[16] *N,N′*-Dimethylselenourea: C–Se 1.866, C(Se)–N 1.32 Å: H. M. K. K. Pathirana, T. J. Weiss, J. H. Reibeneis, R. A. Zingaro, E. A. Meyers, *Z. Kristallogr.* **1994**, *209*, 697.
[17] With the length of the [0] bridge between the six-membered rings, calculated with ab initio methods, as an indicator: **3**: 1.461, **4**: 1.379, **5**: 1.353 Å.
[18] H. Schmidbaur, *Angew. Chem.* **1983**, *95*, 980; *Angew. Chem. Int. Ed. Engl.* **1983**, *22*, 907.
[19] We will report on compounds of this type separately.

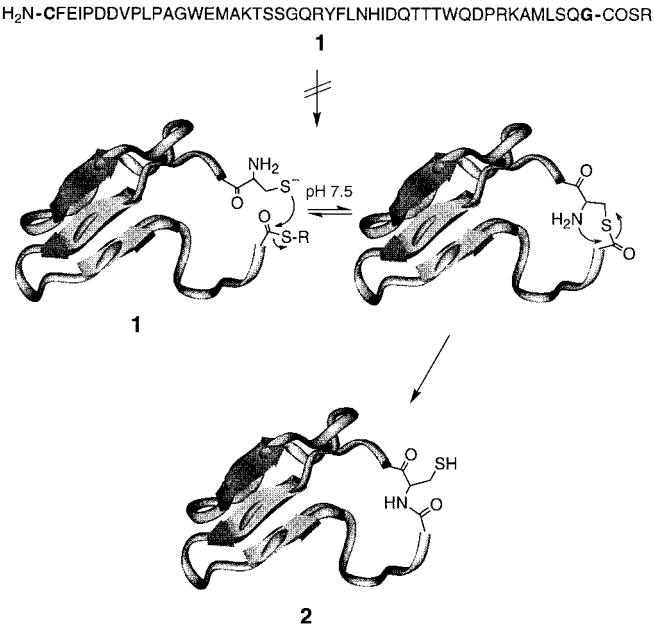
Chemical Synthesis of a Circular Protein Domain: Evidence for Folding-Assisted Cyclization**

Julio A. Camarero, Joanna Pavel, and Tom W. Muir*

The chemical-ligation approach to the condensation of peptide fragments offers a unique route for generating proteins possessing both natural[1] and unnatural topology.[2] Of the various architectures possible, one of the most fascinating from the perspective of both protein engineering and protein folding is the circular topology.[3] Although head-to-tail cyclization is commonly used to constrain the structure of small peptides,[4] relatively little is known about the affect of this modification on the structure and function of a protein. Indeed, the bulk of our current knowledge is based on the pioneering work of Creighton and co-workers on a cyclized version of bovine pancreatic trypsin inhibitor (BPTI) prepared by random chemical cross-linking.[3] Here we demonstrate that a backbone-cyclized protein domain, prepared by an intramolecular version[5] of Kent's native chemical ligation,[6] is able to spontaneously fold into a functional, native-like conformation. In addition, we provide evidence that the rate of cyclization is related to the folded state of the protein.

Our initial studies on protein cyclization focussed on the WW domain from the human Yes kinase-associated protein (YAP).[7] This protein domain possesses a globular structure composed of a three-stranded antiparallel β sheet which positions the flexible N and C termini in close proximity.[8] Given that native chemical ligation is tolerant to the presence of chemical denaturants,[6] it was anticipated that insights into the effect of this juxtaposition on the chemical reaction rate would be forthcoming.

The strategy used to synthesize the circular WW domain is shown in Scheme 1. Key to our approach was the generation



Scheme 1. Synthesis of the circular WW protein domain **2**.

of the linear polypeptide precursor **1**, which contains both groups necessary for native chemical ligation: an N-terminal cysteine residue and a C-terminal thioester group.[9] Cyclization of fully unprotected **1** to give the circular WW domain **2** was remarkably clean and extremely rapid (Figure 1). The ligation product was characterized as **2** by a combination of

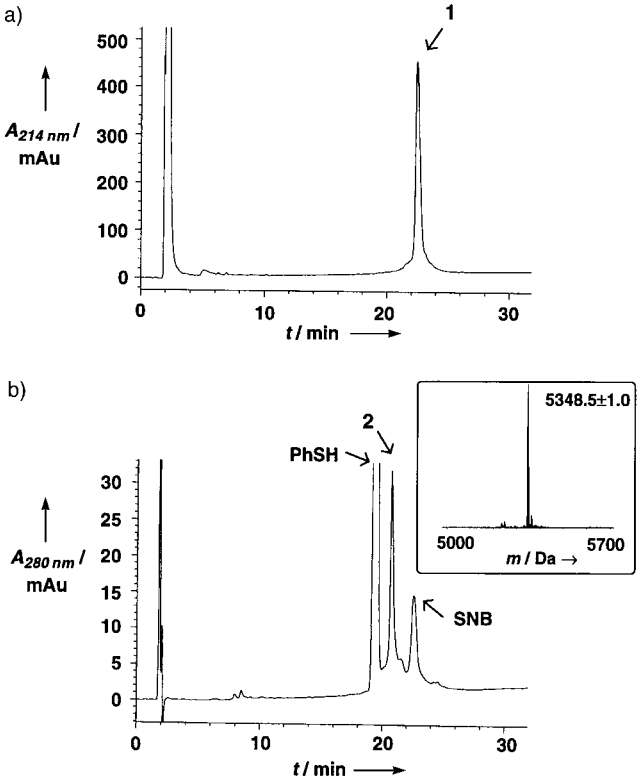


Figure 1. a) Analytical HPLC chromatogram of the purified linear precursor **1**. b) The crude ligation mixture after 2 min of reaction on the same scale as in a). mAu = 10^{-3} of full scale absorbance units. SNB = *S*-(5-sulfenyl-2-nitrobenzoic acid), the by-product. Inset: ES-MS of purified **2**.

[*] Dr. T. W. Muir, Dr. J. A. Camarero, J. Pavel
The Laboratory of Synthetic Protein Chemistry
The Rockefeller University
Box 223, 1230 York Avenue, New York, NY 10021 (USA)
Fax: Int. code + (1) 212 327-7358
E-mail: muirt@rockvax.rockefeller.edu

[**] This work was performed under the auspices of the Rockefeller University. J.A.C. was supported by a postdoctoral fellowship from the Ministerio de Educacion y Ciencia of Spain. T.W.M. gratefully acknowledges the support of the US National Institutes of Health (GM55843-01). We are grateful to Dr. Sean Cahill for help with NMR experiments, and Dr. Harmut Oschkinat for useful discussions and for providing NMR assignments of the recombinant YAP WW domain–ligand complex.

  1433-7851/98/3703-0347 $ 17.50+.50/0

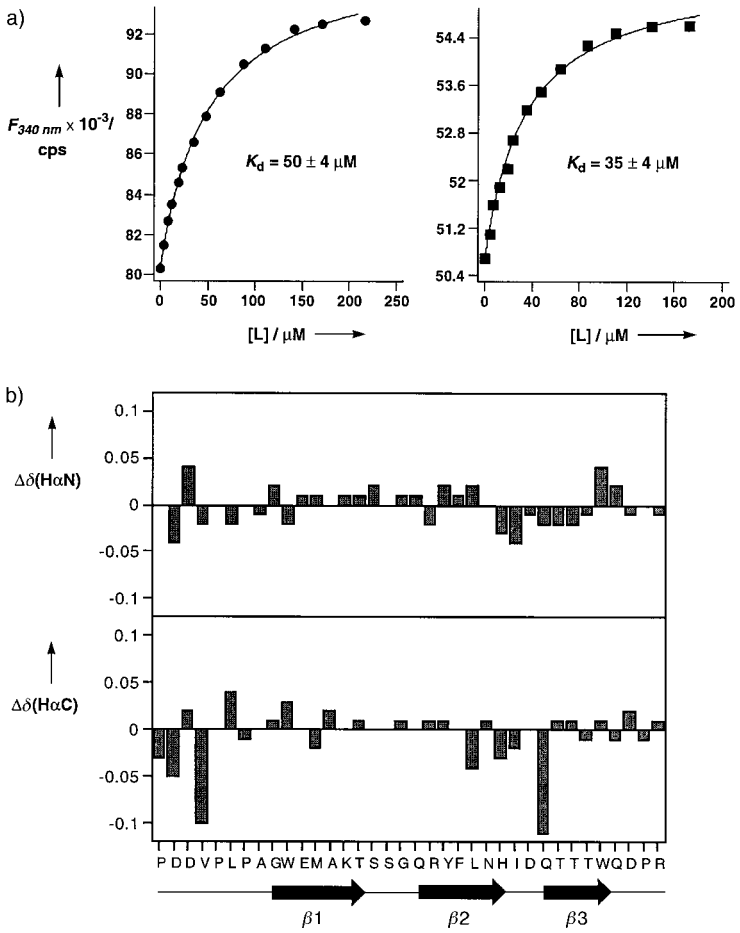


Figure 2. a) Binding of the linear (●) and circular (■) forms of the WW domain to the proline-rich ligand Ac–GTPPPPYTVG–NH_2 (L); cps = counts per scan. b) Summary of the ^{1}H NMR assignments for the main-chain protons: $\Delta\delta$(HαN) and $\Delta\delta$(HαC) are the deviations in the chemical shifts of the backbone protons in the circular WW domain from those in the the linear WW domain. The location of the secondary structure elements are shown at the bottom.[8]

electrospray mass spectrometry (ES-MS), tryptic digestion, and Edman sequencing.[10]

The YAP WW domain binds to proline-rich peptides containing the consensus sequence PPxY.[11] A fluoresence-based ligand-binding assay was used to study the functional consequences of backbone cyclization of the YAP WW domain (Figure 2a). As predicted based on the native structure of the protein, cyclization of the WW domain has no deleterious effect on ligand-binding activity. Indeed, the circular protein has a slightly higher affinity for ligands than its linear counterpart (dissociation constants $K_{d(cyclic)} = 35$ and $K_{d(linear)} = 50\,\mu$M).[12] Interestingly, the ligand affinity was the same regardless of whether the cyclic protein had been prepared in the presence or absence of 6 M guanidinium chloride (GdmCl), suggesting that the folding process is not inhibited by the cyclic topology.

The solution structures of the circular and linear YAP WW domains complexed to ligands were studied by two-dimensional homonuclear NMR spectroscopy. Deviations in the chemical shifts for the main-chain protons in the core β-sheet structure of the linear[13] and circular domains were uniformly less than 0.05 ppm, indicating that the same global fold is common to both versions of the molecule (Figure 2b). The NMR samples were prepared by simply dissolving the proteins, which had been purified by HPLC, in phosphate buffer containing the ligand. Given that HPLC purification is denaturing, this again indicates that the circular protein, like its linear counterpart, can spontaneously fold in the presence of the ligand.

The influence of the fold in the WW domain on the cyclization rate was investigated by performing the reaction in the presence and absence of chemical denaturant. Addition of 6 M GdmCl to the ligation buffer had a marked effect on the rate of cyclization: The reaction was approximately one order of magnitude faster under folding conditions than under unfolding conditions (Figure 3). Interestingly, linear polypeptide **1** was cleanly converted into cyclic **2**, regardless of whether 6 M GdmCl was present or absent in the ligation buffer. This rate enhancement on going from unfolding to folding conditions is significant and can be understood in terms of the proximity of the N and C termini within the native conformation of the protein. Under folding conditions the local concentration of the cysteinyl and thioester reactive functionalities would be expected to be higher than under unfolding conditions; this leads to a more rapid reaction for the former.[14] Therefore, the cyclization reaction is assisted by the native conformation of the protein.

We have generated a functional circular version of a WW protein domain by intramolecular chemical ligation. It is expected that several emerging areas in protein engineering will benefit from the ability to efficiently prepare backbone-cyclized polypeptides from unprotected linear precursors. For example, backbone cyclization provides a general way of constraining minimized protein sequences and complements the use of disulfide bonds.[15] Equally, new opportunities are created in sequence permutation by combining chemoselective cyclization with subsequent main-chain cleavages. Finally, our observation that the kinetics of protein cyclization are sensitive to changes in protein structure

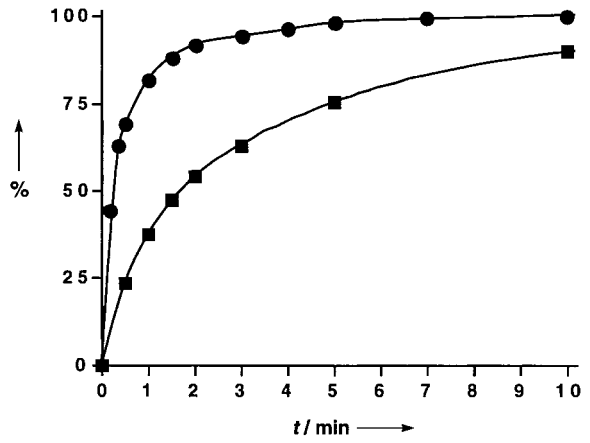


Figure 3. Progress of the cyclization under folding (0 M GdmCl, ●) and unfolding conditions (6 M GdmCl, ■).

  1433-7851/98/3703-0348 $ 17.50+.50/0

suggests a potential application of this approach in the study of protein-structure dynamics.

Experimental Section

All peptide synthesis were carried out according to the in situ neutralization/HBTU activation protocol for Boc solid-phase peptide synthesis (HBTU = benzotriazol-1-yl-*N*-tetramethyluronium hexafluorophosphate).[16] Peptide thioester **1** was prepared by selective thioesterification of the corresponding peptide thioacid (prepared on a glycine–thioester support)[17] with 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB, Ellman's reagent) as previously described[5a] and purified by preparative HPLC. ES-MS: observed mass: 5546.3(1.4) Da; calcd for $C_{244}H_{361}N_{65}O_{76}S_4$: 5545.5 (monoisotopic), 5549.0 (average isotope composition).

Cyclization: Purified **1** (0.4 μmol, 2.2 mg) was dissolved to a final concentration of about 50 μM in a freshly degassed buffer containing 0.1 M sodium phosphate and 1 mM ethylenediamine tetraacetate (EDTA) at pH 7.5 and 0.05 volume % each of PhSH and BnSH. In cyclization reactions carried out under unfolding conditions, 6 M GdmCl was included in the above buffer. To monitor the course of the reaction, small aliquots (50 μL) of the reaction mixture were periodically removed, quickly quenched by addition of 5% trifluoroacetic acid in water (20 μL), and investigated by analytical HPLC. The degree of conversion of **1** into **2** was calculated based on the absorption at 280 nm. The overall yield for the cyclization reaction following HPLC purification was 1.8 mg (0.32 μmol, 81%). Circular protein **2** was characterized by ES-MS and tryptic digestion. ES-MS: observed mass: 5348.5(1.0) Da; calcd for $C_{237}H_{356}N_{64}O_{72}S_3$: 5346.5 (monoisotopic), 5350.0 (average isotope composition).

Tryptic digestion of **2**: Cyclic **2** (100 μg) was dissolved in 0.1 M tris(hydroxymethyl)aminomethane (Tris)/HCl buffer (pH 8.4, 200 μL). Lyophilized trypsin (10 μg, sequencing grade, Sigma) was then added, and the reaction maintained at 37 °C for 3 h; a solution of 1,4-dithiothreitol (DTT, 50 mM, 20 μL) was subsequently added. The three expected fragments T1 (H–AMLSQGCFEIPDDVPLPAGWEMAK–OH), T2 (H–TSSGQR–OH), and T3 (H–YFLNHIDQTTTWQDPR–OH) were isolated by HPLC and characterized by ES-MS and Edman sequencing.

The isotherms for ligand binding were obtained by monitoring the intrinsic Trp fluoresence at 340 nm as a function of the concentration of the ligand at 22 °C. In all cases, the protein concentration was kept at 2 μM in a buffer containing 40 mM sodium phosphate, 50 mM NaCl, and 1 mM DTT at pH 7.2. Equilibrium dissociation constants were extracted from the curves by assuming a 1:1 receptor–ligand complex.[18]

Two-dimensional ^{1}H NMR spectra were measured on a Bruker DPX-400 spectrometer (400.132 MHz). The NMR samples were dissolved in a buffer containing 10 mM sodium phosphate, 100 mM NaCl, 1 mM [D_{10}]DTT, 0.02% NaN_3, and 10% D_2O at pH 6.2. A series of TOCSY and NOESY experiments[19] were recorded at 278 K on samples containing both linear and circular WW domains with WW ligand (protein concentrations 0.6–1.2 mM; ratio of peptide to protein 3:1). Typical mixing times were 30 ms for TOCSY experiments and 50–250 ms for NOESY experiments.

Received: June 30, 1997 [Z 10614 IE]
German version: *Angew. Chem.* **1998**, *110*, 361–364

Keywords: cyclizations · peptides · protein structures · solid-phase synthesis

[1] a) T. W. Muir, *Structure* **1995**, *3*, 649–652; b) C. J. A. Wallace, *Curr. Opin. Biotech.* **1995**, *6*, 403–410.
[2] a) P. E. Dawson, S. B. H. Kent, *J. Am. Chem. Soc.* **1993**, *115*, 7263–7266; b) A. Nefzi, X. Sun, M. Mutter, *Tetrahedron Lett.* **1995**, *36*, 229–230; c) L. E. Canne, A. R. Ferre-D'Amare, S. K. Burley, S. B. H. Kent, *J. Am. Chem. Soc.* **1995**, *117*, 2998–3007; d) J. Shao, J. Tam, *ibid.* **1995**, *117*, 3893–3899; e) K. J. Rose, *ibid.* **1994**, *116*, 30–33 ; f) T. W. Muir, M. J. Williams, M. H. Ginsberg, S. B. H. Kent, *Biochemistry* **1994**, *33*, 7701–7708.
[3] a) D. P. Goldenberg, T. E. Creighton, *J. Mol. Biol.* **1984**, *179*, 527–545; b) D. P. J. Goldenberg, *J. Cell. Biochem.* **1985**, *29*, 321–335; c) D. P. Goldenburg, T. E. Creighton, *J. Mol. Biol.* **1983**, *165*, 407–413.
[4] a) V. J. Hruby, *Life Science* **1982**, *31*, 189–199; b) H. Kessler, *Angew. Chem.* **1982**, *94*, 509–520 ; *Angew. Chem. Int. Ed. Engl.* **1982**, *21*, 512–523; c) J. Rizo, L. M. Gierasch, *Ann. Rev. Biochem.* **1992**, *61*, 387–418.
[5] a) J. A. Camarero, T. W. Muir, *Chem. Commun.*, **1997**, 1369–1370; b) L. Zhang, J. P. Tam, *J. Am. Chem. Soc.* **1997**, *119*, 2363–2370.
[6] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776–779.
[7] WW domains are a novel class of protein modules which are found in a variety of signal-transducing and cytoskeletal proteins. Reviews: a) M. Sudol, *Prog. Biophys. Molec. Biol.* **1996**, *65*, 113–132; b) M. Sudol, H. I. Chen, C. Bougeret, A. Einbond, P. Bork, *FEBS Letters* **1995**, *369*, 67–71.
[8] M. J. Macias, M. Hyvonen, E. Baraldi, J. Schultz, M. Sudol, M. Saraste, H. Oschkinat, *Nature* **1996**, *382*, 646–649.
[9] To facilate native ligation, Cys and Gly were added at the N and C termini, respectively, of the core WW sequence (residues 165–209 of the human YAP protein).
[10] Cyclic protein **2** was digested with trypsin, and the peptide fragment encompassing the ligation site (AMLSQ**GC**FEIPDDVPLPAGWEMAK) characterized by Edman sequencing.
[11] H. I. Chen, M. Sudol, *Proc. Natl. Acad. Sci. USA* 1995, 92, 7819–7823.
[12] The linear WW domain was also chemically synthesized by Boc solid-phase protein synthesis and purified by standard procedures.
[13] Residues located at the immediate N and C termini could not be unambiguously assigned in either the linear or cyclic WW domains.
[14] Large effective molarities were described for many intramolecular organic reactions: F. M. Menger, *Acc. Chem. Res.* **1985**, *18*, 128–134.
[15] B. Li, J. Y. K. Tom, D. Oare, R. Yen, W. J. Fairbrother, J. A. Wells, B. C. Cunningham, *Science* **1995**, *270*, 1657–1660.
[16] M. Schnolzer, P. Alewood, A. Jones, D. Alewood, S. B. H. Kent, *Int. J. Pept. Protein Res.* **1992**, *40*, 180–193.
[17] D. Yamashiro, C. H. Li, *Int. J. Pept. Protein Res.* **1988**, *31*, 322–334.
[18] A. R. Viguera, J. L. Arrondo, A. Musacchio, M. Saraste, L. Serrano, *Biochemistry* **1994**, *33*, 10925–10933.
[19] a) A. Bax, D. G. Davis, *J. Magn. Reson.* **1985**, *65*, 355–360; b) A. Kumar, R. R. Ernst, K. Wuthrich, *Biochem. Biophys. Res. Commun.* **1980**, *95*, 1–6.

A Symmetrically Bridging Triarylsilyl Ligand in a Dinuclear Rhodium Complex: Synthesis and Structure of [LRh(H)(μ-Cl)(μ-SiAr$_2$)(μ-SiAr$_3$)Rh(H)L] (Ar = Ph, *p*-FC$_6$H$_4$; L = P*i*Pr$_3$)**

Kohtaro Osakada,* Take-aki Koizumi, and Takakazu Yamamoto*

Bridging coordination of a secondary silyl group ($SiHR_2$) in dinuclear complexes has become common for late transition metals.[1] The complexes reported so far contain both a σ bond between Si and one metal and a three-center, two-electron bond (3c–2e; π bond) between the Si–H group and the second metal center (Scheme 1a).

[*] Prof. Dr. K. Osakada, Prof. Dr. T. Yamamoto, T. Koizumi
Research Laboratory of Resources Utilization
Tokyo Institute of Technology
4259 Nagatsuta, Midori-ku, Yokohama 226 (Japan)
Fax: Int. code + (81) 45 924-5276
E-mail: kosakada@res.titech.ac.jp

[**] This work was supported by the Japanese Ministry of Education, Science, Culture, and Sports for Grants-in-Aid for Scientific Research (nos. 09239212 and 09440229). T.K. is grateful to the Japan Society for Promotion of Science for a Grant-in-Aid as a JSPS fellow.

  1433-7851/98/3703-0349 $ 17.50+.50/0