J|A|C|S
ARTICLES
Published on Web 01/04/2006

Computational Design of a Single Amino Acid Sequence that Can Switch between Two Distinct Protein Folds

Xavier I. Ambroggio and Brian Kuhlman*

Contribution from the Department of Biochemistry and Biophysics, University of North Carolina, Campus Box 7260, Chapel Hill, North Carolina 27599

Received July 14, 2005; E-mail: bkuhlman@email.unc.edu

Abstract: The functions of many proteins are mediated by specific conformational changes, and therefore the ability to design primary sequences capable of secondary and tertiary changes is an important step toward the creation of novel functional proteins. To this end, we have developed an algorithm that can optimize a single amino acid sequence for multiple target structures. The algorithm consists of an outer loop, in which sequence space is sampled by a Monte Carlo search with simulated annealing, and an inner loop, in which the effect of a given mutation is evaluated on the various target structures by using the rotamer packing routine and composite energy function of the protein design software, RosettaDesign. We have experimentally tested the method by designing a peptide, Sw2, which can be switched from a 2Cys-2His zinc finger-like fold to a trimeric coiled-coil fold, depending upon the pH or the presence of transition metals. Physical characterization of Sw2 confirms that it is able to reversibly adopt each intended target fold.

Introduction

Conformational switches are used by proteins to modulate function in a variety of ways including controlling access to active sites, generating cooperativity, and creating mechanical force. Because protein switches are essential to many cellular processes, there is considerable interest in learning how to design switches. In general, there are two basic features that all protein switches share: their amino acid sequences are compatible with multiple structures, and the relative stabilities of these structures are sensitive to binding events and/or covalent modification. On the latter topic considerable progress has been made in exploiting allosteric systems in natural proteins by the computational design of altered ligand specificity, novel ligand binding sites, and mutations of residues involved in the allosteric mechanism.[1−4] An excellent illustration of design studies concerned with the former topic of finding amino acid sequences compatible with multiple structures are peptide switches that have been designed to undergo a transition from α-helix to associating β-strands.[5−8] Many amino acid sequences are compatible with both α-helices and β-strands, but different sets of interactions stabilize the two states so that changes in environment differentially perturb the stability of the two states.

One approach for identifying an amino acid sequence that is compatible with two structures is to create a sequence that is a hybrid of the naturally occurring sequences for the two target structures. By examining multiple sequence alignments it is possible to identify which residues are most important to each fold and, therefore, which amino acids should be placed at each position in the hybrid sequence. This approach was used to aid the design of globular proteins that switch from being mostly β-sheet to all helical with less than 50% of their residues mutated.[9,10] Taking this approach a step further, Hori and Sugiura searched for proteins from the PDB that have sequences that align with the core motif of 2Cys-2His (2C-2H) zinc fingers.[11] By introducing key residues from the zinc finger sequence into their top match, they were able to create a protein that switched from being all-helical to a zinc finger-like structure with the addition of zinc. A similar approach was recently used to create a protein designed to switch between a coiled coil and a 4H zinc finger.[12]

We have developed an alternative approach for designing protein switches that makes use of computational methods for side-chain packing and design. There has been considerable success in the area of computational protein design as a variety of computer programs have been developed for identifying low-energy sequences for target protein structures.[13,14] With these algorithms it is possible to identify sequences that have

(1) Marvin, J. S.; Hellinga, H. W. *Nat. Struct. Biol.* **2001**, *8*, 795−798.
(2) Zhang, J.; Campbell, R. E.; Ting, A. Y.; Tsien, R. Y. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 906−918.
(3) Looger, L. L.; Dwyer, M. A.; Smith, J. J.; Hellinga, H. W. *Nature* **2003**, *423*, 185−190.
(4) Liu, H.; Schmidt, J. J.; Bachand, G. D.; Rizk, S. S.; Looger, L. L.; Hellinga, H. W.; Montemagno, C. D. *Nat. Mater.* **2002**, *1*, 173−177.
(5) Zhang, S.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 23−28.
(6) Dado, G. P.; Gellman, S. H. *J. Am. Chem. Soc.* **1993**, *115*, 12609−12610.
(7) Mutter, M.; Gassmann, R.; Buttkus, U.; Altmann, K. H. *Angew. Chem., Int. Ed. Engl.* **1991**, *30*, 1514−1516.
(8) Cerpa, R.; Cohen, F. E.; Kuntz, I. D. *Folding Des.* **1996**, *1*, 91−101.
(9) Dalal, S.; Balasubramanian, S.; Regan, L. *Nat. Struct. Biol.* **1997**, *4*, 548−552.
(10) Dalal, S.; Regan, L. *Protein Sci.* **2000**, *9*, 1651−1659.
(11) Hori, Y.; Sugiura, Y. *J. Am. Chem. Soc.* **2002**, *124*, 9362−9363.
(12) Cerasoli, E.; Sharpe, B. K.; Woolfson, D. N. *J. Am. Chem. Soc.* **2005**, *127*, 15008−15009.
(13) Park, S.; Xi, Y.; Saven, J. G. *Curr. Opin. Struc. Biol.* **2004**, *14*, 487−494.
(14) Pokala, N.; Handel, T. M. *J. Struct. Biol.* **2001**, *134*, 269−281.

10.1021/ja054718w CCC: $33.50

well-packed hydrophobic cores and that satisfy the hydrogen-bonding potential of buried polar groups. In general, these methods have been used to identify sequences that are optimal for a single target structure or complex. Particularly relevant to our design goals, protein design algorithms have also been used to explicitly stabilize one protein conformation relative to alternative conformations to increase specificity for the desired state[15,16] or to create a switch between a stable conformation and an unstructured peptide.[17] This was accomplished in one case by setting up a two-layer optimization protocol.[18] The inner layer was used to determine the energy of a fixed sequence threaded on to the various competing structures. The outer layer then used these energies to compare the favorability of one sequence to that of another. To design protein switches we have developed a similar type of algorithm; the key difference is that, instead of favoring one structure over another, we aim to make both structures favorable.

Our procedure makes use of the RosettaDesign program. RosettaDesign uses a Monte Carlo optimization procedure to search for amino acid side-chain rotamers that pack well together and form low-energy hydrogen bonds. RosettaDesign has been used previously to stabilize naturally occurring proteins, perturb protein folding pathways, and design a protein with a novel fold.[19,20] In this study we use RosettaDesign to thread and pack the sequences of interest on the target structures and use the resulting energies along with an additional Monte Carlo search to find sequences that are compatible with both target structures. To experimentally validate the protocol we have designed a sequence (Sw2) that can switch between a coiled coil and a 2C-2H zinc finger.

The 2C-2H zinc finger fold and the coiled-coil fold are among the simplest and most thoroughly studied protein folds. Both folds have been used as targets in successful computational de novo design studies.[21,22] Their ability to accommodate great sequence variability makes them attractive targets for designing a conformational switch. The 2C-2H zinc finger fold contains an N-terminal β-hairpin followed by a short α-helix. The zinc binding site is buried between the helix and hairpin and is formed by two cysteines from the hairpin and two histidines from the helix.[23] The coiled-coil fold is formed by two or more chains of α-helices that wrap around each other.[24] Many coiled coils contain a heptad repeat ($a-g$) in which the *a* and *d* positions (typically leucines) form the hydrophobic core. In this study we use the three-helix coiled coil from hemagglutinin as our coiled-coil template.

Materials and Methods

1. Computational Methods. The switch algorithm uses a Monte Carlo optimization protocol with simulated annealing. Starting from a random amino acid sequence, single amino acid mutations are evaluated by threading the new sequence onto the multiple target structures, and determining if the sum of the threading energies compares favorably (based on the Metropolis criterion) to the energy of the previous sequence. Threading is performed using predetermined sequence-structure mapping and RosettaDesign's side-chain repacking protocol.[20] The threading energy is set equal to the standard RosettaDesign energy[20] which is dominated by a Lennard-Jones potential, a torsion potential, an explicit hydrogen-bonding term, and an implicit solvation model. As the simulation progresses, the temperature is lowered to ensure convergence to low-energy sequences. The energy of a typical simulation for a 50-residue protein usually nears convergence after 1000 mutations, every amino acid having been sampled at every sequence position. The convergence criterion is set to end the simulation when there is no improvement in energy after each amino acid possible for each sequence position is attempted four times.

Sw2 was designed to form a trimeric coiled coil based on the backbone coordinates of residues 13−44 of hemagglutinin (PDB ID: 1HTM)[25] in the absence of zinc and a zinc finger based on the backbone coordinates of residues 3−33 of the Zif268 zinc finger-DNA complex (PDB ID: 1AAY)[26] in the presence of zinc. To preserve the metal binding site in the zinc finger-like state, the rotamers corresponding to the side chains of the Cys and His metal ligands in the zinc finger target structure were held fixed while only the amino acid type was fixed for the coiled-coil states. The structural alignment was determined by performing a sequence alignment of residues 3−33 of Zif268 with the coiled-coil segment of hemagglutinin (residues 3−65) using ClustalX.[27] To make use of the extensive multiple sequence alignment information available for zinc fingers and coiled coils, we also performed a set of simulations in which an additional scoring term was added to the Rosetta energy function to favor amino acids most often seen in coiled coils and zinc fingers. The new score was derived from the probability (*P*) of seeing each amino (*aa)* at each sequence position (*sp*) in the multiple sequence alignments:

$$E_{msa}(aa,sp) = -\ln(0.5P_{ZnFin}(aa,sp) + 0.5P_{cc}(aa,sp))$$

A structure-based multiple sequence alignment of 2C-2H zinc finger domains[28] was used for the zinc finger target structure, and probabilities derived from tropomysoin, myosin, paramyosin, and intermediate filaments[29] were used for the coiled-coil target structure. During the design, 160 independent simulations were performed of which Sw2 had the lowest energy. The multiple sequence alignments in Figure 1 were created using ClustalX.[27]

2. Experimental Methods. 2.1. Sample Preparation. Peptides were purchased from the GenScript Corporation (Piscataway, NJ) where they were synthesized manually on Wang resin through Fmoc chemistry. Cleavage and deprotection were done by treating the Wang resin with trifluoroacetic acid (TFA) in the presence of appropriate scavengers. The peptides were analyzed by reverse phase HPLC on C18 coulmns and mass spectrometry (ESI-MS) using a SHIMADZU LC−MS 2010A. To prevent disulfide formation all samples were made with at least 500 μM TCEP. Standard assays with Ellman's reagent[30] confirmed that cysteines were reduced. The pH was adjusted with NaOH or H_2SO_4. To reduce the propensity of zinc-loaded SW2 to aggregate, all experiments with zinc, and most without zinc were performed in

(15) Shimaoka, M.; Shifman, J. M.; Jing, H.; Takagi, J.; Mayo, S. L.; Springer, T. A. *Nat. Struct. Biol.* **2000**, *7*, 674−678.
(16) Summa, C. M.; Rosenblatt, M. M.; Hong, J.-K.; Lear, J. D.; DeGrado, W. F. *J. Mol. Biol.* **2002**, *321*, 923−938.
(17) Signarvic, R. S.; DeGrado, W. F. *J. Mol. Biol.* **2003**, *334*, 1−12.
(18) Havranek, J. J.; Harbury, P. B. *Nat. Struct. Biol.* **2003**, *10*, 45−52.
(19) Kuhlman, B.; Baker, D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10383−10388.
(20) Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D. *Science* **2003**, *302*, 1364−1368.
(21) Dahiyat, B. I.; Mayo, S. L. *Science* **1997**, *278*, 82−87.
(22) Harbury, P. B.; Plecs, J. J.; Tidor, B.; Alber, T.; Kim, P. S. *Science* **1998**, *282*, 1462−1467.
(23) Michael, S.; Kilfoil, V.; Schmidt, M.; Amann, B.; Berg, J. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 4796−4800.
(24) Crick, F. *Acta Crystallogr.* **1953**, *6*, 689−697.
(25) Bullough, P. A.; Hughson, F. M.; Skehel, J. J.; Wiley, D. C. *Nature* **1994**, *371*, 37−43.
(26) Elrod-Erickson, M.; Rould, M. A.; Nekludova, L.; Pabo, C. O. *Structure* **1996**, *4*, 1171−1180.
(27) Chenna, R.; Sugawara, H.; Koike, T.; Lopez, R.; Gibson, T. J.; Higgins, D. G.; Thompson, J. D. *Nucl. Acids Res.* **2003**, *31*, 3497−3500.
(28) Krishna, S. S.; Majumdar, I.; Grishin, N. V. *Nucleic Acids Res.* **2003**, *31*, 532−550.
(29) Cohen, C.; Parry, D. A. *Proteins* **1990**, *7*, 1−15.
(30) Habeeb, A. F. S. A. Reaction of protein sulfhydryl groups with Ellman's reagent. *Methods in Enzymology*; Academic Press: New York, 1972; Vol. 25, pp 457−464.

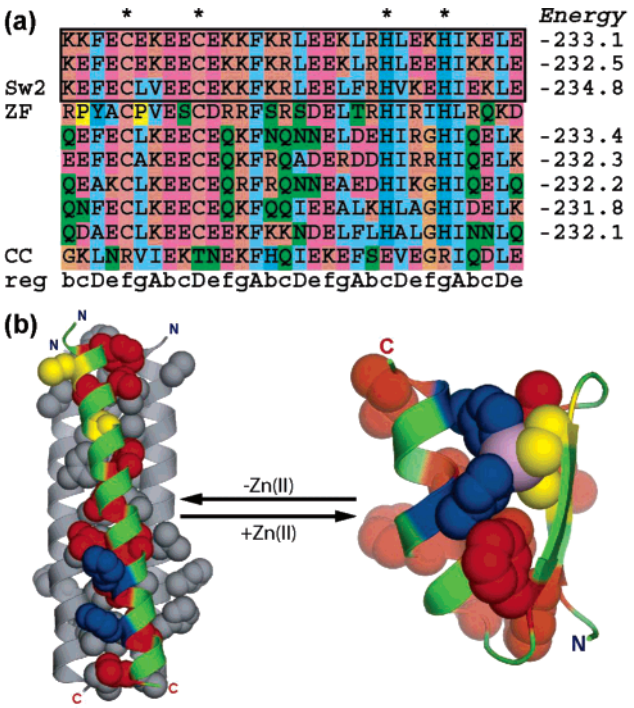

Figure 1. (a) Multiple-sequence alignment of the lowest-energy sequences produced by the switch algorithm with (black box) and without the E_{msa} term, and the native sequences of the target structures (ZF is residues 3−33 of Zif268; CC is residues 13−44 of hemagglutinin). The indicated energies are the sum of the RosettaDesign energies derived from packing the sequence on the two target structures. (b) Models of the Sw2 sequence threaded onto the coiled coil (left) and zinc finger (right) backbones. Hydrophobic residues are colored red, cysteine ligands are colored yellow, histidine ligands are colored blue, and the Zn(II) ion is represented as an exaggerated sphere colored in violet. For clarity, only one of the three helices is colored in the coiled-coil model.

200 mM guanidine hydrochloride. PIPPS, MES, or MOPS was used as a buffer, depending on the target pH. Because zinc phosphate has very low solubility, no phosphate buffers were used in any of the experiments. $ZnSO_4$ was used for zinc titrations, and $CoCl_2$ was used for cobalt experiments.

2.2. Circular Dichroism Spectroscopy. Circular dichroism spectra were obtained using an Applied Photophysics PiStar-80 spectrometer with a thermoelectric temperature controller utilizing an external water bath. Unless stated otherwise, all measurements were collected at 25 °C. All spectra have been converted to mean-residue ellipticity $[\theta_{obs}]$ according to the formula:

$$[\theta_{obs}] = \frac{\theta_{obs} mw}{10clN}$$

where θ_{obs} is the measured ellipticity in mdeg, mw is the size of the protein in Daltons, c is the peptide concentration in mg/mL, l is the path length of the cuvette in centimeters, and N is the number of amino acids in the peptide.

Full spectra of two samples were recorded from 205 to 250 nm (Figure 2). Two spectra were taken of the first sample which consisted of 300 μM Sw2 in a buffer of 20 mM PIPPS pH 4.8, 3 mM TCEP (2× pH 5 buffer) at 20 °C (spectrum 1) and then supplemented with 6 M urea (spectrum 2). The second sample consisted of a 100 μM Sw2 in a buffer of 200 mM guanidine-HCl, 5 mM MES, 5 mM MOPS, 1 mM TCEP. The pH was adjusted using NaOH or H_2SO_4 (spectra 2,3), and $ZnSO_4$ was used for titrations (spectra 4−8). Temperature melts of the first sample and the second sample with or without 200 μM $ZnSO_4$ were measured from 4 to 95 °C at 222 nm (Figure 3).

Concentration dependence of Sw2 was measured at 222 nm in pH 5 buffer, pH 5 buffer supplemented with 200 mM guanidine-HCl, and

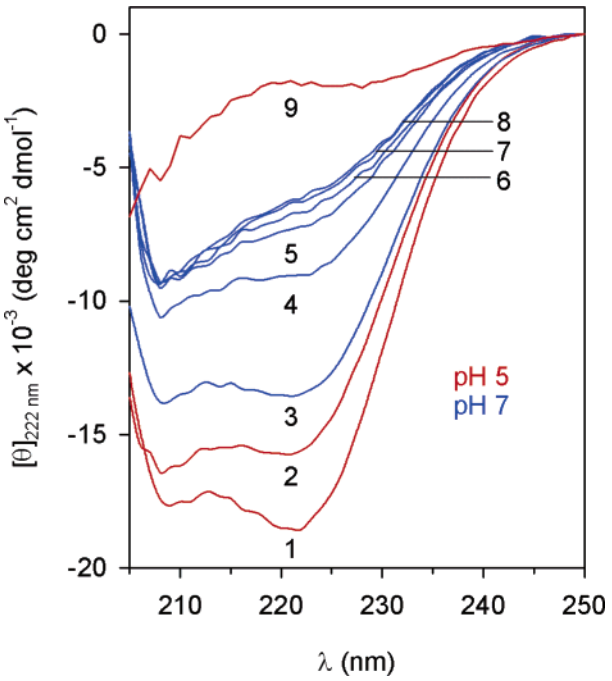

Figure 2. CD spectra showing that Sw2 undergoes a metal dependent conformational change. Spectra **1** and **9** are of 300 μM Sw2 respectively without and with 6 M urea. Spectrum **2** is of 100 μM Sw2 (in 200 mM guanidine) at pH 5 that was titrated to pH 7 (spectrum **3**) and then titrated in spectra **4**−**8** to final Zn(II) concentrations of 25 μM (**4**), 50 μM (**5**), 75 μM (**6**), 100 μM (**7**), and 200 μM (**8**).

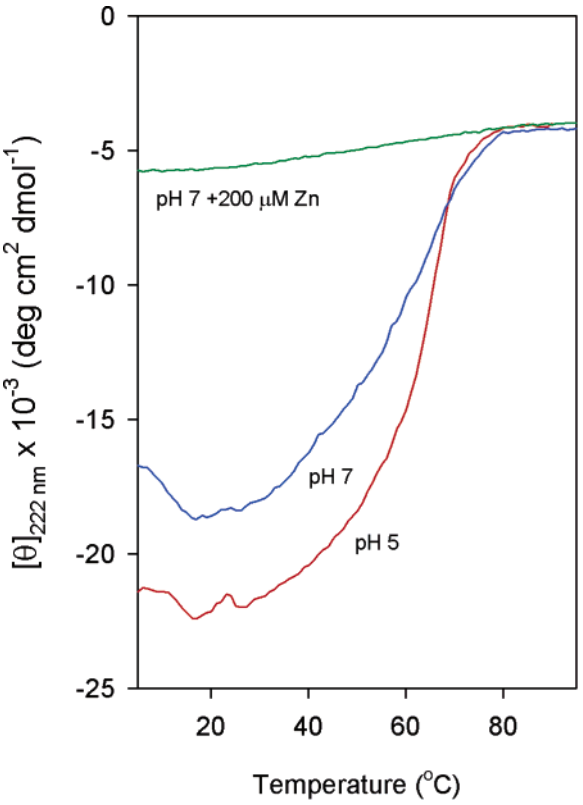

Figure 3. Temperature melts of 100 μM Sw2 at pH 7 and 300 μM Sw2 at pH 5 monitored by CD at 222 nm.

in a buffer of 200 mM guanidine-HCl, 10 mM MOPS pH 7, 1 mM TCEP (pH 7 buffer) with and without 300 μM $ZnSO_4$ (Figure 4). The data for Sw2 in pH 5 buffer were fit to monomer−oligomer equilibrium

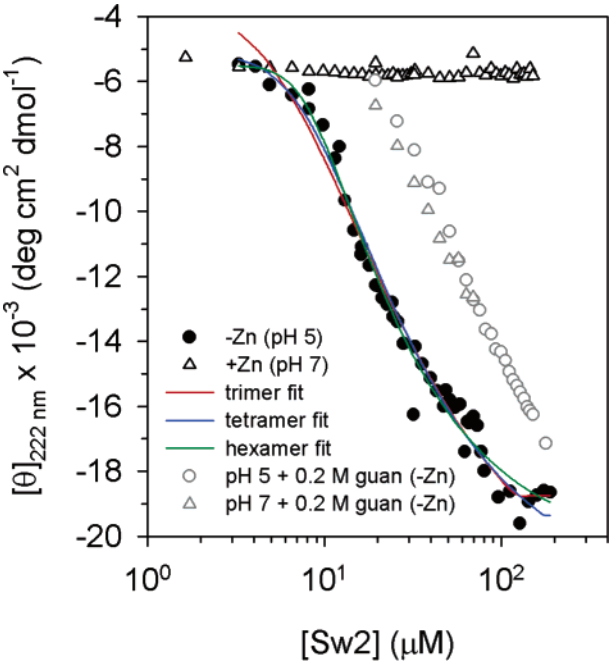

Figure 4. Black symbols, concentration dependence of Sw2 CD signal with zinc and without zinc fit to monomer−oligomer equilibrium models. Gray symbols, concentration dependence of Sw2 without zinc at pH 5 and 7, supplemented with 0.2 M guanidine-HCl.

Table 1. Parameters Derived from CD Data Fit to Oligomeric Equilibrium Models of Sw2 Concentration versus Ellipticity at 222 nm

n-mer	$[\theta_{mon}]^a \times 10^{-3}$	$[\theta_{olig}]^a \times 10^{-3}$	K_D M^{n-1}	R^2
$N = 3$	−3.4	−22.7	4.7×10^{-10}	0.979
$N = 4$	−5.1	−21.7	9.8×10^{-15}	0.981
$N = 6$	−5.5	−20.5	1.3×10^{-24}	0.984

[a] $[\theta_{mon}]$ and $[\theta_{olig}]$ are in units of deg cm^2 dmol^{-1}.

models using the equation shown below where n is the number of chains in the oligomer.[31,32]

$$[\theta_{obs}] = \frac{[\theta_{mon}][P_{mon}]}{[P_{tot}]} + \frac{n[\theta_{olig}][P_{mon}]^n}{K_D[P_{tot}]}$$

The parameters $[\theta_{mon}]$, $[\theta_{olig}]$, and K_D were found empirically for models in which $n > 3$ and fit through least-squares minimization in SigmaPlot 8.02 for the trimer model (Table 1). For the trimer model, the monomer concentration, $[P_{mon}]$, was calculated for a given total peptide concentration, $[P_{tot}]$, by finding the cubic roots to the following equilibrium equation rearranged to apply Cardano's method:

$$[P_{mon}]^3 + \frac{K_D[P_{mon}]}{3} = \frac{K_D[P_{tot}]}{3}$$

$$a = \frac{K_D}{3}; \quad b = \frac{K_D[P_{tot}]}{3}$$

$$[P_{mon}] = \sqrt[3]{\frac{a}{2} + \sqrt{\left(\frac{a}{2}\right)^2 + \left(\frac{b}{3}\right)^3}} + \sqrt[3]{\frac{a}{2} - \sqrt{\left(\frac{a}{2}\right)^2 + \left(\frac{b}{3}\right)^3}}$$

To measure the affinity of Sw2 for zinc, 100 μM Sw2 was titrated with $ZnSO_4$ in pH 7 buffer, and ellipticity measurements were made at 222 nm (Figure 6). The data were fit through least-squares minimization in SigmaPlot 8.02 to the following equations:

$$[\theta_{obs}] = \frac{([\theta_{ZnFin}] - [\theta_{cc}])[P_{Zn}]}{[P_{tot}]} + [\theta_{cc}]$$

$$[P_{Zn}] = \frac{-a - \sqrt{a^2 - 4([P_{tot}][Zn])}}{2}$$

$$a = -[P_{tot}] - [Zn] - K_D$$

The reversibility of the switch was tested by recording spectra from 205 to 250 nm of 130 μM Sw2 in 2 mM PIPPS, 2 mM MES, 2 mM MOPS, 200 mM guanidine-HCl, 500 μM TCEP at pH 7, with 400 μM $CoCl_2$ added, and with 500 μM EDTA added (Figure 7). Changes in secondary structure of zinc-bound Sw2 due to pH were observed by ellipticity measurements at 222 nm of 70 μM Sw2 in the buffer described above supplemented with 200 μM $ZnSO_4$, and titrated with H_2SO_4 or NaOH (Figure 8).

2.3. Analytical Ultracentrifugation. A Beckman XL-A centrifuge equipped with absorbance optics was used for analytical ultracentrifugation measurements. The samples measured at 240 nm were centrifuged at 25000, 30000, and 45000 rpm at 20 °C and were buffered in 10 mM PIPPS, 10 mM MOPS, 200 mM guanidine-HCl, 1 mM TCEP, and adjusted to the desired pH using NaOH. The concentrations of the samples lacking zinc were 70, 140, and 200 μM (Figure 5). The samples containing 600 μM $ZnSO_4$ were at concentrations of 100, 200, and 300 μM (Figure 10b). The samples measured at 220 nm were centrifuged at 45000 rpm at 20 °C in pH 7 buffer and were at concentrations of 10 and 20 μM (Figure 10a). Optima XL-A data analysis software running under Microcal Origin 4.1 was used to fit the data to single-species models. Values for the solvent density and partial specific volume of the peptide were calculated using the Sednterp public domain software package (www.rasmb.bbri.org).

2.4. Cobalt Absorption Spectroscopy. Spectra of Sw2 were recorded using a Varian Cary 50 UV−visible spectrophotometer. To obtain a dissociation constant for the binding of cobalt to Sw2, $CoCl_2$ was titrated into a sample of 130 μM Sw2 in pH 7 buffer, and absorption values were recorded at 640 nm (Figure 6). The data were fit to an equation similar to that described for the zinc titration. The spectrum from 300 to 800 nm was taken with 400 μM $CoCl_2$. All spectra were corrected for protein and free cobalt absorbances in buffer (Figure 9).

Results and Discussion

1. Computational Results. To run the switch algorithm it is first necessary to set which residues map to each other on the target structures. Because of the coiled-coil heptad repeat, each of the seven registers of the coiled coil will present a distinct mapping for the zinc finger sequence. To determine which alignment was most favorable, we created a profile alignment with ClustalX of the first zinc finger domain sequence of Zif268 with the coiled-coil segment of hemagglutinin. In the resulting best alignment, residue 3 of Zif268 mapped to residue 13 in hemagglutinin, a *b* position within the coiled coil. We used this alignment for all subsequent simulations. Next, we ran the switch algorithm with and without a term added to the energy function that incorporated information derived from multiple sequence alignments of both of the target structure protein families (Figure 1a). The low-energy sequences were quite similar, indicating that the additional term was not dominating the energy function. It did, however, lead to a slight preference of hydrophobic residues at positions corresponding to *a* and *d* in the coiled coil, as well as having the unanticipated effect of disfavoring alanine, glutamine, and asparagines residues. The additional energy term also resulted in greater convergence of low-energy sequences.

(31) Fairman, R.; Chao, H. G.; Mueller, L.; Lavoie, T. B.; Shen, L.; Novotny, J.; Matsueda, G. R. *Protein Sci.* **1995**, *4*, 1457−1469.
(32) DeGrado, W. F.; Prendergast, F. G.; Wolfe, H. R., Jr.; Cox, J. A. *J. Cell Biochem.* **1985**, *29*, 83−93.

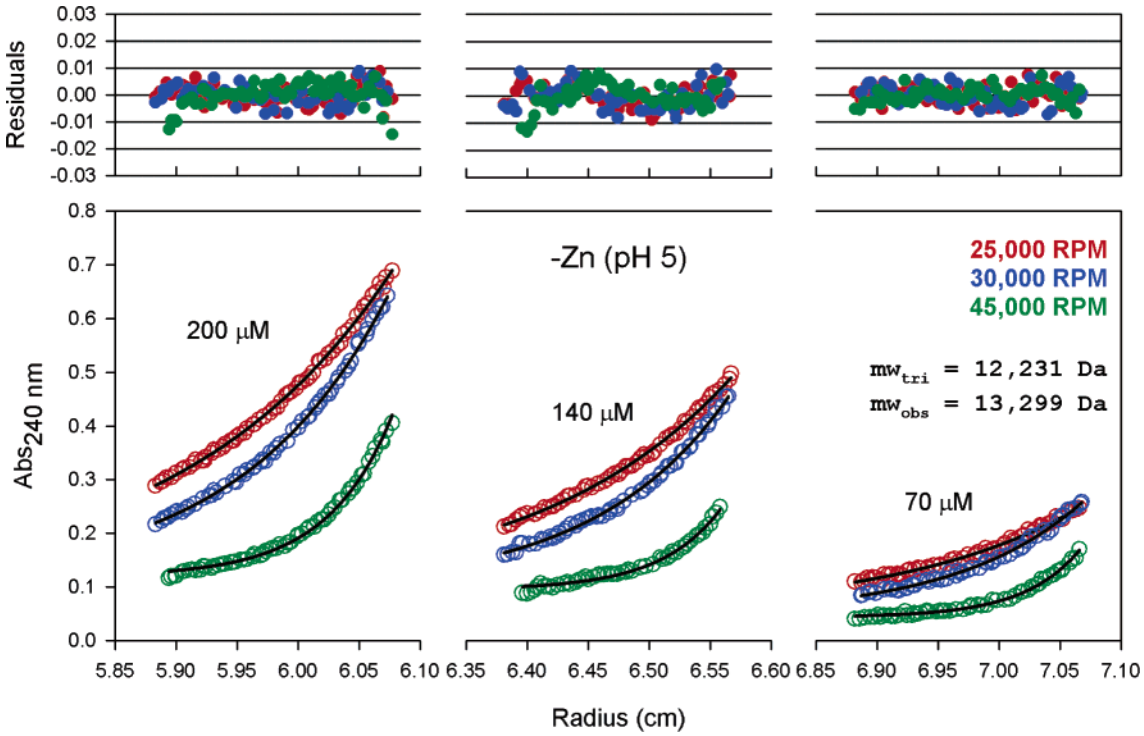

Figure 5. Analytical ultracentrifugation data of Sw2 with a global fit to a single-species model for three concentrations of Sw2 at three speeds.

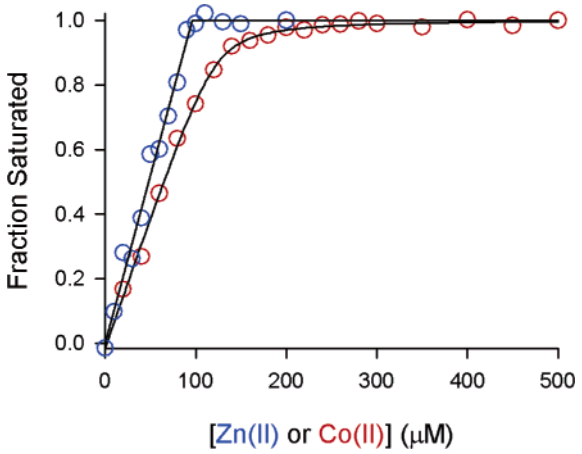

Figure 6. Co(II) and Zn(II) binding of Sw2 at pH 7 fit to single site binding models. Zn(II) binding of 100 μM Sw2 (blue circles) was monitored by ellipticity at 222 nm, and the fit gave an $[\theta]_{ZnFin}$ of -4.4×10^{-3} deg cm^2 dmol^2, an $[\theta]_{cc}$ of -13.3×10^{-3} deg cm^2 dmol^2, and a K_D of 0.44 nm. Co(II) binding of 130 μM Sw2 (red circles) was monitored at the $d \rightarrow d$ transition peak of 640 nm and the fit gave an Abs_{ZnFin} of 0.091, an Abs_{cc} of 0.001, and a K_D of 1.4 μM.

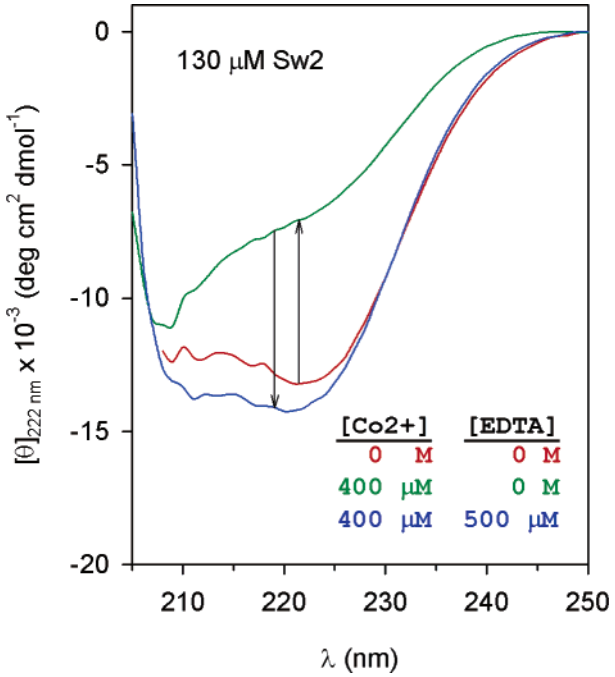

Figure 7. CD spectra showing a reversible switch in the conformation of Co(II) loaded Sw2 (130 μM) at pH 7 by EDTA.

The lowest-energy sequence, Sw2, was selected for synthesis and experimentation. The model of the Sw2 sequence packed onto the coiled-coil state shows that all buried positions are occupied by hydrophobic amino acids (Figure 1b). Both histidines from the metal binding site in the zinc finger align to surface postions on the coiled coil, while one cysteine is surface exposed and one is buried. In the zinc finger state, the buried positions are occupied by the metal binding residues and hydrophobic amino acids. The surface of the zinc finger contains a mixture of polar and hydrophobic amino acids. The hydrophobic amino acids are favored on the surface of the zinc finger by the switch algorithm because these positions map to buried positions in the coiled-coil state, and the penalty in the energy function for burying a polar group is greater than the penalty for placing a surface exposed hydrophobic.

2. Evidence for a Coiled Coil State. In the absence of zinc, the CD spectrum of Sw2 (> 100 μM) shows strong ellipticity minima at 208 and 222 nm indicative of a helical structure (Figure 2). Additionally, Sw2 has a cooperative and reversible thermal denaturation transition as monitored by CD at 222 nm (Figure 3). A cooperative thermal unfolding is indicative of a well-packed protein. Many *de novo* designed proteins do not unfold cooperatively with temperature, and it has been suggested that these proteins are forming molten globules.[33]

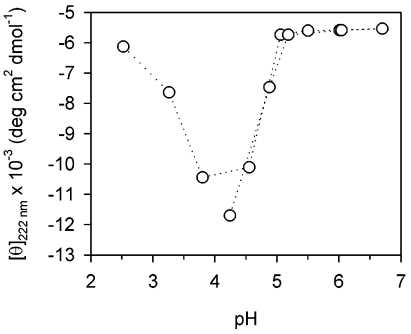

Figure 8. CD spectra monitored at 222 nm showing a pH-dependent, reversible switch in the conformation of 70 μM Sw2 in the presence of 200 μM Zn(II). During the titration the pH was raised from pH 2.5 to 7, and then lowered back to pH 4.2.

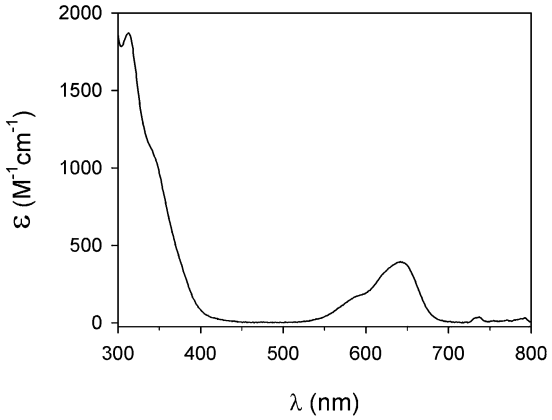

Figure 9. Optical absorption spectrum of 130 μM Sw2 with 400 μM Co(II) at pH 7.

Because coiled coils are oligomeric, their stability is concentration dependent. We examined the concentration-dependent stability of Sw2 by monitoring CD at 222 nm as a function of Sw2 concentration (Figure 4). Below 10 μM the ellipticity at 222 nm is weak suggesting that the protein is unfolded. As the protein concentration is raised from 10 μM to 100 μM the CD signal strengthens significantly suggesting the formation of a helical oligomer. Similar curves are seen at pH 5 and 7, although at pH 7 Sw2 aggregates at concentrations above 130 μM in the absence of Zn(II). Between pH 5 and pH 7 it is expected that the surface exposed histidines on Sw2 will deprotonate to the neutral state which may be promoting the aggregation. Using a two-state transition model for a monomer−oligomer self-associating equilibrium, the concentration dependent CD spectrum of Sw2 can be fit almost equally well to trimeric and higher order oligomers (Table 1).

To discriminate between the various models, analytical ultracentrifugation experiments were performed with Sw2 (Figure 5). The target structure for Sw2 in the absence of zinc was trimeric, with a predicted molecular mass of 12.2 kDa. The AUC data was best fit to a single-species model with a molecular mass of 13.3 kDa suggesting that Sw2 is predominantly trimeric and possibly in equilibrium with a tetrameric state. Taken together, the CD and AUC results provide good evidence that Sw2 forms a coiled coil in the absence of zinc. The structure is helical, unfolds cooperatively, and is oligomeric.

(33) Bryson, J. W.; Betz, S. F.; Lu, H. S.; Suich, D. J.; Zhou, H. X.; O'Neil, K. T.; DeGrado, W. F. *Science* **1995**, *270*, 935−941.

3. Evidence for a Structural Switch. When Sw2 is titrated with Zn(II) at pH 7, the CD spectra undergo a dramatic transition (Figure 2). The helicity decreases, and upon saturation the spectra resemble those of zinc finger proteins with a shoulder at 222 nm and a minimum at 205−208 nm.[34] In addition, the thermal denaturation profile of Sw2 saturated with Zn(II) (Figure 3) is remarkably different from that of Zn(II)-free Sw2 and shows little change in ellipticity over the entire temperature range, consistent with the observed thermal denaturation profiles of a natural 2C-2H zinc finger.[34]

Previously, we had observed some visible aggregation of Sw2 with and without Zn(II) at pH 7 (data not shown). We found that we could prevent this aggregation by adding a small amount of denaturant (urea or guanidine hydrochloride), and thus all experiments performed at pH 7 contained 200 mM guanidine hydrochloride. A binding curve was generated by plotting the CD signal at 222 nm against Zn(II) concentration (Figure 6). The ellipticity shifts linearly with zinc concentration until a one-to-one ratio of zinc to Sw2 concentration is achieved. Because zinc binds so tightly to Sw2, we were not able to determine a dissociation constant for binding. Generally, cobalt also binds to zinc fingers, but with weaker affinity. We followed cobalt binding to Sw2 by monitoring absorbance at the $d \rightarrow d$ transition peak of 640 nm (Figure 6). In this case there is curvature in the binding curve and we were able to fit a single site binding model with a K_D of 1.4 μM with an R^2 value of 0.99. Similar dissociation constants have been observed between naturally occurring zinc fingers and cobalt.[35] The key difference between our results and what is typically seen for zinc fingers is that we see a transition from a helical structure to a zinc finger-like spectra, whereas naturally occurring zinc fingers are generally unfolded in the absence of zinc so that the CD spectra switches from random coil to zinc finger-like spectra.

According to the design, the conformational switch should be reversible since folding of the zinc finger is driven by the formation of a metal binding site. To check for reversibility we added the metal chelating agent EDTA. At pH 7, stoichiometric amounts of EDTA have no effect on the zinc-loaded Sw2 CD spectrum, indicating either that the off rate for zinc from Sw2 is very slow (samples were monitored for approximately 1 h) or that Sw2 binds zinc more tightly than EDTA. Similar results have been observed for other 2C-2H zinc fingers.[36] When the same experiment is performed with cobalt which binds more weakly to Sw2, Sw2 returns to the helical state when a stiochiometric amount of EDTA is added, indicating that the switch is reversible (Figure 7).

We also tested if the state of the switch was pH dependent in the presence of zinc (Figure 8). The affinity of histidine and cysteine for zinc will decrease as the pH is lowered, and naturally occurring zinc fingers have been shown to unfold at lower pHs.[37,38] Monitoring the ellipticity at 222 nm of zinc-loaded Sw2 we observe a sharp change in mean residue ellipticity (from -6 to -12×10^{-3} deg cm^2 dmol^2) between

(34) Frankel, A. D.; Berg, J. M.; Pabo, C. O. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 4841−4845.
(35) Berg, J. M.; Godwin, H. A. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 357−371.
(36) Shi, Y.; Berg, J. M. *Chem. Biol.* **1995**, *2*, 83−89.
(37) Krizek, B. A.; Amann, B. T.; Kilfoil, V. J.; Merkle, D. L.; Berg, J. M. *J. Am. Chem. Soc.* **1991**, *113*, 4518−4523.
(38) Parraga, G.; Horvath, S.; Hood, L.; Young, E.; Klevit, R. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 137−141.

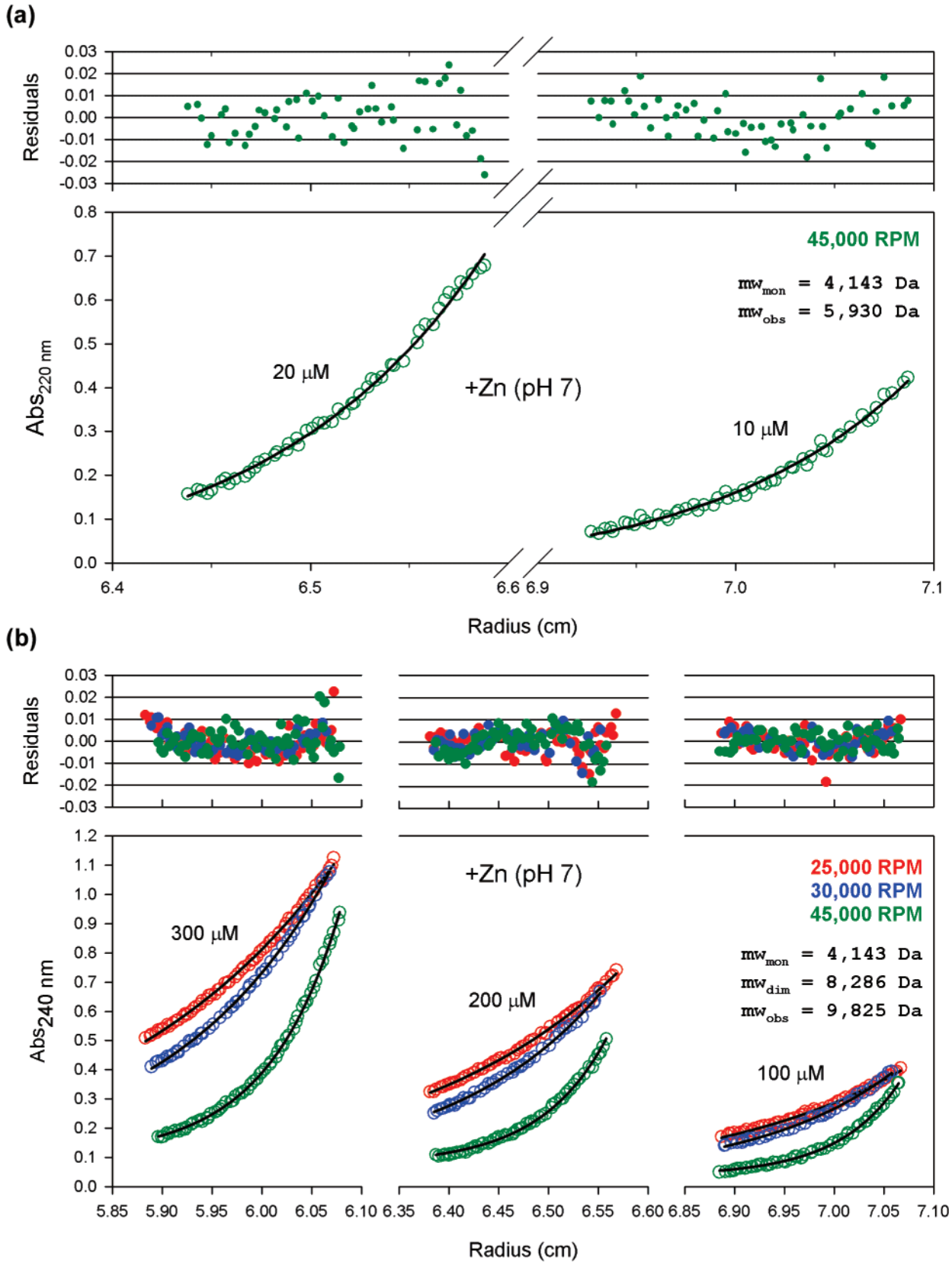

Figure 10. Analytical ultracentrifugation of Zn(II) loaded Sw2 at low concentrations **(a)** and high concentrations **(b)** with global fits to single-species models.

pH 5.0 and 4.5, suggesting that Sw2 switches to the coiled-coil state at pH 4.5. The transition is fully reversible. As the pH is lowered further, the CD signal at 222 nm becomes weaker, suggesting that the coiled coil is unstable below pH 4.

4. Evidence for a Zinc Finger State. To further test whether metal-loaded Sw2 is adopting a zinc finger, we used cobalt absorption spectroscopy to probe the coordination state and ligand environment of the metal binding site.[35] The optical

absorption spectrum of cobalt-substituted Sw2 shows peaks at 310, 340, and 640 nm (Figure 9) similar to the spectrum observed for other C2−H2 zinc fingers.[34,37,38] The peaks at 310 and 340 nm are characteristic of cysteine ligands, and the peak at 640 nm, with an extinction coefficient of ∼415 M^{-1} cm^{-1}, is consistent with the tetrahedral coordination of metals in 2C-2H zinc finger domains.[35] These results strongly suggest that Sw2 is forming a zinc finger-like structure.

According to the design model, zinc binding should provide the driving force for folding rather than interprotein interactions. Indeed, the CD spectra of Sw2 with zinc do not vary significantly with increasing concentrations of protein (Figure 4). This lack of concentration dependence does not rule out the possibility that zinc-bound Sw2 may nonspecifically associate or form a tight oligomer mediated by zinc binding from cysteine ligands between separate chains. The latter hypothesis is unlikely, considering the cobalt absorption spectrum presented above. To test the former possibility we performed analytical ultracentrifugation of zinc-loaded Sw2 at a variety of protein concentrations. The AUC results indicate that Sw2 is primarily monomeric below 20 μM (Figure 10a) but that above 100 μM the apparent molecular weight is near that of a dimer (Figure 10b). The change in apparent molecular weight with increasing protein concentration coupled with the lack of concentration-dependent folding suggests that zinc-bound Sw2 nonspecifically associates above 100 μM. It is interesting to note that zinc-loaded Sw2 is probably aggregation prone because of the hydrophobic amino acids placed on its surface to promote the switch to the coiled-coil state. To prevent this nonspecific association it may be necessary to identify buried positions in the coiled coil that can accommodate polar amino acids by forming a hydrogen-bond network, such as the conserved asparagine residue in GCN4 that has been found to selectively stabilize the coiled coil in a parallel dimeric state.[39] Though, given the strong orientational and distance dependence of hydrogen bonds, this will be a difficult constraint to satisfy.

Conclusion

Conformational switches are central to the function of many naturally occurring proteins. To aid in the design of novel switches, we have developed a computer-based approach for simultaneously optimizing a single amino acid sequence for multiple protein structures. One of the attractive features of our methodology is that it is based on a physical model of proteins, and hence should be applicable toward a wide array of structures and switches. As a first test of our approach we have designed a single amino sequence that is compatible with both the zinc finger fold and a trimeric coiled coil. We have demonstrated with a variety of biophysical techniques that Sw2 undergoes a large and reversible conformational change with the addition of zinc, and all of the data are consistent with this being a switch from a coiled coil to a zinc finger-like structure.

During the preparation of our manuscript a study was published from Derek Woolfson's laboratory that describes a protein (called ZiCo) they have designed to switch between a coiled-coil conformation and a 4His zinc finger.[12] It is interesting to compare the two studies because different design approaches were used in each case. The 27-residue sequence of ZiCo was chosen by merging the consensus sequences of coiled coils and zinc fingers. Aside from the C-terminal histidine residues and the conserved Phe and Leu residues found in the zinc finger motif, there is no significant similarity between ZiCo and Sw2. As does Sw2, ZiCo forms a coiled coil in the absence of zinc. The Sw2 coiled coil appears to be more stable as its T_m is approximately 45 °C higher than that of ZiCo at 100 μM protein concentration. This could partially be because Sw2 is longer, but it also indicates that the switch algorithm effectively optimized the coiled-coil interface. With the addition of zinc, ZiCo converts to a monomer with an IR spectrum that indicates the formation of β-sheet. ZiCo binds zinc more weakly than Sw2, but it is difficult to determine if this is because the overall sequences are differentially optimized for the zinc finger state, or if it is solely because cysteines chelate zinc more tightly than histidines. In total, a comparison of the Sw2 and ZiCo studies suggests that the switch algorithm was able to find sequences that compare favorably with designs derived from merging sequences.

One undesirable characteristic of Sw2 is that it has low solubility in the zinc-loaded state because hydrophobic amino acids that are buried in the coiled coil are exposed in the zinc finger state. This result highlights one of the primary challenges of designing well-behaved conformational switches. The amino acid sequence is a compromise between what is ideal for each target structure. In our current implementation of the switch algorithm we simply minimize the sum of the calculated energies for each structure. One potential danger of this approach is that one structure will be optimized at the expense of the other. Ideally, one would like to optimize the relative stability of the two proteins in addition to the total energy of the system. The challenge is that current energy functions and conformational sampling techniques are not accurate enough to predict the relative free energies of two dramatically different but favorable protein structures. However, now that we have experimental results that provide information about the relative stability and solubility of the two target structures, it should be possible to modify the switch algorithm within RosettaDesign to more aggressively optimize the sequence for one structure versus the other.

Acknowledgment. X.A. and B.K. thank Ashutosh Tripathy at the UNC Macromolecular Interactions Facility and Xiaozhen Hu and Ziad Eletr for helpful discussions. B.K. acknowledges funding from the Searle Scholar's Program and the Beckman Foundation.

JA054718W

(39) Gonzalez, L., Jr.; Brown, R. A.; Richardson, D.; Alber, T. *Nat. Struct. Biol.* **1996**, *3*, 1002−1009.