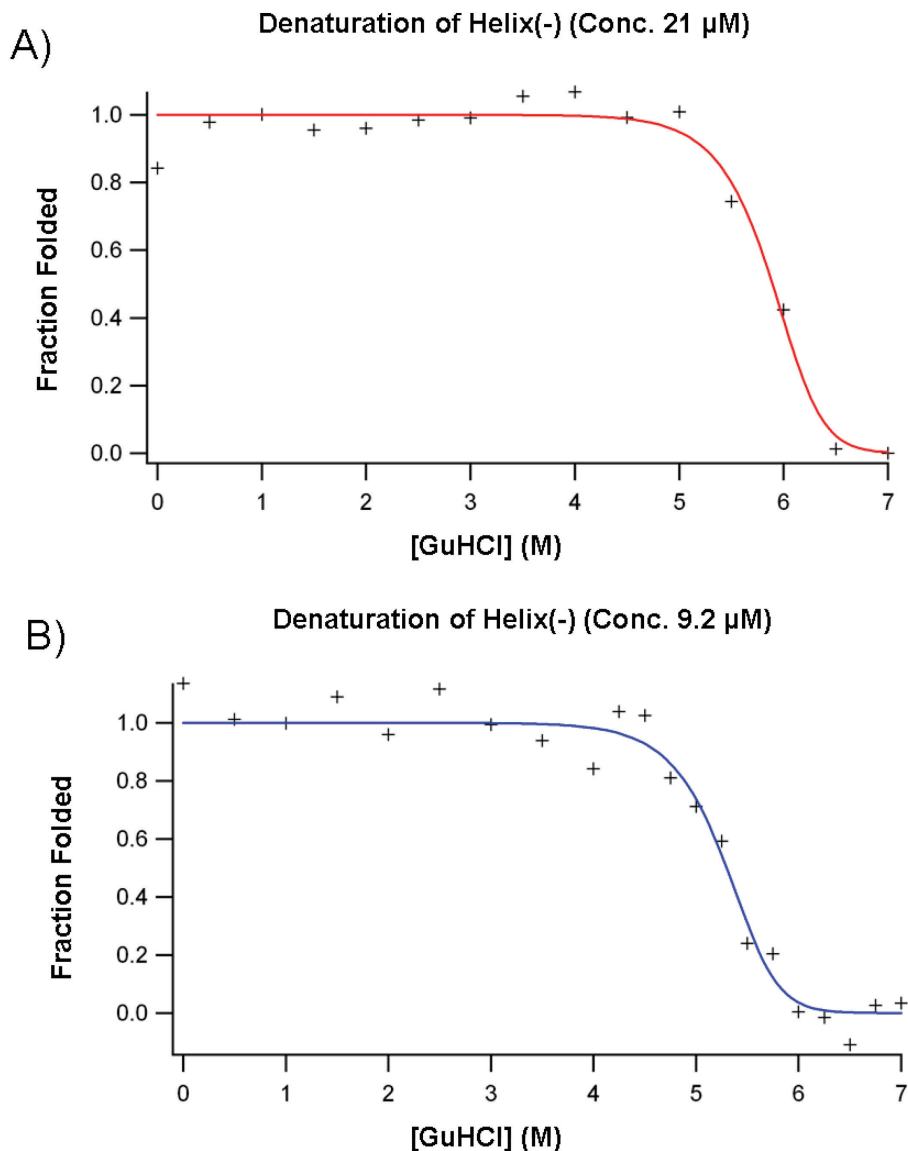
**Figure S4.**

SE AUC data and fits with residuals for Helix(-). Helix(-) SE AUC data traces (blue circles) are at speeds of 35,000 and 40,000 rpm fit to monomer-dimer equilibrium model (red line). Molecular weight of 15,180 Da determined from fit with an expected molecular weight of 15,634 Da for a dimer. Plots of the residuals for the fits of the scans are given below the sedimentation equilibrium scan data and indicate good fitting to the selected model.



**Figure S5.**

SE AUC data and fits with residuals for Helix(+). Helix-(+) SE AUC data traces (blue circles) at speeds of 30,000 and 35,000 rpm fit to a single component species (red line). Molecular weight of 14,430 Da determined from fit with an expected molecular weight of 15,376 for dimer. Plots of the residuals for the fits of the scans are given below the sedimentation equilibrium scan data and indicate good fitting to the selected model.



**Figure S6.**

Denaturation of Helix(-) and determination of  $\Delta G_{\text{unfold}}$  and  $m$ . A) Fraction folded data for Helix(-) at a concentration of 21  $\mu$ M and fit to a dimeric species. B) Fraction folded data for the Helix(-) at a concentration of 9.2  $\mu$ M and fit to a dimeric species. The fits gave  $\Delta G_{\text{unfold}}$  of  $24.5 \pm 2.6$  kcal/mol and  $26.8 \pm 3.9$  kcal/mol and  $m$  values of  $3.3 \pm 0.45$  kcal mol<sup>-1</sup> M<sub>GuHCl</sub><sup>-1</sup> and  $3.5 \pm 0.65$  kcal mol<sup>-1</sup> M<sub>GuHCl</sub><sup>-1</sup> for the data at peptide concentrations of 9.2  $\mu$ M and 22  $\mu$ M, respectively.