Design and Synthesis of a Protein Catenane**

Liang Z. Yan and Philip E. Dawson*

The structural characterization of thousands of proteins has revealed over 550 different backbone topologies.[1] Interestingly, none of these proteins form a backbone structure that can be described as topologically linked. Examples of knotted and catenaned structures have been observed within a single peptide chain through disulfide bonds and cofactors,[2] a handful of linear "knotted" proteins have been described.[3] Recently, the bacteriophage HK97 capsid[4, 5] was found to be composed of a network of interlocking protein subunits connected through a side chain isopeptide bond. Here, we describe the design and synthesis of the first protein with topologically linked backbone structure, a protein [2]catenane.

In contrast to proteins, interlocked macromolecules represent an intensely studied area of supramolecular chemistry,[6] and natural and synthetic DNA molecules form both knotted and catenane structures.[7] Catenanes have been synthesized by both statistical and template-directed strategies.[6] Since proteins have evolved to fold into defined three-dimensional structures, a folding-directed strategy was devised. The tetramerization domain of the tumor suppressor protein p53 (residues 325–356) has been structurally characterized by crystallographic and NMR methods.[8, 9] The structure can be described as a "dimer of dimers" in which the repeated dimer unit consists of two antiparallel β-strands and two antiparallel α-helices (Figure 1). These structures are formed by the folding of the peptides in an intertwined manner that has been described as a bisecting U motif.[10] It was reasoned that if the N- and C-termini of each folded peptide could be joined to form a cyclic peptide, the polypeptides would become interlocked, creating a catenane structure.

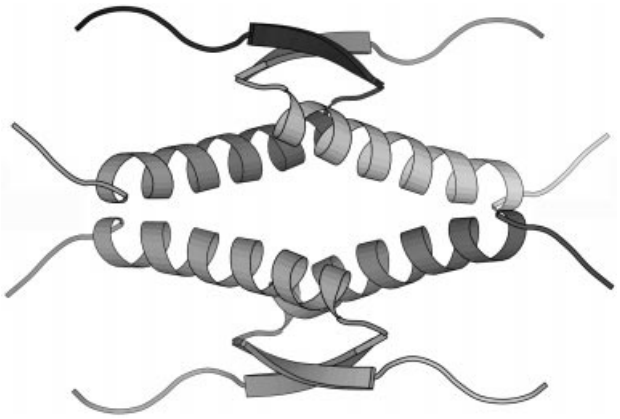
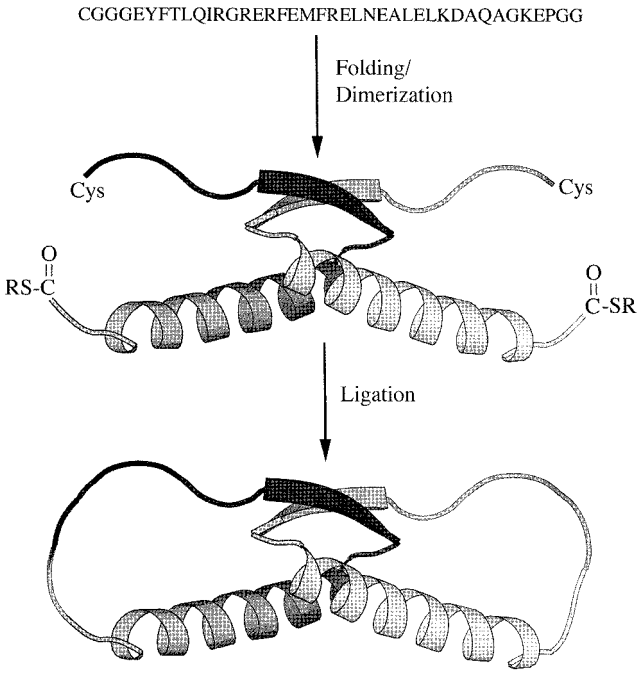


Figure 1. Structure (Molscript plot[24]) of the tetramerization domain of the tumor suppressor protein p53. The backbone is represented as a ribbon to highlight the interlocking structure of the dimers.

To form a peptide bond between the N- and C-termini of the p53 peptides following folding, a highly chemoselective reaction is necessary. Native chemical ligation utilizes the rapid and specific reaction of an N-terminal cysteine residue with a C-terminal α-thioester group to form a native amide bond. This reaction is mediated by thioester exchange followed by a rapid S to N acyl shift.[11] If the N-terminal Cys and C-terminal thioester groups are on the same polypeptide, cyclic products (e.g. cyclic peptides and proteins) can be generated with remarkable efficiency.[12, 13]

The synthetic strategy for the assembly of the p53 catenane is described in Scheme 1. Since protein dimerization and folding are rapid[14] compared to the native chemical ligation



Scheme 1. Synthetic design of the p53 catenane.

reaction at pH 6, the p53 peptides should assemble before circularizing. The resulting structure is expected to consist of two p53 catenanes folded together to form a dimer of catenanes. A polypeptide corresponding to residues 325–361 with an N-terminal CGG extension and a C-terminal thioester was synthesized by standard solid-phase protein synthesis using butyloxycarbonyl protecting groups (Boc-SPPS) on a thioester resin. The folding and ligation of the CGG-p53-COSR polypeptide was highly efficient, yielding a single major peak by HPLC after 6 h (Figure 2a).[15]

Since [2]catenanes are held together by a mechanical bond, they are a single molecular compound which cannot dissociate into separate components without breaking at least one covalent bond. Consistent with this structure, the observed molecular weight of 8971.5 Da by ESI-MS corresponded to

[*] Prof. P. E. Dawson, L. Z. Yan
Departments of Cell Biology and Chemistry
The Skaggs Institute of Chemical Biology, The Scripps Research Institute
10550 N. Torrey Pines Road, CVN-6, La Jolla, CA 92037 (USA)
Fax: (+1) 858-784-7319
E-mail: dawson@scripps.edu

[**] This work was supported by The Skaggs Institute for Chemical Biology, The Sloan Foundation, and NIH-GM570132 (PED). We thank Dr. Songpong Deechongkit for assistance with the analytical ultracentrifuge and CD measurements.

Supporting information for this article is available on the WWW under http://www.angewandte.com or from the author.

  1433-7851/01/4019-3625 $ 17.50+.50/0

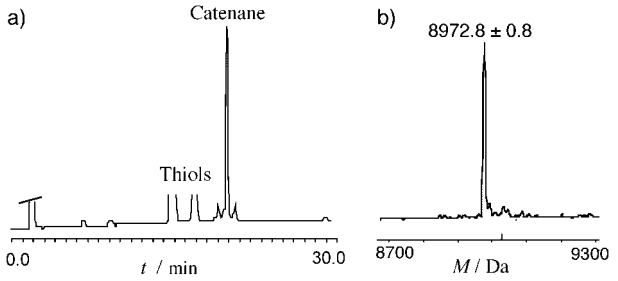


Figure 2. Synthesis of p53 catenane. a) HPLC trace of crude reaction mixuture after 6 h. b) Electrospray mass spectrum of the major peak, deconvoluted to a single mass.

two cyclic p53 peptides (Figure 2b). An alternative structure that would have the same mass is a cyclic dimer. To confirm the catenane structure, a partial cleavage at the single cysteine residue was performed by cyanylation followed by aminolysis.[16] Analysis of the cleavage products by HPLC and ESI-MS detected only linear monomer (4529 Da), cyclic monomer (4487 Da), and cyanylated cyclic monomer (4511 Da) products. The presence of cyclic monomer and absence of linear dimer products is consistent with the interlocked structure.

The folded structure of the p53 catenane was characterized by circular dichroism. The CD spectrum of the p53 catenane (Figure 3) has the same minima as that of the linear form although the mean residual ellipticity is significantly higher

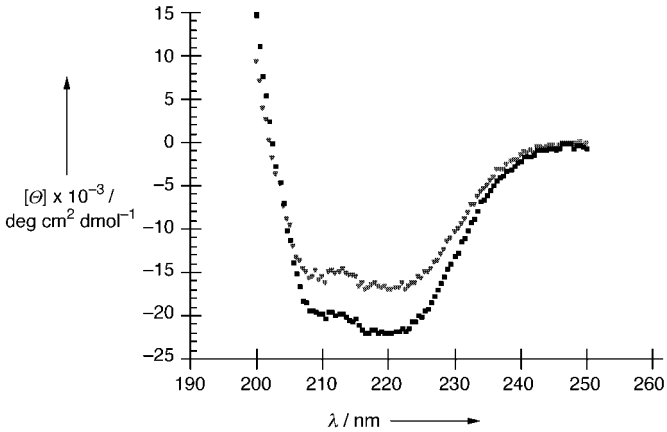


Figure 3. Circular dichroism spectra of linear (▼) and catenaned (■) p53 tetramerization domains.

for the catenane structure. This enhancement could result from structural differences in the CGG-linker or from the region corresponding to the "ends" of the linear p53 peptides.[17] Two p53 catenanes are expected to fold together to form a dimer of catenanes of 17942 Da. Consistent with this folded conformation, a molecular weight of 16250 Da was determined by analytical ultracentrifugation.

Linear proteins can be stabilized by engineered disulfide bonds and by tethering dimeric proteins.[18, 19] The resulting increase in stability has been attributed to reduction in entropy and the resulting destabilization of the unfolded state.[20] By analogy, the p53 catenane was expected to be more stable than the linear p53 tetramer as a result of peptide cyclization and topological linking. As anticipated, the linear p53 tetramer started denaturing at about 60 °C with a T_m of 83 °C, while the p53 catenane was folded up to 98 °C, as monitored by CD at 222 nm. Interestingly, the p53 catenane is significantly stablized even though only one of the dimer interfaces is topologically linked.[21]

The design and synthesis of the p53 protein catenane demonstrates a novel method for assembling and stabilizing proteins. The protein folding-assisted assembly of the linear polypeptide chains followed by backbone circularization using native chemical ligation should be applicable to the entire family of bisecting U proteins using synthetic or biologically expressed C-terminal thioester peptides. In addition, other homo- or heterodimeric proteins may be accessible to topological engineering through the use of appropriate peptide linker sequences. Protein catenanes may find application in many areas of protein science including the study of folding and dynamics and the development of novel protein materials.

Experimental Section

Peptide synthesis and assembly: The polypeptide corresponding to CGGGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEPGG was synthesized manually using established in situ neutralization protocols for Boc-SPPS[22] on a thioester resin.[23] The crude peptide was purified by reversed-phase HPLC, and characterized by ESI-MS: 4704.3 ± 1.0 Da. The purified linear peptide (3.5 mg) was dissolved in a solution (1.5 mL; pH 7.5) containing 0.1 M NaH_2PO_4 and 1 mM EDTA, and benzyl mercaptan (30 μL; 2%) and thiophenol (30 μL; 2%) were added; the final ligation mixture was pH 6. The circularized product was purified by HPLC. (1.6 mg, 47% recovered yield). ESI-MS: 8971.5 ± 1.5 Da.

Cyanylation and aminolysis: The cyclic product (1.6 mg, 0.18 μmol) and 1-cyano-4-(dimethylamino)pyridinium tetrafluoroborate (CDAP) (63.4 μg, 0.27 μmol) were dissolved in a solution (1.2 mL; pH 2.3) containing degassed 6 M guanidine and 0.1 M AcOH for 20 min at 25 °C. After desalting by HPLC, the cyanylated peptides (0.4 mg) were treated with 3.0 M ammonia (0.2 mL) at 0 °C for 20 min. The following species were found by ESI-MS: 8972.0 ± 1.3, 4528.8 ± 1.0, 4512.1 ± 0.3, 4486.6 ± 0.8 Da.

Circular dichroism and thermal denaturation: The CD spectra were recorded on an Aviv spectropolarimeter with a 2-mm pathlength cuvette. Samples were prepared in 0.1 M NaH_2PO_4 (pH 6.4), and concentrations were determined by amino acid analysis. Thermal denaturation was monitored by CD at 222 nm from 2 °C to 98 °C in increments of 2 °C with 2 min equilibration.

Analytical ultracentrifugation: Experiments were carried out in a Beckman XL-A analytical ultracentrifuge at 20 °C with a rotor speed of 25000 rpm and detection at 254 nm. Samples were prepared in 0.1 M NaH_2PO_4 (pH 6.4). The data were edited by using XL-A software.

Received: April 19, 2001
Revised: July 4, 2001 [Z16965]

[1] A. G. Murzin, S. E. Brenner, T. Hubbard, C. Chothia, *J. Mol. Biol.* **1995**, *247*, 536–540. (1.53 release of SCOP database)
[2] C. Liang, K. Mislow, *J. Am. Chem. Soc.* **1995**, *117*, 4201–4213.
[3] a) M. L. Connolly, I. D. Kuntz, G. M. Crippen, *Biopolymers* **1980**, *19*, 1167–1182. b) M. L. Mansfield, *Nat. Struct. Biol.* **1997**, *4*, 166–167; c) W. R. Taylor, *Nature* **2000**, *406*, 916–919.
[4] R. L. Duda, *Cell* **1998**, *94*, 55–60.
[5] W. R. Wikoff, L. Liljas, R. L. Duda, H. Tsuruta, R. W. Hendrix, J. E. Johnson, *Science* **2000**, *289*, 2129–2133.
[6] a) G. Schill, *Catenanes, Rotaxanes and Knots*, Academic Press, New York, **1971**; b) F. M. Raymo, J. F. Stoddart, *Chem. Rev.* **1999** *99*, 1643–1663; c) *Catenanes, Rotaxanes and Knots* (Eds.: J. P. Sauvage, C. Dietrich-Buchecker), Wiley-VCH, Weinheim, **1999**.
[7] N. C. Seeman, *Angew. Chem.* **1998**, *110*, 3408–3428; *Angew. Chem. Int. Ed.* **1998**, *37*, 3220–3238.

  1433-7851/01/4019-3626 $ 17.50+.50/0

[8] P. D. Jeffrey, S. Gorina, N. P. Pavletich, *Science* **1995**, *267*, 1498 – 1502.
[9] G. M. Clore, J. Ernst, R. Clubb, J. G. Omichinski, W. P. Kennedy, K. Sakaguchi, E. Appella, A. M. Gronenborn, *Nat. Struct. Biol.* **1995**, *2*, 321 – 333.
[10] R. B. Hill, D. P. Raleigh, A. Lombardi, W. F. DeGrado, *Acc. Chem. Res.* **2000**, *33*, 745 – 754.
[11] a) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776 – 779; b) P. E. Dawson, S. B. H. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923 – 960.
[12] a) J. A. Camarero, J. Pavel, T. W. Muir, *Angew. Chem.* **1998**, *110*, 361 – 364; *Angew. Chem. Int. Ed.* **1998**, *37*, 347 – 349; b) J. A. Camarero, T. W. Muir, *J. Am. Chem. Soc.* **1999**, *121*, 1369 – 1370.
[13] L. Zhang, J. P. Tam, *J. Am. Chem. Soc.* **1997**, *119*, 2363 – 2370.
[14] M. G. Mateu, M. M. Sanchez Del Pino, A. R. Fersht, *Nat. Struct. Biol.* **1999**, *6*, 191 – 198.
[15] Peptides with a CG- or C-linker were also synthesized and successfully circularized in lower yield.
[16] a) G. R. Jacobson, M. H. Schaffer, G. R. Stark, T. C. Vanaman, *J. Biol. Chem.* **1973**, *248*, 6583 – 6591. b) J. Wu, J. T. Watson, *Protein Sci.* **1997**, *6*, 391 – 398.
[17] This effect is much more pronounced but consistent with the reported subtle changes in CD and enhanced ligand affinity that was observed with a cyclic SH3 domain.[13]
[18] T. Zhang, E. Bertelsen, T. Alber, *Nat. Struct. Biol.* **1994**, *1*, 434 – 438.
[19] C. R. Robinson, R. T. Sauer, *Biochemistry* **2000**, *39*, 12494 – 12502.
[20] B. W. Matthews, H. Nicholson, W. J. Becktel, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 6663 – 6665.
[21] The dimer of catenanes could dissociate into individual catenanes without a change in CD intensity at 222 nm. However, mutations to the dimer interface have produced p53 dimers that are much less stable than the wildtype.[14]
[22] M. Schnölzer, P. Alewood, A. Jones, D. Alewood, S. B. H. Kent, *Int. J. Pept. Protein Res.* **1992**, *40*, 180 – 193.
[23] T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068 – 10073.
[24] P. T. Kraulis, *J. Appl. Crystallogr.* **1991**, *24*, 946 – 950.

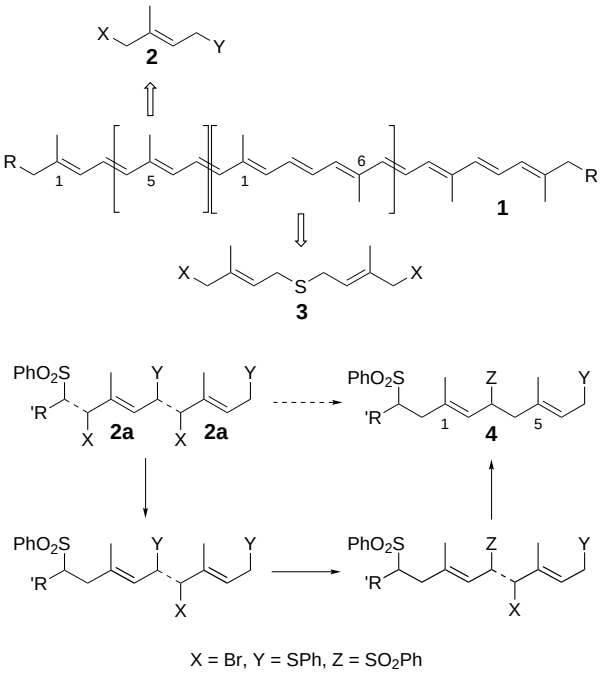
A Highly Efficient Chain-Extension Process in the Systematic Syntheses of Carotenoid Natural Products**

Minkoo Ji, Hojin Choi, Minsoo Park, Minyong Kee, Young Cheol Jeong, and Sangho Koo*

Dedicated to Professor Masato Koreeda on the occasion of his 60th birthday

Nature adopts isopentenyl pyrophosphate (IPP) or dimethylallyl pyrophosphate (DMAP) as building blocks for de novo syntheses of various isoprenoids, such as terpenoids, steroids, and carotenoids.[1] Enzymatic assembly of these basic C_5 units provides a diverse range of natural products. We are currently searching for such a mimic for IPP or DMAP in the chemical syntheses of isoprenoids, especially for carotenoids,[2] which show antioxidant activities for low-density lipoprotein and which are widely used as natural pigments. Certain carotenoids, such as lycopene,[3] also lower the propensity for humans to develop prostate cancers.

Carotenoids **1** have the general structure of two consecutive methyl groups attached to a conjugated polyene chain in either a 1,5 or 1,6 arrangement (Scheme 1). The symmetrical



Scheme 1. Generalized disconnection of the carotenoid structure **1** into the chain-extension unit **2** and the chain-termination unit **3**, as well as the proposed chain-extending process by repeated use of **2a**.

carotenoid structure can be divided into three parts for a systematic construction of such compounds: the chain-initiating unit, the chain-extending unit, and the chain-terminating unit. The Julia sulfone olefination protocol[4] is a perfect methodology to put the three components together, and allylic sulfone compounds were thus selected as the initiating unit. It was envisioned that bifunctionalized C_5 prenyl compound **2** might be an efficient chemical mimic for IPP or DMAP as a chain-extending unit, insofar as X is a good leaving group and Y can be easily transformed to a sulfonyl group. The carbanion-stabilizing ability of Y should be much less than that of the sulfonyl group of the chain-initiating unit to prevent the unfavorable base-promoted halide elimination process from occuring in compound **2**. It was thus appropriate to use 4-bromo-3-methyl-2-butenyl phenyl sulfide **2a** (X = Br, Y = SPh in **2**) to meet the above criteria. Repeated use of this chain-extending unit **2a** with each additional oxidation step to produce the corresponding sulfone would then produce the required 1,5-dimethyl-substituted carbon skeleton, such as structure **4**. This compound could accomplish the chain-extension process. We have already demonstrated that bis(haloallylic) sulfide **3** is a stable substitute for highly unstable 1,8-dihalo-2,7-dimethyl-2,4,6-octatriene in β-caro-

[*] Prof. S. Koo, M. Ji, H. Choi, M. Park, M. Kee, Y. C. Jeong
Department of Chemistry
Myong Ji University
San 38-2, Nam-Dong
Yongin, 449-728, Kyunggi-Do (Korea)
Fax: (+82) 31-335-7248
E-mail: sangkoo@mju.ac.kr

[**] This work was supported by the RRC program of MOST and KOSEF.

Supporting information for this article is available on the WWW under http://www.angewandte.com or from the author.

  1433-7851/01/4019-3627 $ 17.50+.50/0