Cavitands Are Effective Templates for Inducing Stability and Nativelike Structure in de Novo Four-Helix Bundles

Adam R. Mezo and John C. Sherman*

Contribution from the Department of Chemistry, 2036 Main Mall, University of British Columbia, Vancouver, British Columbia, V6T 1Z1, Canada

Received February 16, 1999

Abstract: We have designed, synthesized and characterized eight cavitand-based de novo four-helix bundles, where each helix contains the basis sequence EELLKKLEELLKKG. We find that each de novo protein is highly helical and extremely stable to the chemical denaturant guanidine hydrochloride (GuHCl). We studied the effect of the cavitand−peptide linker on the stability of each de novo protein. Flexible linkers render the helical structures more susceptible to denaturation by GuHCl. Linker structure and length also dictate monomer/dimer equilibria of the proteins. Proteins containing 0−3 Gly units possess varying degrees of nativelike structure. In contrast, proteins containing two or four methylene linkers are more characteristic of molten globules. These differences can be attributed to the additional hydrogen bonding capabilities of the glycine linker variants and to the distance between the cavitand template and the helical bundle. The cavitand pendent group was also changed in one de novo protein from a methyl group to a propyl-phosphate moiety. This alteration does not affect the stability or packing within the helical bundle; however, it does affect the monomer/dimer equilibrium.

Introduction

The protein folding problem remains one the most important unsolved mysteries in biochemistry. De novo protein design offers an attractive route to studying protein folding by testing our current knowledge of the mechanism by which proteins fold.[1] By designing small model tertiary structures, researchers can isolate and study specific protein-stabilizing interactions; such interactions may not be amenable to study in more complex natural proteins as a result of the multitude of stabilizing and destabilizing effects present. Currently, one of the greatest challenges in de novo protein design is moving from the small and fairly well understood secondary structural realm toward the design of larger and more nativelike tertiary structures.

α-Helices have been well studied and thus multi-helical bundles are a logical starting point for the development of a nativelike de novo design research program. De novo helical bundles have been studied by several research groups[2,3] and recently some of the rules governing the formation of nativelike[4] helical bundles have been determined.[5] In addition, there have been several reports using de novo four-helix bundles as scaffolds for catalyzing chemical reactions[6] and for exploring the properties of functionalized heme-proteins.[7]

Since the introduction of template-assembled synthetic proteins (TASPs) by Mutter,[8] the technique of using template molecules to direct peptide strands into desired folds has been used by several researchers to aid in their study of de novo designs.[9] Examples of templates used for α-helical bundles

(1) (a) 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*, 935−94. (b) Harbury, P. B.; Plecs, J. J.; Tidor, B.; Alber, T.; Kim, P. S. *Science* **1998**, *282*, 1462−1467. (c) Struthers, M. D.; Cheng, R. P.; Imperiali, B. *Science* **1996**, *271*, 342−345. (d) Dahiyat, B. I.; Mayo, S. L. *Science* **1997**, *278*, 82−87. (e) Kortemme, T.; Ramirez-Alvarado, M.; Serrano, L. *Science* **1998**, *281*, 253−256.

(2) (a) Eisenberg, D.; Wilcox, W.; Eshita, S. M.; Pryciak, P. M.; Ho, S. P.; DeGrado, W. F. *Proteins: Struct., Funct., Genet.* **1986**, *1*, 16−22. (b) Ho, S. P.; DeGrado, W. F. *J. Am. Chem. Soc.* **1987**, *109*, 6751−6758. (c) Hill, C. P.; Anderson, D. H.; Wesson, L.; DeGrado, W. F.; Eisenberg, D. *Science* **1990**, *249*, 543−546. (d) Osterhout, J. J.; Handel, T.; Na, G.; Toumadje, A.; Long, R. C.; Connolly, P. J.; Hoch, J. C.; Johnson, W. C., Jr.; Live, D.; DeGrado, W. F. *J. Am. Chem. Soc.* **1992**, *114*, 331−337. (e) Betz, S.; Fairman, R.; O'Neil, K.; Lear, J.; DeGrado, W. *Philos. Trans. R. Soc. London B* **1995**, *348*, 81−88. (f) Regan, L.; DeGrado, W. F. *Science* **1988**, *241*, 976−978. (g) Handel, T. M.; Williams, S. A.; Menyhard, D.; DeGrado, W. F. *J. Am. Chem. Soc.* **1993**, *115*, 4457−4460. (h) Handel, T. M.; Williams, S. A.; DeGrado, W. F. *Science* **1993**, *261*, 879−885.

(3) (a) Kamtekar, S.; Schiffer, J. M.; Xiong, H.; Babik, J. M.; Hecht, M. H. *Science* **1993**, *262*, 1680−1685. (b) Hecht, M. H.; Richardson, J. S.; Richardson, D. C.; Ogden, R. C. *Science* **1990**, *249*, 884−891. (c) Fezoui, Y.; Connolly, P. J.; Osterhout, J. J. *Protein Sci.* **1997**, *6*, 1869−1877. (d) Chmielewski, J.; Lipton, M. *Int. J. Pept. Protein Res.* **1994**, *44*, 152−157. (e) Ciesla, D. J.; Gilbert, D. E.; Feignon, J. *J. Am. Chem. Soc.* **1991**, *113*, 3957−3961. (f) Olofsson, S.; Johansson, G.; Baltzer, L. *J. Chem. Soc., Perkin Trans. 2* **1995**, 2047−2056. (g) Kaumaya, P. T. P.; Berndt, K. D.; Heidorn, D. B.; Trewhella, J.; Kezdy, F. J.; Goldberg, E. *Biochemistry* **1990**, *29*, 13−23.

(4) (a) Betz, S. F.; Raleigh, D. P.; DeGrado, W. F. *Curr. Opin. Struct. Biol.* **1993**, *3*, 601−610. (b) Betz, S. F.; Bryson, J. W.; DeGrado, W. F. *Curr. Opin. Struct. Biol.* **1995**, *5*, 457−463.

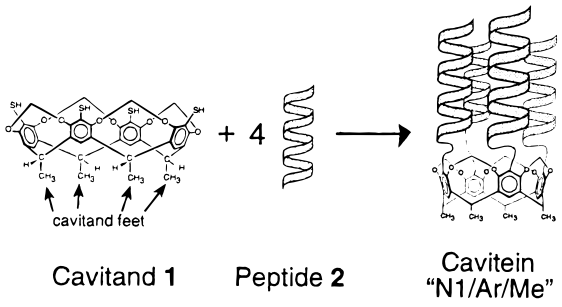
(5) (a) Raleigh, D. P.; Betz, S. F.; DeGrado, W. F. *J. Am. Chem. Soc.* **1995**, *117*, 7558−7559. (b) Betz, S. F.; Liebman, P. A.; DeGrado, W. F. *Biochemistry* **1997**, *36*, 2450−2458. (c) Hill, R. B.; DeGrado, W. F. *J. Am. Chem. Soc.* **1998**, *120*, 1138−1145. (d) Jiang, X.; Bishop, E. J.; Farid, R. S. *J. Am. Chem. Soc.* **1997**, *119*, 838−839. (e) Roy, S.; Ratnaswamy, G.; Boice, J. A.; Fairman, R.; McLendon, G.; Hecht, M. H. *J. Am. Chem. Soc.* **1997**, *119*, 5302−5306. (f) Brive, L.; Dolphin, G. T.; Baltzer, L. *J. Am. Chem. Soc.* **1997**, *119*, 8598−8607. (g) Gibney, B. R.; Rabanal, F.; Skalicky, J. J.; Wand, A. J.; Dutton, P. L. *J. Am. Chem. Soc.* **1997**, *119*, 2323−2324.

(6) (a) Broo, K. S.; Nilsson, H.; Nilsson, J.; Baltzer, L. *J. Am. Chem. Soc.* **1998**, *120*, 10287−10295. (b) Broo, K. S.; Nilsson, H.; Nilsson, J.; Flodberg, A.; Baltzer, L. *J. Am. Chem. Soc.* **1998**, *120*, 4063−4068. (c) Broo, K. S.; Brive, L.; Ahlberg, P.; Baltzer, L. *J. Am. Chem. Soc.* **1997**, *119*, 11362−11372.

(7) (a) Robertson, D. E.; Farid, R. S.; Moser, C. C.; Urbauer, J. L.; Mulholland, S. E.; Pidikiti, R.; Lear, J. D.; Wand, A. J.; DeGrado, W. F.; Dutton, P. L. *Nature* **1994**, *368*, 425−432. (b) Sharp, R. E.; Diers, J. R.; Bocian, D. F.; Dutton, P. L. *J. Am. Chem. Soc.* **1998**, *120*, 7103−7104. (c) Shifman, J. M.; Moser, C. C.; Kalsbeck, W. A.; Bocian, D. F.; Dutton, P. L. *Biochemistry* **1998**, *37*, 16815−16827. (d) Gibney, B. R.; Rabanal, F.; Reddy, K. S.; Dutton, P. L. *Biochemistry* **1998**, *37*, 4635−4643. (e) Rau, H. K.; Haehnel, W. *J. Am. Chem. Soc.* **1998**, *120*, 468−476.

10.1021/ja990487f CCC: $18.00
Published on Web 09/18/1999

Scheme 1



include: cyclic peptides,[10] porphyrins,[11] metal ions,[12] and more recently various phenyl-based structures.[13] Such template assembly offers a different approach to studying de novo protein structure by overcoming large entropic barriers to folding. Hence, such "artificial" stabilization can simplify matters immensely and lead to folded structures that might not otherwise be possible.

We have previously reported[14a] the design and synthesis of parallel de novo four-helix bundle "N1/Ar/Me" using cavitand[15] macrocycle **1**,[16] and demonstrated that it manifests helicity in water (Scheme 1). We named this new family of de novo proteins "*cavi*teins" as a result of the constituent parts (*cavi*tand + pro*tein*).[14]

Very little is known about the relationship between templates and the resulting tertiary structure in TASPs. One recent study concluded that "for suitably long linkers, the shape, size, and directionality of the template are not critical" for four-helix bundle formation.[13] What about short linkers? One major advantage of using cavitands as templates is their extreme rigidity. To exploit this rigidity, a short linker was incorporated into cavitein N1/Ar/Me (see Table 2 for notation).[14] Here, we hold the core peptide sequence constant and explore the effect of the linker by comparing caviteins containing one, two, and four methylene units between the cavitand and the attached peptides. In addition, we incorporated one to three glycine residues in the linkers.

Table 1. List of Peptides Used in Cavitein Synthesis

peptide	sequence[a]
2	$ClCH_2CO$-NH-[*peptide*]-$CONH_2$
3	$ClCH_2CO$-NH-[Gly][*peptide*]-$CONH_2$
4	$ClCH_2CO$-NH-[GlyGly][*peptide*]-$CONH_2$
5	$ClCH_2CO$-NH-[GlyGlyGly][*peptide*]-$CONH_2$
6	$BrCH_2CH_2$-NH-[*peptide*]-$CONH_2$
7	$BrCH_2CH_2CH_2CH_2$-NH-[*peptide*]-$CONH_2$
8	CH_3CO-NH-[*peptide*]-$CONH_2$

[a] "*Peptide*" refers to the sequence: GluGluLeuLeuLysLysLeuGluGluLeuLeuLysLysGly.

Typically, cavitands possess hydrophobic pendent groups (cavitand "foot", Scheme 1) to aid their solubility in organic solvents. Here, cavitein structure is studied in water, so the cavitein prototype[14a] used small methyl groups in an effort to minimize the hydrophobicity of the cavitand. We also report the use of a cavitand bearing propyl-phosphate groups at the foot position.

Results

1. Design. The basis peptide sequence, EELLKKLEELLKKG, was designed[14] to be amphiphilic with glutamic acids and lysines placed four residues apart in the sequence to allow for potential intra- and inter-helical salt-bridges. The C-terminus was amidated to reduce unfavorable interactions with the helix macrodipole, and glycine was incorporated as a C-cap.[17]

Cavitands are extremely rigid macrocycles and hence their incorporation into a four-helix bundle is expected to limit the degrees of freedom afforded to the constituent peptides.[18] Adjacent "rim" functionalities on the cavitand anchor the peptide strands at fixed positions approximately 7 Å apart[19] and are thus well spaced for positioning helices, where inter-helix distances range between 7 and 14 Å.[20]

2. Synthesis. The peptides were synthesized using standard solid-phase methods. For peptides **2**,[14,21] **3**, **4,** and **5**, the free N terminus of the appropriate peptide resin (all protecting groups intact) was reacted with chloroacetyl chloride. For peptides **6** and **7**, the free N terminus of the peptide resin was reacted with 3-bromopropionyl chloride and 5-bromovaleryl chloride, respectively. For peptide **8**, peptide resin was reacted with acetic anhydride (Table 1). Subsequent treatment of each peptide resin with 95% TFA removed the protecting groups and cleaved the peptides from the resin in one step. Each activated peptide was subsequently reacted with cavitand **1**[16] in the presence of DMF and DIEA as base to afford the corresponding caviteins (Table 2).

Two additional caviteins were synthesized using cavitands **10** and **11**. Cavitand **10** has been reported previously;[22] however, a more direct synthetic route is accessible via Sorrell's direct

(8) (a) Mutter, M.; Altmann, E.; Altmann, K.-H.; Hersperger, R.; Koziej, P.; Nebel, K.; Tuchscherer, G.; Vuilleumier, S. *Helv. Chim. Acta* **1988**, *71*, 835−847; Mutter, M.; (b) Mutter, M.; Vuilleumier, S. *Angew. Chem., Int. Ed. Engl.* **1989**, *28*, 535−554.

(9) (a) Schneider, J. P.; Kelly, J. W. *Chem. Rev.* **1995**, *95*, 2169−2187. (b) Tuchscherer, G.; Mutter, M. *J. Biotechnol.* **1995**, *41*, 197−210.

(10) (a) Mutter, M.; Tuchscherer, G. G.; Miller C.; Altmann, K.-H.; Carey, R. I.; Wyss, D. F.; Labhardt, A. M.; Rivier, J. E. *J. Am. Chem. Soc.* **1992**, *114*, 1463−1470. (b) Tuchscherer, G.; Dömer, B.; Sila, U.; Kamber, B.; Mutter, M. *Tetrahedron* **1993**, *49*, 3559−3575.

(11) (a) Sasaki, T.; Kaiser, E. *J. Am. Chem. Soc.* **1989**, *111*, 380−381. (b) Sasaki, T.; Kaiser, E. T. *Biopolymers* **1990**, *29*, 79−88. (c) Akerfeldt, K. S.; Kim, R. M.; Camac, D.; Groves, J. T.; Lear, J. D.; DeGrado, W. F. *J. Am. Chem. Soc.* **1992**, *114*, 9656−9657. (d) Arai, T.; Kobata, K.; Mihara, H.; Fujimoto, T.; Nishino, N. *Bull. Chem. Soc. Jpn.* **1995**, *68*, 1989−1998. (e) Mihara, H.; Nishino, N.; Hasegawa, R.; Fujimoto, T. *Chem. Lett.* **1992**, 1805−1808.

(12) (a) Ghadiri, M. R.; Case, M. A. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 1594−1597. (b) Ghadiri, M. R.; Soares, C.; Choi, C.; *J. Am. Chem. Soc.* **1992**, *114*, 825−831. (c) Ghadiri, M. R.; Soares, C.; Choi, C. *J. Am. Chem. Soc.* **1992**, *114*, 4000−4002. (d) Lieberman, M.; Sasaki, T. *J. Am. Chem. Soc.* **1991**, *113*, 1470−1471.

(13) Wong, A. K.; Jacobsen, M. P.; Winzor, D. J.; Fairlie, D. P. *J. Am. Chem. Soc.* **1998**, *120*, 3836−3841.

(14) (a) Gibb, B. C.; Mezo, A. R.; Sherman, J. C. *Tetrahedron Lett.* **1995**, *36*, 7587−7590. (b) Causton, A. S.; Sherman, J. C. *Bioorg. Med. Chem.* **1999**, *7*, 23−27.

(15) Cavitands are defined as rigid organic macrocycles with an enforced cavity. See: Cram, D. J.; Cram, J. M. *Container Molecules and Their Guests, Monographs in Supramolecular Chemistry*; Stoddart, J. F., Ed.; Royal Society of Chemistry:Cambridge, 1994; Vol. 4.

(16) Gibb, B. C.; Mezo, A. R.; Causton, A. S.; Fraser, J. R.; Tsai, F. C. S.; Sherman, J. C. *Tetrahedron* **1995**, *51*, 8719−8732.

(17) Aurora, R.; Rose, G. D. *Protein Sci.* **1998**, *7*, 21−38.

(18) One particularly exciting aspect of using cavitands as templates is their enforced cavity which has been shown to bind neutral guest molecules. See: (a) Tucker, J. A.; Knobler, C. B.; Trueblood, K. N.; Cram, D. J. *J. Am. Chem. Soc.* **1989**, *111*, 3688−3699. (b) Rudkevich, D. M.; Hilmersson, G.; Rebek, J. *J. Am. Chem. Soc.* **1997**, *119*, 9911−9912. In the future, this enforced cavity may prove useful in the design of caviteins with catalytic properties.

(19) The crystal structure of a tetrabromo analogue of tetrathiol **1** shows that the distance between adjacent bromines is 6.8 Å and that of opposing bromines is 9.6 Å. See: Cram, D. J.; Karbach, S.; Kim, H.-E.; Knobler, C. B.; Maverick, E. F.; Ericson, J. L.; Helgeson, R. C. *J. Am. Chem. Soc.* **1988**, *110*, 2229−2237.

(20) Reddy, B. V. B.; Blundell, T. L. *J. Mol. Biol.* **1993**, *233*, 464−479.

(21) Mezo, A. R.; Sherman, J. C. *J. Org. Chem.* **1998**, *63*, 6824−6829.

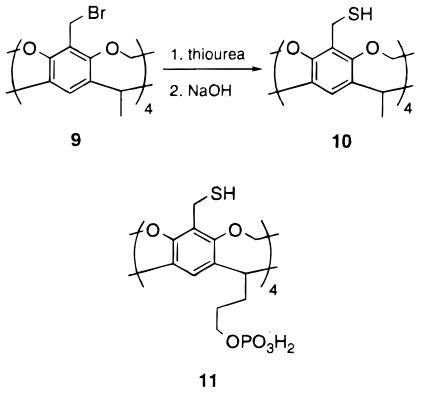
(22) Cram, D. J.; Karbach, S.; Kim, Y. H.; Baczynskyj, L.; Marti, K.; Sampson, R. M.; Kalleymeyn, G. W. *J. Am. Chem. Soc.* **1988**, *110*, 2554−2560.

Table 2. List of Caviteins Synthesized, Their Constituent Parts, Their Cavitand−Peptide Linkages, and Cavitand Feet

cavitein[a]	peptide used	cavitand used	cavitand−peptide linkage[b,c]	cavitand foot[d]
N1/Ar/Me	**2**	**1**	*cavitand*−*(S*-CH_2CO-*peptide)*$_4$	Me
N1G/Ar/Me	**3**	**1**	*cavitand*−*cavitand*−*(S*-CH_2CO-[Gly]-*peptide)*$_4$	Me
N1GG/Ar/Me	**4**	**1**	*cavitand*−*(S*-CH_2CO-[GlyGly]-*peptide)*$_4$	Me
N1GGG/Ar/Me	**5**	**1**	*cavitand*−*(S*-CH_2CO− [GlyGlyGly]-*peptide)*$_4$	Me
N2/Ar/Me	**6**	**1**	*cavitand*−*(S*-$(CH_2)_2$CO-*peptide)*$_4$	Me
N4/Ar/Me	**7**	**1**	*cavitand*−*(S*-$(CH_2)_4$CO-*peptide)*$_4$	Me
N1/Bzl/Me	**2**	**10**	*cavitand*−*(CH*$_2$*S*-CH_2CO-*peptide)*$_4$	Me
N1/Bzl/PO_3H_2	**2**	**11**	*cavitand*−*(CH*$_2$*S*-CH_2CO-*peptide)*$_4$	$(CH_2)_3PO_3H_2$

[a] Each name is indicative of the each cavitein's components. For example, "N1GG" refers to one methylene past the N terminus (N1) with two glycines (GG) as a linker. "Ar" refers to an aryl thiol linkage, while "Bzl" refers to a benzylthiol linkage. "Me" and "PO_3H_2" refer to the type of cavitand foot. All peptides are attached via their N-terminal ends. [b] Italicized words or atoms represent part of the cavitand or core 14-amino acid sequence. [c] "Peptide" refers to the amino acid sequence: NH-[EELLKKLEELLKKG]-$CONH_2$. [d] "Cavitand foot" refers to the pendent group of the cavitand (see Scheme 1).

Scheme 2



bromination method to generate benzyl bromide cavitand **9** (Scheme 2).[23] Subsequent treatment of **9** with thiourea and then NaOH afforded cavitand **10**. Treatment of peptide **2** with cavitand **10** in the presence of DMF and DIEA afforded cavitein N1/Bzl/Me. Cavitand **11**[21] was reacted with peptide **2** in phosphate buffer to afford phosphate-footed cavitein N1/Bzl/PO_3H_2, as described previously.[21]

3. Characterization. i. Far-UV CD Spectra. All eight caviteins possess the signature shape for α-helices: minima at 222 and 208 nm and a maximum near 195 nm (Figure 1). For the prototype cavitein, N1/Ar/Me, the mean residue ellipticity at 222 nm is $[\theta]_{222} = -24\ 000$ deg·cm^2 dmol^{-1}, which is indicative of a highly helical (>75% helicity)[24] structure. Addition of the helix-inducing solvent trifluoroethanol[25] to N1/Ar/Me showed only a ∼20% increase in helicity.[26] Incorporation of 1−3 Gly in the linker successively reduces the absolute helicity, as measured by $[\theta]_{222}$. However, this is likely due to the helix length-dependent equation[24] used in calculating molar ellipticity *per residue*: if each glycine linker is not considered part of the α-helical structure, the $[\theta]_{222}$ values are within 3 000 deg·cm^2 dmol^{-1} of each other (data not shown). Very little differences exist in absolute helicity with incorporation of methylene units into the linker or when the cavitand foot is changed from methyl to propyl-phosphate. Interestingly, the shape of the CD spectra are not the same. Typically, ratios for $[\theta]_{222}/[\theta]_{208}$ that are >1.0 are indicative of coiled-coil-like structures, and ratios <1.0 are indicative of noninteracting helices.[5g,27] However, the cavitand chromophore makes this issue moot, as the aromatic template can make a modest contribution to the CD spectra in this region.[28] This contribution will be linker-dependent;[29] thus $[\theta]_{222}/[\theta]_{208}$ ratios are not informative.

3. ii. Oligomeric States. The helicity of peptide **8** is dependent on concentration from 2 to 500 μM,[26,30] whereas the helicity of N1/Ar/Me is concentration independent from 1 to 175 μM, and all of the caviteins synthesized possess concentration independent CD spectra over at least a 50-fold range. However, sedimentation equilibrium[31,32] experiments (Table 3)[26] indicate that prototype N1/Ar/Me exists in a monomer/dimer equilibrium,[33] N1G/Ar/Me is predominantly dimeric, and N1GG/Ar/Me and N1GGG/Ar/Me are predominantly monomeric. The N2/Ar/Me cavitein is monomeric, while N4/Ar/Me and N1/Bzl/Me are dimers. "Mutation" of the cavitand foot from methyl (N1/Bzl/Me) to propyl-phosphate (N1/Bzl/PO_3H_2) eliminates dimerization. From Table 3, it is apparent that there are considerable errors in using sedimentation equilibrium as a method of molecular weight estimation for these caviteins.[31,34] However, despite these errors, much useful information can be gleaned from the data when taken together with data acquired from other experiments (vide infra).

3. iii. Cavitein Stabilities. The stability of peptide **8** and each cavitein was investigated via denaturation using GuHCl at 25 °C (Figure 2). Prototype N1/Ar/Me is not yet half-unfolded at 8.0 M GuHCl! The incorporation of one and three extra methylenes does not significantly change the stability of the

(23) Sorrell, T. N.; Pigge, F. C. *J. Org. Chem.* **1993**, *58*, 784−785.
(24) Chen, Y.; Yang, J.; Chau, K. *Biochemistry* **1974**, *13*, 3350.
(25) Nelson, J. W.; Kallenbach, N. R. *Proteins* **1986**, *1*, 211−217.
(26) See Supporting Information.
(27) Kohn, W. D.; Kay, C. M.; Hodges, R. S. *Protein Sci.* **1995**, *4*, 237−250.
(28) Woody, R. W. *Biopolymers* **1978**, *17*, 1451−1467.
(29) Chakrabartty, A.; Kortemme, T.; Padmanabhan, S.; Baldwin, R. L. *Biochemistry* **1993**, *32*, 5560−5565.
(30) Oligomerization of single peptides (similar to that of **8**) into helical bundles has been studied extensively by others (refs 2a−e, 3c−e) and thus is not emphasized here. Nevertheless, some elements of the structure and stability of **8** (e.g., GuHCl melts) are reported for comparative purposes.
(31) *Analytical Ultracentrifugation in Biochemistry and Polymer Science*; Harding, S. E., Rowe, A. J., Horton, J. C., Eds.; Royal Society of Chemisty: Oxford, 1992.
(32) We also conducted gel filtration experiments using a Waters Protein Pak 125 column in an attempt to assess the oligomeric state of each cavitein. Unfortunately, we deemed these experiments unreliable as some of the caviteins appeared to interact with the stationary phase.
(33) Due to the errors associated with this experiment, calculation of the association constant for dimerization is not reported.
(34) The largest error arises in estimating υ, the partial specific volume of each cavitein, as the large cavitand component of each cavitein is not considered in estimating υ. In addition, small errors in υ can manifest themselves into large errors in the estimated molecular weight. We therefore conducted a sedimentation equilibrium experiment on N1GG/Ar/Me in the presence of 7.2 M guanidine hydrochloride (GuHCl). At this concentration of denaturant, N1GG/Ar/Me is fully unfolded and therefore is expected to be monomeric in solution. We find that the estimated molecular weight for this experiment is near 11 kDa, suggesting that this may be considered a "baseline" value for this monomeric cavitein. Of course, the presence of a small amount of higher-order aggregates (<10%) may go undetected when analyzed by single-species analysis despite giving rise to the best fit. See: Condino, J. Determination of Molecular Weights by Sedimentation Equilibrium. In *Technical Information DS820*; Beckman Coulter Inc.:Palo Alto, California,1992.

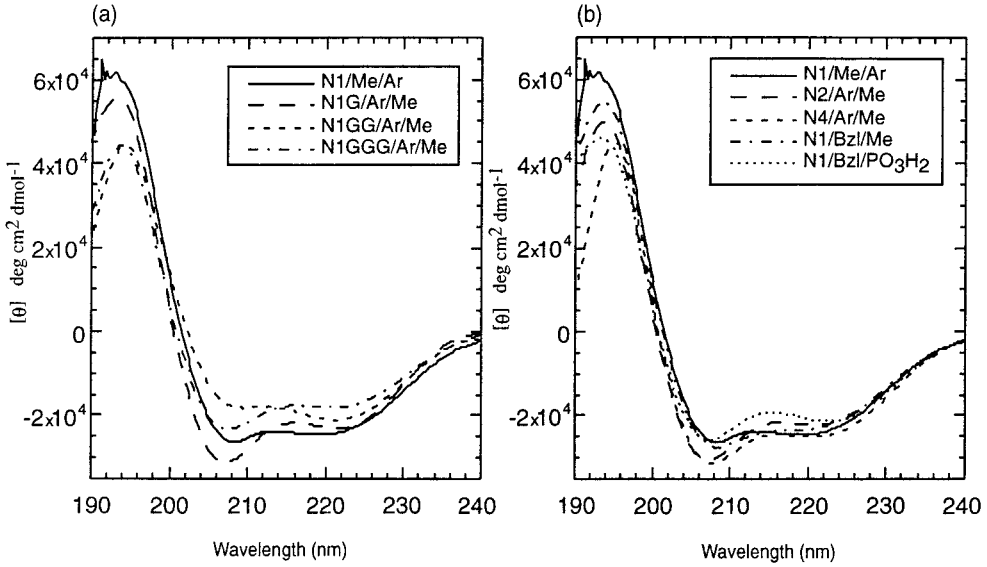


Figure 1. Far-UV CD spectra of (a) N1/Ar/Me and the Gly cavitein variants N1G/Ar/Me, N1GG/Ar/Me, and N1GGG/Ar/Me and (b) N1/Ar/Me, the methylene cavitein variants N2/Ar/Me, N4/Ar/Me, and N1/Bzl/Me and the cavitand foot cavitein variant N1/Bzl/PO_3H_2. Each spectrum was acquired at 25 °C in 50 mM phosphate buffer at pH 7.0 using a cavitein concentration of 30 μM.

helices toward GuHCl (Figure 2b); both N2/Ar/Me and N4/Ar/Me are still not fully unfolded in the presence of 8.0 M GuHCl. From the incorporation of one to the incorporation of two glycines (Figure 2a), the concentration of GuHCl required to half-unfold the helical structure is reduced, whereas on going from two to three glycines, there is little difference in stability. Even the flexible three-Gly linker in N1GGG/Ar/Me shows greatly enhanced stability (half-unfolds at 5.5 M GuHCl) and a cooperative unfolding transition in comparison to single-stranded peptide **8**, which half-unfolds near 1.0 M GuHCl. The unfolding curves of N1/Bzl/Me and N1/Bzl/PO_3H_2 are nearly identical, which demonstrates that changing the cavitand foot, at least in this case, does not affect the stability of the helical bundle.

Concentration-independent GuHCl unfolding experiments (high, 30 μM; low, 3 μM; data not shown) support the sedimentation equilibrium evidence for monomeric caviteins N1GG/Ar/Me and N1GGG/Ar/Me. Interestingly, N1/Ar/Me and N1/Bzl/Me also showed concentration independent unfolding curves, which suggests that in their case, dimerization is *not* contributing significantly to the stability of the four-helix bundle (data not shown). As an additional check, we conducted a sedimentation equilibrium experiment on N1/Ar/Me in the presence of 6.0 M GuHCl. At this concentration of denaturant, N1/Ar/Me is still fully folded (Figure 2a); however, the concentration distribution in the sedimentation equilibrium experiment was best described as a single noninteracting species of molecular weight 12 850 ± 50 Da (Table 3);[26] this molecular weight is indicative of a monomer.[34] We therefore conclude that the apparent dimerization of N1/Ar/Me is mediated by surface interactions that do not interfere with the packing or stability of the four-helix bundle.

Unfolding experiments at high and low concentrations of cavitein N1G/Ar/Me yield a concentration-*dependent* unfolding curve (Figure 3). This suggests that the dimerization of N1G/Ar/Me is an inherent component to the stability of its helical structure. Whereas N1/Ar/Me and N1/Bzl/Me likely go through a dimer-to-monomer-to-unfolded transition, N1G/Ar/Me likely goes through a dimer-to-unfolded transition. Hence, it is likely that N1G/Ar/Me is not forming a four-helix bundle but a helical structure of higher topology.

We estimated the stabilities of the caviteins using the "linear extrapolation method."[35] Calculated values for $\Delta G°_{H_2O}$ are listed in Table 4. Note that some entries were omitted since a full unfolding transition was not observed. The stabilities of N1GG/Ar/Me, N1GGG/Ar/Me, and N1/Bzl/Me are near or more stable than that of previously designed de novo four-helix bundles.[5a,b,f,10a,11b,12c]

The most exciting result is that of N1G/Ar/Me, which displays a much more cooperative unfolding curve than any of the other caviteins (Figure 2a). Assuming a monomeric structure, its $\Delta G°_{H_2O}$ is near double that of the next most stable cavitein. If we calculate its $\Delta G°_{H_2O}$ assuming a two-state dimer-to-unfolded transition process, we find that its stability is in the neighborhood of 40 kcal/mol! This corresponds to a dissociation constant of $\sim 10^{-30}$! Obviously these numbers should not be overinterpreted: (1) there is no theoretical basis for assuming the linear dependence of free energy on denaturant concentration, (2) unfolding may not be an ideal two-state process, and (3) errors increase drastically as the length of extrapolation is increased.[36] Nevertheless, it is clear that N1G/Ar/Me is *extremely* stable (especially considering its dimeric topology) and is the most stable of all the cavitein variants.

Denaturant *m* values have been correlated well with solvent accessible surface areas: larger *m* values generally correspond to larger protein surfaces.[37] This is consistent with the data (vide supra) suggesting that N1G/Ar/Me is a dimer with an overall topology comprising a larger surface area than the other cavitein variants.

We also tested the stability of prototype cavitein N1/Ar/Me to the chemical denaturant urea and found no change (<5% and within error) in $[\theta]_{222}$ from 0 to 10 M urea.[26] The fact that GuHCl is a far more effective denaturant suggests that there may be a considerable electrostatic component to the stability of N1/Ar/Me.[38] Temperature was also used to monitor unfolding using N1/Ar/Me: no unfolding transition was observed even in the presence of 6.0 M GuHCl.[26]

(35) Pace, C. N. *Methods Enzymol.* **1986**, *131*, 266−280.
(36) Santoro, M. M.; Bolen, D. W. *Biochemistry* **1988**, *27*, 8063−8068.
(37) Myers, J. K.; Pace, C. N.; Scholtz, J. M. *Protein Sci.* **1995**, *4*, 2138−2148.
(38) It has been found that the salt GuHCl is able to screen electrostatic interactions at low [GuHCl], while urea, in general, does not: Monera, O. D.; Kay, C. M.; Hodges, R. S. *Protein Sci.* **1994**, *3*, 1984−1991; Zou, Q.; Habermann-Rottinghaus, S. M.; Murphy, K. P. *Proteins* **1998**, *31*, 107−115.

Table 3. Experimentally Estimated Molecular Weights Determined by Sedimentation Equilibrium Experiments at 25 °C[a]

cavitein	cavitein concn and conditions	exptl estimated MW (Da)	calcd MW[B]	predominant species in solution
N1/Ar/Me	30 μM in buffer[c]	7557[d]	7557	monomer−dimer equilibrium
N1/Ar/Me	300 μM in buffer	7557[d]	7557	monomer−dimer equilibrium
N1/Ar/Me	30 μM in 6.0 M GuHCl	12850 ± 50	7557	monomer
N1/Ar/Me	30 μM in 2 M NaCl	8200 ± 50	7557	monomer
N1/Ar/Me	30 μM in 10% methanol	9800 ± 50	7557	monomer
N1G/Ar/Me	3 μM in buffer	18800 ± 1000	7788	dimer
N1G/Ar/Me	30 μM in buffer	18700 ± 300	7788	dimer
N1GG/Ar/Me	30 μM in buffer	10000 ± 600	8019	monomer
N1GG/Ar/Me	30 μM in 7.2 M GuHCl	10800 ± 1400	8019	monomer
N1GGG/Ar/Me	30 μM in buffer	8200 ± 400	8250	monomer
N1GGG/Ar/Me	300 μM in buffer	8900 ± 200	8250	monomer
N2/Ar/Me	30 μM in buffer	10800 ± 100	7615	monomer
N4/Ar/Me	30 μM in buffer	22100 ± 500	7725	dimer
N1/Bzl/Me	30 μM in buffer	19300 ± 100	7615	dimer
N1/Bzl/Me	300 μM in buffer	19400 ± 100	7615	dimer
N1/Bzl/PO_3H_2	30 μM in buffer	11600 ± 400	8112	monomer

[a] See Supporting Information for plots of absorbance vs radius and the residuals for each experiment. [b] These values were calculated from the chemical composition of each cavitein. [c] "Buffer" refers to pH 7, 50 mM phosphate buffer. [d] In this case, the molecular weight was fixed at 7557 and the association constant for dimerization was varied to obtain the best fit.

3. iv. Effect of pH, Salt, and Methanol. For N1/Ar/Me, we monitored $[\theta]_{222}$ as a function of pH (2−12)[26] and KCl concentration (0−3 M, data not shown). Neither had any effect on $[\theta]_{222}$, which suggests that, although electrostatic effects (e.g., salt-bridges) may be relevant, they are not the main contributor to stability. Rather, it is likely that the hydrophobic effect and packing interactions dictate bundle formation and stability.

We further explored the cause of the monomer−dimer equilibrium of N1/Ar/Me by performing sedimentation equilibrium experiments in the presence of 2 M NaCl and 10% methanol (Table 3).[26] Note that neither of these conditions affects the CD spectrum of N1/Ar/Me, which suggests that neither additive affects the bundle structure. In both cases, the concentration distribution of N1/Ar/Me was best described as a single noninteracting species of molecular weight near that of a monomer (Table 3). This suggests that both hydrophobic and electrostatic effects are responsible for the observed aggregation and that these interactions are occurring cooperatively: removal of either of these interactions is sufficient to inhibit dimerization.[39]

3. v. Near-UV CD Spectra. Molten globules[40] typically possess a reduction or absence of near-UV CD bands when compared to nativelike states due to the time-averaged fluxional environment of their aromatic chromophores.[41] The near-UV CD spectra of the glycine and methylene cavitein variants are shown in Figure 4. Prototype N1/Ar/Me displays bands near 265 and 247 nm. The only aromatic chromophores in each cavitein lie in the arenes of the cavitand. The observation of such bands suggests that the cavitand template is in an asymmetric environment, which is indicative of nonaveraged specific structural elements (i.e., nativelike structure) near the peptide-cavitand linkage. The incorporation of glycine linkers augments these aromatic bands, while the incorporation of methylene linkers diminishes them. This suggests that the glycine linker variants manifest nativelike structure near the cavitand−peptide linkage, while the methylene linker variants relegate this region to molten globular states. Interestingly, N1G/Ar/Me displays the most enhanced near-UV absorption of all the caviteins, which suggests that it possesses the most defined structure at the cavitand−peptide linkage. Finally, altering the cavitand foot from methyl to propyl-phosphate does not significantly alter the near-UV bands.

3. vi. ANS Binding. 8-Anilino-1-naphthalenesulfonic acid (ANS) is a hydrophobic dye that has been found to bind only the molten globule state of proteins and not nativelike structures.[42] The binding of ANS has thus been used to probe nativelike protein structure.[5] We observe no binding of ANS to the prototype N1/Ar/Me or any of the Gly cavitein variants under typical conditions for the experiment (Figure 5).[5a,e] In this respect, each of these caviteins possess nativelike characteristics. In contrast, all of the methylene variants bound ANS under identical conditions. Note that single-stranded peptide **8** showed no binding of ANS at both 50 and 200 μM. This demonstrates that the sequence itself may be well designed for specific packing interactions (vide infra).

3. vii. NMR Chemical Shift Dispersion. Typically, considerable chemical shift dispersion is observed in natural proteins and is indicative of nativelike structure.[43] In contrast, signals arising from conformationally mobile groups, such as molten globules, are averaged and result in broad, less disperse signals.[40] Prototype N1/Ar/Me displays ~11 well-dispersed amides (Figure 6), which is indicative of a well-defined amide backbone with a high content of secondary structure, considering that there are only four different types of amino acids present. The presence of just 11 distinct amide peaks also suggests that the four helices of cavitein N1/Ar/Me are in degenerate environments as might be expected from the 4-fold symmetric template.

Relatively sharp aliphatic signals are observed in the spectrum of N1/Ar/Me near 2.2 ppm and are likely due to specific side-chain interactions typical of nativelike states.[43] However, the leucine methyl peaks near 0.8 ppm show little dispersion, which is more characteristic of molten globule-type states.

N1G/Ar/Me displays the most significant chemical shift dispersion of all the cavitein variants. Despite the sequence and helix degeneracies, approximately 20 amide protons are visible; in addition, the symmetry of the cavitand template (at 6 ppm)

(39) Possible aggregation via the hydrophobic cavitand foot is unlikely since several monomeric caviteins possessing methyl feet are reported here (see Discussion).

(40) Ptitsyn, O. B. *Adv. Protein. Chem.* **1995**, *47*, 83−229.
(41) Kuwajima, K. *Proteins* **1989**, *6*, 87−103.
(42) Semisotnov, G.; Rodionova, N. A.; Razjulyaev, O. I.; Uversky, V. N. *Biopolymers* **1991**, *31*, 119−128.
(43) (a) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G., III; Skelton, N. J. *Protein NMR Spectroscopy: Principles and Practice*; Academic Press: San Diego, 1996. (b) Roy, S.; Helmer, K. J.; Hecht, M. H. *Folding Des.* **1997**, *2*, 89−92.

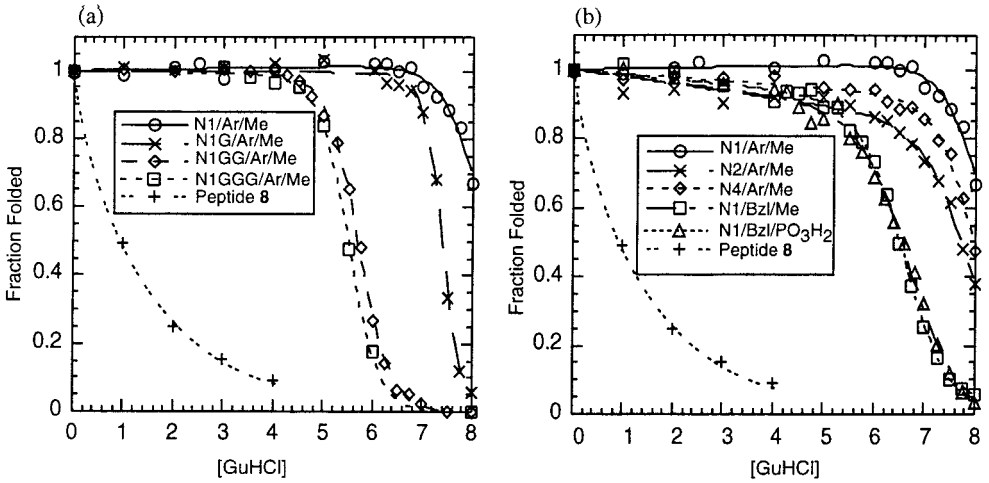


Figure 2. GuHCl denaturation curves for 30 μM solutions of (a) N1/Ar/Me and the Gly cavitein variants N1G/Ar/Me, N1GG/Ar/Me, and N1GGG/Ar/Me, and (b) N1/Ar/Me, the methylene cavitein variants N2/Ar/Me, N4/Ar/Me, and N1/Bzl/Me and the cavitand foot cavitein variant N1/Bzl/PO_3H_2. Each unfolding curve was acquired at pH 7.0 and 25 °C. Peptide **8** is shown in both diagrams for comparison and was at a concentration of 200 μM.

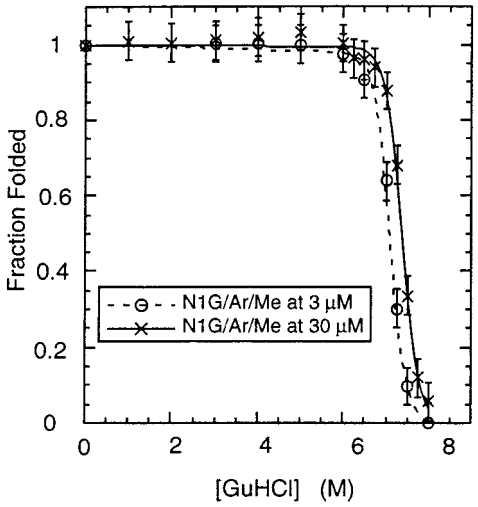


Figure 3. GuHCl denaturation of N1G/Ar/Me at 3 and 30 μM illustrating its concentration dependence. Unfolding was monitored at pH 7.0 and 25 °C.

has been broken.[44] Thus, the four helices are *not* in degenerate environments, which is consistent with an asymmetric dimeric structure for N1G/Ar/Me.[44] N1G/Ar/Me displays significantly more disperse peaks than N1/Ar/Me in the α-carbon region (at ∼4 ppm) and more notably in the leucine-methyl region (at ∼1 ppm). Such dispersion (one leucine-methyl peak is at ∼0 ppm!) is indicative of nativelike packing.

N1GG/Ar/Me displays slightly less dispersion than N1G/Ar/Me, yet more than N1/Ar/Me. The helices in N1GG/Ar/Me appear to be in degenerate environments as only ∼14 amide protons are apparent and the cavitand peak near 6 ppm is a singlet.[44] Thus, the presence of two glycines in N1GG/Ar/Me still appears to allow the helical bundle to adopt a more nativelike structure than N1/Ar/Me.

The ^{1}H NMR spectrum of cavitein N1GGG/Ar/Me manifests little chemical shift dispersion. Most notable is the poor dispersion of the leucine-methyl peaks in the hydrophobic core, which is similar to that of N1/Ar/Me. Despite the similar stabilities for N1GG/Ar/Me and N1GGG/Ar/Me (Table 3), the former is considerably more nativelike; thus, high structural stability does not necessarily indicate nativelike structure.

Table 4. Calculated Values for $\Delta G°_{H_2O}$ for the Cavitein Variants

cavitein	cavitein concn (μM)	$[GuHCl_{1/2}$ (M)[a]	m kcal mol^{-1} M^{-1}	$\Delta G°_{H_2O}$ kcal/mol
N1/Ar/Me	30	>8.0	*b*	*b*
N1G/Ar/Me	3	7.1 ± 0.1	−3.2 ± 0.3[c]	−22.6 ± 2.3[c] (∼40)[d]
N1G/Ar/Me	30	7.4 ± 0.4	−3.1 ± 0.3[c]	−22.9 ± 1.9[c] (∼40)[d]
N1GG/Ar/Me	30	5.7 ± 0.1	−1.7 ± 0.1	−9.9 ± 0.3
N1GGG/Ar/Me	30	5.5 ± 0.1	−1.9 ± 0.1	−10.8 ± 0.5
N2/Ar/Me	30	7.7 ± 0.1	*b*	*b*
N4/Ar/Me	30	8.0 ± 0.1	*b*	*b*
N1/Bzl/Me	30	6.4 ± 0.1	−1.3 ± 0.1[c]	−8.6 ± 0.8[c]
N1/Bzl/PO_3H_2	30	6.5 ± 0.1	−1.2 ± 0.1[c]	−8.0 ± 0.5[c]

[a] $[GuHCl]_{1/2}$ is the concentration of GuHCl required to unfold half of the helical structure. [b] A full unfolding transition was not observed, and therefore these data are not reported. [c] The post-transitional baseline was not observed and was therefore estimated. This may introduce additional errors to those listed. [d] This value was calculated taking into account the dimerization of N1G/Ar/Me (see Results and Experimental Section).

Both the amide and aliphatic regions of N1/Bzl/Me are broad compared to those of the N1/Ar/Me and the Gly cavitein variants. This result is consistent with a fluctuating or dynamic molten globule-like tertiary structure.[45]

The NMR spectrum of single-stranded peptide **8** retains much of the chemical shift dispersion observed for N1/Ar/Me, and is sharper and more disperse than that of N1/Bzl/Me. This suggests a highly specific aggregate structure for peptide **8** (vide infra).[30]

viii. Hydrogen/Deuterium Exchange. Another method for probing the dynamics of protein structures is by monitoring the rate of backbone amide hydrogen/deuterium exchange.[46] Typically, amide hydrogens in nativelike structures exchange far slower than amide hydrogens in more dynamic, partially folded or molten globule structures.[47] We performed H/D exchange

(44) In the NMR spectrum of N1G/Ar/Me, the cavitand peak near 6.0 ppm is split into two peaks, which suggests asymmetry in the cavitand template, likely a result of the observed dimerization.

(45) The effect of dimerization of N1/Bzl/Me cannot be ruled out completely as a possible cause of the broadening; however, the NMR spectra of both N1/Ar/Me (monomer−dimer mixture) and N1G/Ar/Me (dimer) possess significantly sharper and more disperse peaks.

(46) (a) Englander, S. W.; Downer, N. W.; Teitelbaum, H. *Annu. Rev. Biochem.* **1972**, *41*, 903−924. (b) Englander, S. W.; Kallenbach, N. *Q. Rev. Biophys.* **1984**, *16*, 521−655. (c) Raschke, T. M.; Marqusee, S. *Curr. Opin. Biotechnol.* **1998**, *9*, 80−86.

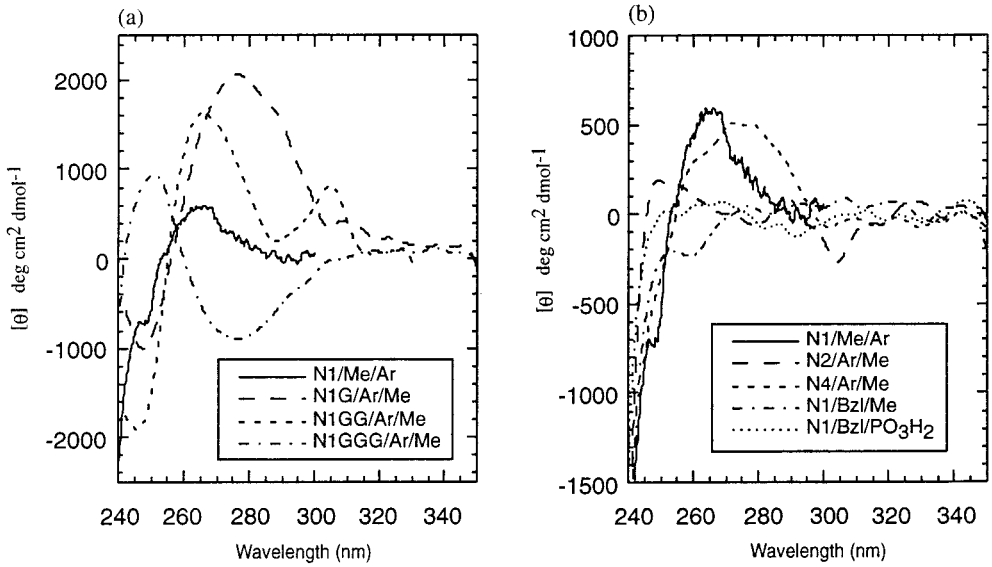


Figure 4. Near-UV CD spectra for (a) N1/Ar/Me and the Gly cavitein variants N1G/Ar/Me, N1GG/Ar/Me, and N1GGG/Ar/Me and (b) N1/Ar/Me, the methylene cavitein variants N2/Ar/Me, N4/Ar/Me, and N1/Bzl/Me and the cavitand foot cavitein variant N1/Bzl/PO_3H_2. Each spectrum was acquired at 25 °C in 50 mM phosphate buffer at pH 7.0 using a cavitein concentration of 30 μM. Note that the magnitude of the *y*-axis in (b) has been adjusted to better illustrate the spectra.

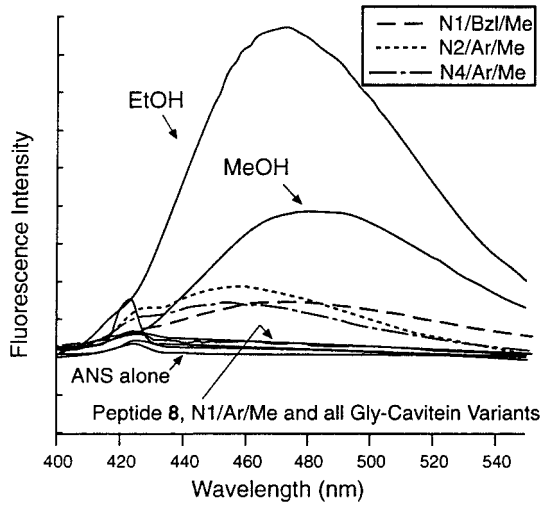


Figure 5. Fluorescence emmission spectra of 1 μM ANS in the presence of 95% ethanol, 100% methanol, 50 μM solutions of caviteins N1/Ar/Me, N1G/Ar/Me, N1/GG/Ar/Me, N1GGG/Ar/Me, N2/Ar/Me, N4/Ar/Me, N1/Bzl/Me, peptide **8** at 50 and 200 μM (indistinguishable), and ANS alone at pH 7.0, 25 °C in 50 mM phosphate buffer. Note that, apart from ethanol and methanol, only caviteins N2/Ar/Me, N4/Ar/Me and N1/Bzl/Me show any detectable emission signals above the baseline. The emission spectra of the other caviteins and peptide **8** were essentially indistinguishable from the baseline.

experiments on prototype N1/Ar/Me[26] and the most stable and nativelike cavitein, N1G/Ar/Me, at pD[48] 5.0 and 25 °C. For N1/Ar/Me, many amides were visible for up to 24 h and yield protection factors[46a,49] of about 4×10^3. These protection factors correspond to some of the more stable amide protons found in molten globule structures or to amides that are intermediary between a molten globule state and a nativelike state.[47] The longest lived amide proton (~8.4 ppm) was observed for several days under these conditions and yields a protection factor of $(2 \pm 0.5) \times 10^4$. This proton exhibits significant protection from exchange and its protection factor is consistent with an amide intermediary between a molten globule and nativelike state.[47]

H/D exchange in N1G/Ar/Me is substantially slower (Figure 7) than for N1/Ar/Me. At least three amide peaks (8.3, 8.4, and 8.5 ppm), are visible even after three weeks. Protection factors for these amides[26] range from $(8 \pm 1) \times 10^4$ to $(8 \pm 2) \times 10^5$. Although there are large errors associated with these protection factors, these amide protons, nevertheless, exhibit *substantial* protection from exchange and these protection factors are typical of nativelike structures.[47]

Discussion

An essential goal in de novo protein design is to create well-defined, nativelike structures, as opposed to fluxional, molten globules.[50] Although we still have a long way to go before we can confidently produce such structures, we have met with significant success here. We now provide our take on the features that govern successful design of nativelike structure in our system.

Stability. All eight caviteins are highly helical, extremely stable to chemical denaturants, and some possess nativelike features. All are far more stable than nontemplated peptide **8**. Thus, "template assembly" plays an enormous role in creating stable helical bundles. The linkers also affect stability, as shown with N1/Ar/Me and N1G/Ar/Me versus N1GG/Ar/Me and N1GGG/Ar/Me (Figure 2a). Also notable are the relative stabilities of N1/Bzl/Me and N2/Ar/Me (Figure 2b), despite being constitutional isomers: Computational analysis (MM2)[51] using cavitand-*S*-methyl derivatives reveals that the most stable

(47) (a) Hughson, F. M.; Wright, P. E.; Baldwin, R. L. *Science* **1990**, *249*, 1544−1548. (b) Jeng, M.-F.; Englander, S. W.; Elove, G. A.; Wand, A. J.; Roder, H. *Biochemistry* **1990**, *29*, 10433−10437. (c) Chamberlain, A. K.; Marqusee, S. *Biochemistry* **1998**, *37*, 1736−1742.

(48) pH was corrected for isotope effects by using the equation: pD = pH (apparent) +0.4. See: Glasoe, P. K.; Long, F. A. *J. Phys. Chem.* **1960**, *64*, 188−190.

(49) Assignments of individual amide protons was not performed, and therefore the listed protection factors are not sequence-corrected as described in: Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W. *Proteins* **1993**, *17*, 75−86. Instead, a context-independent equation was used as described in ref 46a.

(50) Note that in natural proteins molten globules are often meta-stable intermediates on the path toward native proteins. In de novo proteins, molten globules are often the most stable state, as nativelike structures are rarely present at all on the folding-energy landscape.

(51) MM2 calculations were performed with the computer program *CS Chem3D Pro*, version 3.2; CambridgeSoft Corporation: Cambridge, MA, 1995.

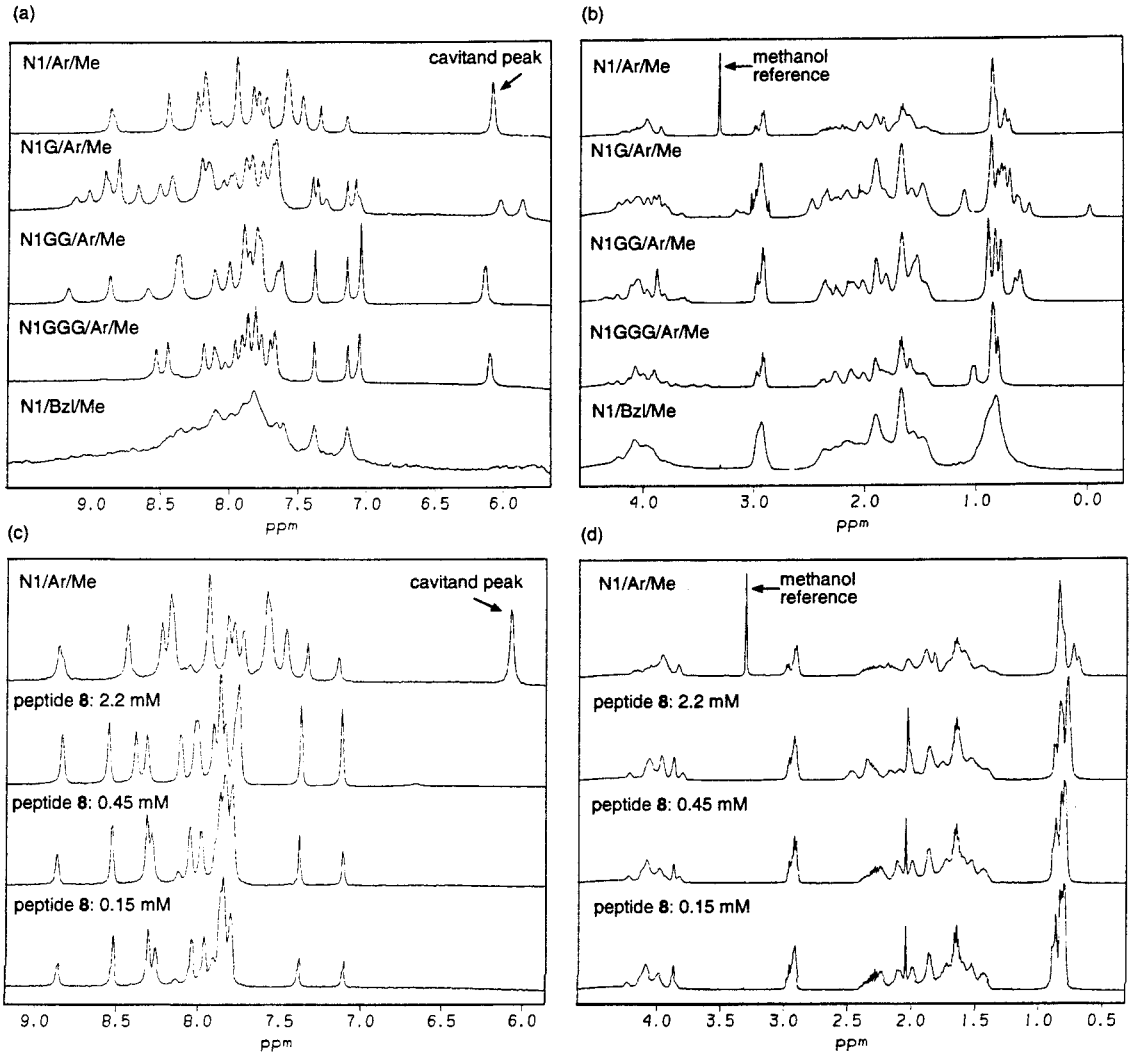


Figure 6. Amide (a) and aliphatic (b) expansions of 500 MHz ^{1}H NMR spectra of N1/Ar/Me, N1G/Ar/Me, N1GG/Ar/Me, N1GGG/Ar/Me, N1/Bzl/Me (each at 0.2−0.3 mM). Parts (c) and (d) compare the spectrum of N1/Ar/Me to that of peptide **8** at different concentrations. All spectra were recorded at 25 °C, in phosphate buffer at pH 7.0 in the presence of 10% D_2O.

conformation (by ∼5 kcal/mol) directs the sulfur lone pairs away from the fixed lone pairs of the cavitand "bridge oxygens" in order to minimize dipole−dipole repulsions (Figure 8). This conformational restriction effectively relegates the peptide strands to different environments for the two caviteins and thus provides an explanation for their different unfolding curves, and indicates that perturbations as subtle as conformational alteration can be significant to overall stability.

Nativelike Structure. The hydrophobic core of the caviteins contain only leucine residues—a feature that has led to molten globule-like structures in other systems.[4] Surprisingly, even template-free peptide **8** manifests nativelike properties in its lack of ANS binding and its relatively well-dispersed NMR spectrum. However, comparison of peptide **8** with the caviteins should be done with caution: peptide **8** may fold into parallel or antiparallel/multimeric topologies,[30] which are precluded for the caviteins. Comparison between caviteins is more fruitful, as they are composed of the same basis sequence and four parallel helices. Linker effects yield both nativelike and molten globule-like features. Control over such features would be invaluable, and may be realized. The Gly linker variants are more nativelike than the methylene linker variants. Noteworthy are N1G/Ar/Me and N4/Ar/Me, which have different ANS-binding properties, vastly different NMR dispersion, and different near-UV CD spectra. These caviteins are constitutionally similar, as they possess the same number of atoms linking the cavitand template to the basis sequence (Figure 9). The peptide bond in the linker of N1G/Ar/Me adds an element of rigidity due to the planarity of the peptide bond. The added peptide bond also provides an additional hydrogen bond acceptor (CO), which may act as a helix-stabilizing "N-cap."[17] The extra peptide bond also possesses an NH group that has been shown to hydrogen-bond to the bridge-oxygen of the cavitand and the thioether linkage when placed in that position with respect to the cavitand.[16] The combination of these effects increases the structural rigidity and packing specificity for N1G/Ar/Me, thus demonstrating how judicious choice of the linker group can impart nativelike structure to the helical bundle.

Oligomerization. One "problem" that appeared while studying these caviteins was their different oligomeric properties. In all cases but one (N1G/Ar/Me), the oligomerization was found to have no effect on the stability of the bundle. It is likely that very short linkers (e.g., N1/Ar/Me, N1G/Ar/Me, and N1/Bzl/Me) distort the helices by clamping them down too tightly upon the rigid cavitand template. Long linkers (e.g., N1GGG/Ar/Me) avoid aggregation but compromise nativelike structure. Linkers with no hydrogen-bonding capability also sacrifice nativelike structure. An ideal linker may be N1GG/Ar/Me, which yields a stable, monomeric, parallel four-helix bundle and retains significant of nativelike structure.

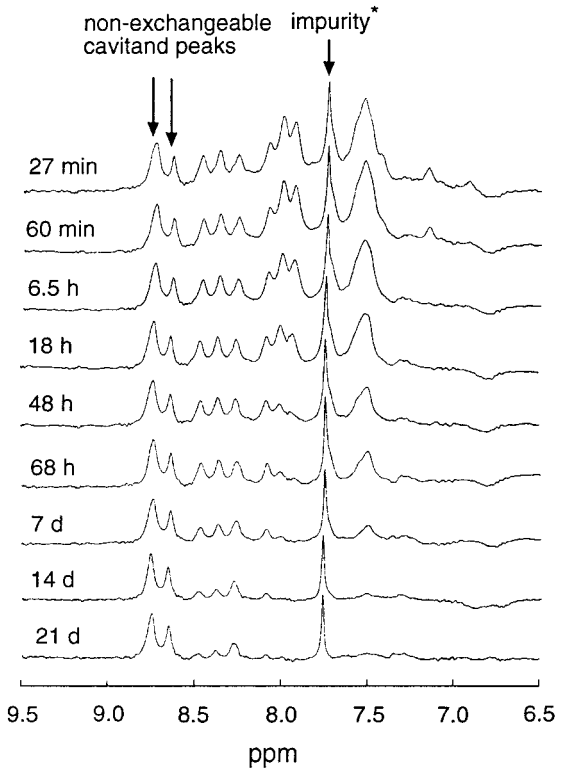


Figure 7. Stack plot of 500 MHz ^{1}H NMR spectra illustrating the time-dependent amide H/D exchange of N1G/Ar/Me in 50 mM pD CD_3COOD/NaOD buffer at pD 5.0 and 25 °C. The impurity (∗), likely residual $CHCl_3$ from a microsyringe, was not present when this experiment was first attempted on a 400 MHz spectrometer. The H/D exchange behavior, during this prior 400 MHz experiment, was very similar to that of the above experiment. Thus, the impurity likely does not affect H/D exchange in the above 500 MHz experiment.

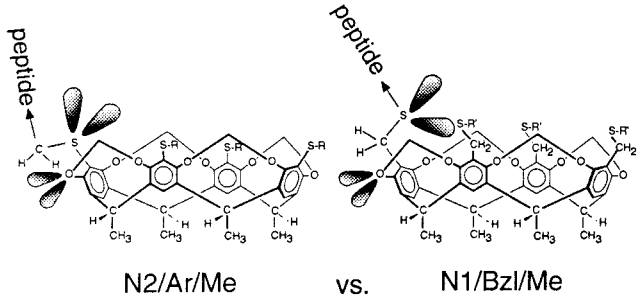


Figure 8. Illustration of the likely preferred linker conformations near the sulfur atoms of N2/Ar/Me and N1/Bzl/Me. In each case, the sulfur is oriented such that its lone pairs are directed away from the lone pairs of the oxygen near the cavitand. In *S*-methyl derivatives of each cavitand, this conformation was predicted by MM2 calculations to be ∼5 kcal/mol more stable by minimizing dipole−dipole repulsions between the lone pairs. A possible consequence of this conformational restriction is that the peptides are likely arranged differently with respect to the cavitand, giving rise to the different unfolding curves.

The structure of N1G/Ar/Me is perhaps the most exciting result. This cavitein was found to be a stable helical dimer with considerable nativelike structure. However, its structure is not the result of two independently associated four-helix bundles (as is the case for N1/Ar/Me and N1/Bzl/Me), but rather a dimer with a somewhat integrated helical structure. The asymmetry of the cavitand peaks in the NMR spectrum (1:1 near 6.0 ppm, Figure 6a; and 3:1 near 8.7 ppm, Figure 7) suggests that perhaps a single helix from one subunit is associating with three helices of the other subunit (and vice versa). Such a "helix-swapped" topology would retain the four-helix bundle topology as promoted by the peptide design. Higher resolution structural information is required to identify the precise mode and cause of dimerization, and such studies are underway. Thus, this cavitein represents an ideal candidate for exploring quaternary interactions between helical bundles.

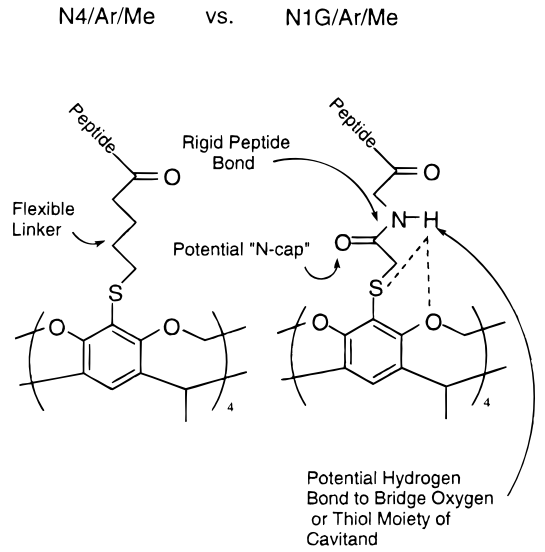


Figure 9. Illustration of the different cavitand−peptide linkers in N4/Ar/Me and N1G/Ar/Me. In each case, "peptide" refers to the identical peptide sequence NH-[EELLKKLEELLKKG]-$CONH_2$.

We also altered the cavitand foot from a hydrophobic methyl group (N1/Bzl/Me) to a propyl-phosphate group (N1/Bzl/PO_3H_2). From their near-identical unfolding curves (Figure 2), we conclude that the cavitand foot is not affecting the stability of the four-helix bundle. However, the addition of the propyl-phosphates eliminates the dimerization (Table 3). This result is significant and may be interpreted as follows:[52] Self-association is hampered by the introduction of the charged phosphate feet, which serve to repel one another. Introduction of an overall charge of at least −4 (depending on the pK_a of the second ionizable phosphate proton, ∼7) to the cavitein via the phosphate feet would provide significant electrostatic repulsion. The dimerization is likely driven by distortion of the helices due to the short linkers, which may be conformationally unsuited for ideal bundling (Figure 8). Regardless, electrostatic repulsion appears to supersede such deleterious effects.

Conclusions

We have designed, synthesized, and studied eight cavitand-based de novo parallel four-helix bundle proteins. Each possesses the same 14-amino acid basis sequence, but a different cavitand−peptide linkage or cavitand foot. The cavitand is instrumental in imparting stability to the proteins, presumably via a severe savings in entropy of organization. Despite the extreme stabilities of all the cavitein variants, not all possess nativelike structure, which demonstrates that very high stability is not necessarily accompanied by structural specificity. The cavitand−peptide linkage has a dramatic effect on the stabilities and structures of the bundles. All of the Gly cavitein variants

(52) Two less plausible explanations for elimination of dimerization by incoproration of phosphates are: (1) the caviteins are self-associating via the hydrophobic methyl feet of the cavitand. This is unlikely since several monomeric caviteins presented herein possess methyl feet. (2) The helices are interacting significantly with the cavitand foot such that altering the foot can alter the overall structure and result in monomers or dimers. This cannot be ruled out entirely; however, it is unlikely that the peptides could be in such dissimilar conformations and possess the observed near-identical GuHCl denaturation curves and stabilities.

possess some of the properties of nativelike structure, while the methylene linker variants were more characteristic of molten globules. The additional hydrogen bonding capabilities of the Gly linkers likely aids in inducing nativelike structure. Changing the cavitand feet was found to have little effect on the stability of the bundle, however, the introduction of an overall charge via the cavitand feet likely inhibits oligomerization via electrostatic repulsion. Short linkers may distort the helices and lead to oligomerization, while ideal linkers (e.g., N1GG/Ar/Me) promote nativelike structure without inducing oligomerization. The highly stable, nativelike N1G/Ar/Me appears to be an exciting candidate for probing quaternary structure, which could lead to well designed multicomponent de novo proteins.

Experimental Section

Peptide Synthesis. All peptides were synthesized using standard solid-phase methods and the Fmoc/*t*-Bu strategy on a Applied Biosystems 431A peptide synthesizer and Rink's amide resin to afford C-terminal amides.[53] All chemical were purchased from Advanced Chemtech (Louisville, KY) except for DMF and the activating reagents which were purchased from Aldrich. After activation with the appropriate linker (see below), each peptide resin was washed with CH_2Cl_2, cleaved from the resin using a 2 h treatment of 95% TFA, and purified by reversed-phase HPLC using a linear gradient of acetonitrile (0.05% TFA) and water (0.1% TFA) on a Phenomenex Selectosil C_{18} column (250 × 22.5 mm) at a flow rate of 10 mL/min. Each peptide eluted at ~50% acetonitrile. The purity of each peptide was assessed by the observation of a single peak by analytical reversed-phase HPLC (>95% pure) and their identities confirmed with mass spectrometry ($LSIMS^+$, thioglycerol matrix). The reported yields are only approximate due to adsorbed water present after lyophilization and variable loading of the peptide resin.

The synthesis of peptide **2** has been described previously.[14,21]

Peptide 3. The free N terminus of the 15-amino acid peptide resin (300 mg resin, ~150 mg peptide, ~0.089 mol) was reacted with chloroacetyl chloride (42 μL, 0.53 mmol) and DIEA (92 μL, 0.53 mmol) in dry DMF (5 mL) for 1.5 h at room temperature. Subsequent cleavage from the resin and purification by HPLC afforded peptide **3** as a white solid (35 mg, ~25%). MS: *m*/*z* 1802 ($(M + H)^+$, 100).

Peptide 4. The free N terminus of the 16-amino acid peptide resin (300 mg resin, ~15Peptide 3.0 mg peptide, ~0.089 mmol) was reacted with chloroacetyl chloride (42 μL, 0.53 mmol) and DIEA (92 μL, 0.53 mmol) in dry DMF (5 mL) for 1.5 h at room temperature. Subsequent cleavage from the resin and purification by HPLC afforded peptide **4** as a white solid (42 mg, ~25%): MS *m*/*z* 1859 ($(M + H)^+$, 100).

Peptide 5. The free N terminus of the 17-amino acid peptide resin (300 mg resin, ~150 mg peptide, ~0.089 mmol) was reacted with chloroacetyl chloride (42 μL, 0.53 mmol) and DIEA (92 μL, 0.53 mmol) in dry DMF (5 mL) for 1.5 h at room temperature. Subsequent cleavage from the resin and purification by HPLC afforded peptide **5** as a white solid (45 mg, ~25%): MS *m*/*z* 1916 ($(M + H)^+$, 100).

Peptide 6. The free N terminus of the 14-amino acid peptide resin (450 mg, ~150 mg peptide, 0.094 mmol) was treated with 3-bromopropionyl chloride (47 μL, 0.47 mmol), DIEA (82 μL, 0.47 mmol) and NaBr (97 mg, 0.94 mmol) in dry DMF (5 mL) for 45 min at room temperature. Subsequent cleavage from the resin and purification by HPLC afforded peptide **6** as a white solid (36 mg, ~20%): MS *m*/*z* 1805 ($(M + H)^+$, 100).

Peptide 7. The free N terminus of the 14-amino acid peptide resin (350 mg, ~120 mg peptide, 0.071 mmol) was treated with 5-bromovaleryl chloride (47 μL, 0.36 mmol) and DIEA (82 μL, 0.47 mmol) in dry DMF (5 mL) for 1 h at room temperature. Subsequent cleavage from the resin and purification by HPLC afforded peptide **7** as a white solid (49 mg, ~40%): MS *m*/*z* 1805 ($(M + H)^+$, 100).

Peptide 8. The free N terminus of the 14-amino acid peptide resin (350 mg, ~120 mg peptide, 0.071 mmol) was treated with 10% acetic anhydride in NMP for 1 h at room temperature. Subsequent cleavage from the resin and purification by HPLC afforded peptide **8** as a white solid (60 mg, ~50%): MS *m*/*z* 1711 ($(M + H)^+$, 100).

Cavitand Synthesis. The synthesis of cavitands **1** and **11** have been described previously.[16,21] The synthesis of cavitand **10** has also been described previously;[22] however, we used the more direct direct synthetic route described below.

Cavitand 10. Thiourea (18 mg, 0.24 mmol) was added to a solution of benzylbromide **9**[23] (50 mg, 55 μmol) in degassed DMF (5 mL). The reaction was stirred for 2 h at room temperature and ice-cold degassed 2 M NaOH was added. The reaction mixture was then stirred for an additional 1 h and the solvent evaporated in vacuo. Acetic acid (5%) was added until the mixture became acidic, and the precipitate was collected by filtration over a fine frit. The crude product was then dissolved in $CHCl_3$, dried with $MgSO_4$, and precipitated with hexanes to afford cavitand **10**. The product gave a ^1H NMR spectrum identical to that previously reported.[22]

Cavitein Synthesis. Each cavitein was purified by reversed-phase HPLC as described above. Their purities were assessed by the observation of a single peak by reversed-phase HPLC (>95% pure), and their identities were confirmed with electrospray mass spectrometry (reported molecular weights are from deconvoluted raw spectra). The synthesis of $N1/Bzl/PO_3H_2$ has been described previously.[21]

N1/Ar/Me.[14] Procedure A: DIEA (2.6 μL, 15 μmol) was added to a solution of cavitand **1** (1.1 mg, 1.5 μmol) and peptide **2** (11.9 mg, 6.8 μmol) in degassed DMF (2 mL), and the mixture was stirred at 25 °C for 16 h. The crude reaction mixture was evaporated in vacuo and purified to afford N1/AAr/Me as a white solid (7.0 mg, 62%): MS 7556.0 ± 0.9 Da [calcd 7557.1 (average isotope composition)].

N1G/Ar/Me. Procedure A was employed using cavitand **1** (1.4 mg, 1.9 μmol), peptide **3** (12 mg, 6.7 μmol) and DIEA (2.3 μL, 13 μmol) at 30 °C to afford N1G/Ar/Me (6.0 mg, 58%): MS 7787.9 ± 0.3 [calcd 7785.4 (average isotope composition)].

N1GG/Ar/Me. Procedure A was employed using cavitand **1** (1.4 mg, 1.9 μmol), peptide **4** (22 mg, 12 μmol) and DIEA (3.3 μL, 19 μmol) at 30 °C to afford N1GG/Ar/Me (10 mg, 65%): MS 8015.1 ± 0.8 [calcd 8013.6 (average isotope composition)].

N1GGG/Ar/Me. Procedure A was employed using cavitand **1** (2.8 mg, 3.9 μmol), peptide **5** (33 mg, 17 μmol) and DIEA (6.8 μL, 39 μmol) at 30 °C to afford N1GGG/Ar/Me (6.8 mg, 21%): MS 8242.8 ± 0.9 [calcd 8241.8 (average isotope composition)].

N2/Ar/Me. Procedure A was employed using cavitand **1** (0.9 mg, 1.3 μmol), peptide **6** (10 mg, 5.5 μmol) and DIEA (2.2 μL, 13 μmol) to afford N2Ar/Me (0.9 mg, 9%): MS 7615.1 ± 1.0 [calcd 7613.3 (average isotope composition)].

N4/Ar/Me. Procedure A was employed using cavitand **1** (1.0 mg, 1.3 μmol), peptide **7** (11 mg, 5.8 μmol) and DIEA (2.3 μL, 13 μmol) to afford N4Ar/Me (1.2 mg, 12%): MS 7727.9 ± 0.3 [calcd 7725.5 (average isotope composition)].

N1/Bzl/Me. Procedure A was employed using cavitand **10** (1.0 mg, 1.4 μmol), peptide **2** (14 mg, 8.1 μmol) and DIEA (2.4 μL, 14 μmol) to afford N2Ar/Me (0.9 mg, 9%): MS 7615.1 ± 1.0 [calcd 7613.3 (average isotope composition)].

Circular Dichroism. All CD spectra were recorded on a JASCO J-710 spectropolarimeter except for the temperature dependence study of N1/Ar/Me which was recorded on a JASCO J-720 equipped with a computer-controlled water bath. Quartz cells of either 1 mm or 1 cm path length were used in all scans. Each spectrum was an average of three scans with the solvent baseline recorded separately and subtracted. All spectra were recorded at 25 °C and all data points were performed in triplicate unless otherwise noted. Calibration of the instrument was performed routinely with *d*-10-(+)-camphorsulfonic acid. Error bars represent an average of the standard deviations for each point and determined to be approximately ±5%. Raw spectra were normalized to mean residue ellipticity, $[\theta]$, using the equation

$$[\theta] = \theta_{obs}/10lcn$$

where θ_{obs} is the ellipticity measured in millidegrees, l in the path length of the cell in centimeters, c is the peptide concentration in mol/L, and n is the number of residues in the de novo protein. Percent helicity was calculated using the equation described in ref 24.

(53) Rink, H. *Tetrahedron Lett.* **1987**, *28*, 3787−3788.

Guanidine hydrochloride denaturation experiments were performed using an 8.0 M solution of GuHCl buffered at pH 7.0 with 50 mM phosphate. Each data point was performed in duplicate and acquired with a separate sample of cavitein in pH 7.0, 50 mM phosphate buffer. To attain the high concentrations of GuHCl, cavitein **3** was dissolved directly in the 8 M GuHCl stock solution, and diluted accordingly to give the appropriate concentrations. Samples were equilibrated for 5 min only since it was found that any effect of GuHCl was immediate as the mean residue ellipticity did not change from 5 min to 24 h at 25 °C. The urea experiment was performed in a similar manner.

Peptide and cavitein concentrations were determined by amino acid analysis and thus have potential errors of up to 10%.

Unfolding was analyzed using the linear extrapolation method of Santoro and Bolen.[36] This method assumes that unfolding is a reversible, two-state process and that the free energy of folding is a linear function of the GuHCl concentration. Accordingly, the data were fit to the following equation using nonlinear least-squares methods to fit the pretransitional baseline and, due to the extreme stability of the caviteins, approximates the post-transitional baseline:[54]

$$\theta_{obs} = \theta_N(f_n)(1 - a[\text{GuHCl}]) + \theta_U(1 - f_n)$$

where θ_{obs} is the ellipticity at 222 nm at a given concentration of GuHCl, θ_N is the ellipticity of the folded state in the absence of GuHCl, θ_U is the ellipticity of the unfolded state, a is a constant, and f_n is the fraction of the cavitein in the folded state:

$$f_n = \exp\{-((\Delta G^\circ_{H_2O} - m[\text{GuHCl}])/RT)\}/[1 + \exp\{-((\Delta G^\circ_{H_2O} - m[\text{GuHCl}])/RT)\}]$$

where $\Delta G^\circ_{H_2O}$ is the free energy of unfolding in the absence of denaturant, m represents the change in free energy with respect to the concentration of GuHCl, R is the universal gas constant, and T is the temperature. The values of $\Delta G^\circ_{H_2O}$, m, and a were determined by nonlinear least-squares analysis by the program KaleidaGraph version 3.08d (Synergy Software). The value of θ_N was normalized in all cases to one. The errors reported were calculated by this analysis.

In the case of N1G/Ar/Me, we observed a concentration dependent denaturation curve suggesting a possible dimer (folded) to monomer (unfolded) transition. Therefore, in addition to the above analysis, we analyzed the data according to[55]

$$N_2 \rightleftharpoons 2U$$

$$K_u = [U]^2/[N_2] = 2P_t(f_u^2/(1 - f_u)$$

where $[N_2]$ and $[U]$ are the concentrations of the native dimer and the unfolded monomer, respectively, P_t is the total cavitein concentration in mol/L, and f_u is the fraction of cavitein unfolded. If we assume (as above) that the free energy of unfolding is a linear function of denaturant concentration, we find that

$$f_u = k^2/2P_t(1 + (1 + 8P_t/k^2)^{1/2})$$

where

$$k = \exp\{-((\Delta G^\circ_{H_2O} - m[\text{GuHCl}])/RT)\}$$

The extrapolated value was calculated using a nonlinear least-squares curve fitting as above. Note that this equation assumes that the monomeric form of N1G/Ar/Me is entirely unstable which may not be valid.

NMR Spectroscopy. 1D-^{1}H NMR spectra used to evaluate chemical shift dispersion were run at 500 MHz on a Varian Unity equipped with shimming hardware. Each sample was run at 25 °C and dissolved in 45 mM phosphate buffer (90:10, H_2O:D_2O) at pH 7.0 using a sweep width of 8000.00 Hz, a data size of 4096 complex points and 128 scans. A relaxation delay of 1.0 s was used, during which time the water signal was saturated using a frequency-selective low-power decoupling pulse. Spectra were processed using Felix v2.30 (Biosym, San Diego, CA). A convolution function was applied to remove the residual water signal. No other line-broadening functions were used. A residual amount of methanol was used as a reference (3.31 ppm). The concentration of each cavitein was ~0.2−0.3 mM; concentrations of peptide **8** were ~0.15 mM, 0.45 mM and 2.2 mM.

The H/D exchange experiment for N1/Ar/Me was run at 400 MHz on a Bruker WH-400 at 25 °C using presaturation of the water signal. For N1G/Ar/Me, the experiment was run on a Bruker AMX-500 running at 500 MHz for protons. Each experiment was initiated by dissolving the cavitein directly into 50 mM acetic acid-d_4 buffer in D_2O at pD 5.0 and transferred quickly to the NMR tube. pD was corrected for isotope effects.[48] The first scan was acquired as soon as possible followed by scans at arbitrary time increments. The peak heights were integrated and normalized with the nonexchangeable proton H_{out} (near 6 ppm) from the cavitand. For N1/Ar/Me, the first-order rate constant for the longest-lived amide proton[26] at ca. 8.4 ppm was calculated to be $(2 \pm 0.5) \times 10^{-2}$ h^{-1} and a half-life of $t_{1/2} = 37 \pm 9$ h. Its protection factor was then calculated to be $(2 \pm 0.5) \times 10^4$ where $t_{1/2}$ for an "unprotected" amide proton at pH 5.0 and 25 °C is 0.11 min.[46a] For N1/Ar/Me, errors represent one standard deviation from four rate constant estimates. Data for the N1G/Ar/Me experiment are reported in Table S2 in Supporting Information.

Fluorescence Spectroscopy. ANS fluorescence measurements were determined using a 1 cm path length where all runs contained 1 μM ANS in 50 mM phosphate buffer pH 7.0 at 25 °C with either 50 μM cavitein, 50 μM peptide **8**, 200 μM peptide **8**, 95% ethanol, or 100% methanol. Excitation was at 370 nm and emission was recorded between 400 and 550 nm.

Analytical Ultracentrifugation. Sedimentation equilibrium measurements were performed on a temperature-controlled Beckman XL-I analytical ultracentrifuge equipped with an An60Ti rotor and photoelectric scanner. A double sector cell, equipped with a 12 mm Epon centerpiece and sapphire windows was used. Data were collected at 25 °C at a rotor speed of 30 000 rpm until equilibrium was established across the cell. Samples were allowed to equilibrate for 24−32 h, and duplicate scans 3 h apart were overlaid to determine that equilibrium had been reached. Scans were acquired by UV detection at different wavelengths for each cavitein.[26] The partial specific volumes of the each cavitein were calculated based on their amino acid composition only.[26,31] The solution densities were calculated based on the buffer composition.[26,31] The data were analyzed by a nonlinear least-squares analysis using the Origin software provided by Beckman. The accuracy of the fit was evaluated on the basis of randomness and magnitude of the residuals, expressed as the difference between the theoretical curve and the experimental data. The fit was also checked for physical reasonability. All data fit to a single species model except for N1/Ar/Me (in buffering conditions) which was found to exist in a monomer−dimer equilibrium (Table 3). The association constant for dimerization was not calculated due to the errors involved in its calculation (see note 34).

Acknowledgment. We would like to thank the Natural Sciences and Engineering Research Council of Canada for financial support in the form of a grant to J.C.S. and a postgraduate fellowship to A.R.M. Acknowledgment is also made to the donors of the Petroleum Research Fund, administered by the ACS, for partial support of this research. Additional thanks to E. Koepf, M. Petrowski and J. Kelly for acquiring the sedimentation equilibrium data, F. Rosell and G. Mauk for use of their CD temperature apparatus, L. Burtnick for use of his fluorimeter, M. Fryzuk for use of his HPLC, and G. Connolly and L. McIntosh for acquiring the 500 MHz 1D-NMR dispersion spectra.

(54) Regan, L.; Rockwell, A.; Wasserman, Z.; DeGrado, W. F. *Protein Sci.* **1994**, *3*, 2419−2427.

(55) (a) Neet, K. E.; Timm, D. E. *Protein Sci.* **1994**, *3*, 2167−2174. (b) Mok, Y.-K.; Gay, G. D.; Butler, P. J.; Bycroft, M. *Protein Sci.* **1996**, *5*, 310−319. (c) De Francesco, R. D.; Pastore, A.; Vecchio, G.; Cortese, R. *Biochemistry* **1991**, *30*, 143−147. (d) Bowie, J. U.; Sauer, R. T. *Biochemistry* **1989**, *28*, 7140−7143.

Supporting Information Available: Experimental details of each sedimentation equilibrium experiment, plots of absorbance vs radius, the residuals for each plot, and plots of $[\theta]_{222}$ for N1/Ar/Me as a function of concentration (including peptide **8**), pH, temperature, TFE, urea concentration. In addition, a stack plot of the H/D exchange experiment for N1/Ar/Me and a table summarizing the half-lives, rate constants and protection factors for three amide protons in the H/D exchange experiment of N1G/Ar/Me are also included (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA990487F