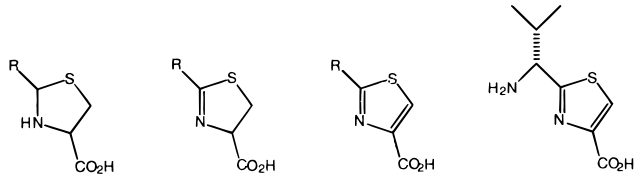
Cyclooligomerization of Thiazole-Containing Tetrapeptides. Symmetrical Macrocycles with up to 76 Amino Acids

Nikolai Sokolenko,† Giovanni Abbenante,† Martin J. Scanlon,† Alun Jones,† Lawrence R. Gahan,‡ Graeme R. Hanson,§ and David P. Fairlie*,†

Centre for Drug Design and Development Department of Chemistry, and Centre for Magnetic Resonance The University of Queensland, Brisbane, Qld 4072, Australia

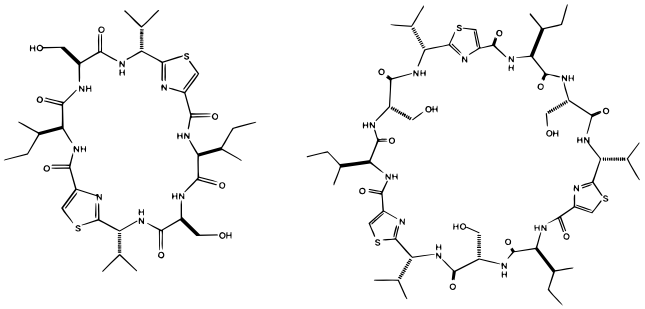
Received September 21, 1998

Thiazolidine (**1a**), thiazoline (**1b**) and thiazole (**1c**), derived from the amino acid cysteine, are frequently found in naturally occurring cyclic peptides[1] many of which possess cytotoxic,



1a **1b** **1c** **1d**

antitumor and antibiotic properties. The structural effects of these five-membered rings have only been elucidated in a few cyclic peptides, but it is clear that they can profoundly affect macrocycle conformation.[1c] We previously demonstrated that the dipeptide derivative D-valine(thiazole) (**1d**) was a β-turn inducing constraint in cyclic octapeptides.[2] We now report that a one-pot cyclooligomerization of the tetrapeptide derivative H-Ile-Ser-D-Val(Thz)-OH produces a novel series (**2**−**19**) of constrained macrocycles cyclo-[-Ile-Ser-D-Val(Thz)-]$_n$ where $n = 2-19$, containing up to 76 amino acids (228 atoms) in the cycle. Their extraordinary size, high symmetry, conformational restraints, and potential to carry multiple functional groups may facilitate development of extended symmetrical scaffolds for new macromolecular devices, protein surface mimics[3], multiple metal-ion carriers[4] or novel antibiotics.



2 ($n = 2$) **3** ($n = 3$)

The tetrapeptide analogue H-Ile-Ser-D-Val(Thz)-OH reacts with BOP/DIPEA in DMF[5] to give a concentration-dependent series of novel cyclooligomers **2**−**19** of formula cyclo-[Ile-Ser-D-Val-(Thz)-]$_n$, $n = 2-19$ (Figure 1). At low peptide concentrations ($10^{-2}-10^{-4}$ M) the predominant product is cyclodimer, while insoluble polymers are formed at the limit of solubility (0.35M). At 0.15 M tetrapeptide, up to 50% soluble cyclic peptides accompanied insoluble polymer. This cyclic peptide mixture, based upon mass ratios,[5] was 16% dimer, 18% trimer, and 66% higher oligomers including 3% ($n = 10$) and 24% ($n = 11-19$).[5] Cyclic products gave uniformly separated rpHPLC retention times (Figure 1) and were readily identified by tandem HPLC/MS.

Electrospray[6] mass spectrometry was used to characterize each macrocycle. Compounds **2** and **3** (dimer, trimer) show the expected molecular ions at *m*/*z* 765 and 1147 (Figure 2a). However cyclo-[Ile-Ser-D-Val(Thz)-]$_n$, $n \geq 5$, produced the most abundant molecular ions at *m*/2*z* (e.g., Figure 2a, $n = 5$, 8, 10). Isotopic component peaks varied by 0.5 mass unit, demanding their formulation as *m*/2*z*, and reconstructed *m*/*z* molecular ions are shown in Figure 2b. Molecular weights were confirmed by the precise match of peak shapes to calculated isotopic distribution patterns (Figure 2b) and by their MALDI-TOF mass spectra.[6] Thus, incremental differences in mass units of [-Ile-Ser-D-Val-(Thz)-] and unique isotopic patterns together characterized each macrocycle.

The ^{1}H NMR spectra[7] for cyclo-[Ile-Ser-D-Val(Thz)]$_n$ in d_6-DMSO (Figure 3, Supporting Information) or d_6-acetone display only three amide NH resonances. When $n = 2$, intramolecular hydrogen bonding is implicated for Val-NH and Ile-NH protons. In the variable temperature ^{1}H NMR spectra (Figure 4, Supporting Information) for cyclo-[Ile-Ser-D-Val(Thz)]$_2$ between 280 and 320 K in d_6-acetone, the amide chemical shifts for ValNH and IleNH resonances are temperature independent[8] ($\Delta\delta/T$ (ppb/deg) = 1.4, ValNH; 2.8, IleNH) but not for the SerNH signal (5.4 ppb/deg). Also ValNH and IleNH protons undergo much slower H−D exchange than SerNH in d_6-acetone/CD_3OD (9:1) (Figure 5, Supporting Information). Amide coupling constants $^3J_{NH-CH\alpha}$ (5.5 Hz, Ile; 9.5 Hz, Val; 6.0 Hz, Ser) are typical of conformationally averaged random coil values (∼6−7 Hz)[9] for SerNH but atypical for ValNH or IleNH residues where local structure is indicated. These data are consistent with intramolecular hydrogen bonds rendering the cycle planar and symmetric.

For higher oligomers, cyclo-[Ile-Ser-D-Val(Thz)]$_n$ $n \geq 3$, the conformation of the molecules is more ambiguous. Coupling

* To whom correspondence should be addressed. E-mail: d.fairlie@mailbox.uq.edu.au.
† Centre for Drug Design and Development.
‡ Department of Chemistry.
§ Centre for Magnetic Resonance.

(1) (a) Faulkner, D. J. *Nat. Prod. Rep.* **1998**, *15*, 113−158; **1997**, *14*, 259−302; **1996**, *13*, 75; **1995**, *12*, 233; **1994**, *11*, 355. (b) Wipf, P. *Chem. Rev.* **1995**, *95*, 2115. (c) Fairlie, D. P.; Abbenante, G.; March, D. R. *Curr. Med. Chem.* **1995**, *2*, 654−686. (d) McGeary, R. P.; Fairlie, D. P. *Curr. Opin. Drug Discovery Devel.* **1998**, *1*, 208−217. (e) Fate, G. D.; Benner, C. P.; Grode, S. H.; Gilbertson, T. J. *J. Am. Chem. Soc.* **1996**, *118*, 11363−11368.

(2) Abbenante, G.; Fairlie, D. P.; Gahan, L. R.; Hanson, G. R.; Pierens, G.; van den Brenk, A. L. *J. Am. Chem. Soc.* **1996**, *118*, 10384−8.

(3) (a) Fairlie, D. P.; West, M. W.; Wong, A. K. *Curr. Med. Chem.* **1998**, *5*, 29−62. (b) Schneider, J. P.; Kelly, J. W. *Chem. Rev.* **1995**, *95*, 2169−2187. (c) Regan, L.; DeGrado, W. F. *Science* **1988**, *241*, 976−978. (d) Ho, S. P.; DeGrado, W. F. *J. Am. Chem. Soc.* **1987**, *109*, 6751−6758.

(4) Dagani, R. *Chem. Eng. News* **1998**, June 8, 35−46.

(5) The synthesis of H-Ile-Ser-D-Val(Thz)-OH is described in Supporting Information and Boc-D-Val(Thz)-OH was made by an established method.[11] TFA·H-Ile-Ser-D-Val(Thz)-OH (20 mg, 0.04 mmol) and BOP (50 mg, 0.11 mmole) were dissolved in DMF (247 μL) to which DIPEA (20 μL, 0.12 mmole) was added. The mixture was stirred for 2 h at 20 °C, the solvent was evaporated in vacuo, and the residue was dissolved in water/acetonitrile (1:1) and lyophilised to a powder. Crude reaction mixtures were analysed by rpHPLC using acetonitrile/water eluant gradients, and cyclic peptides, [-Ile-Ser-D-Val(Thz)-]$_n$ for $n = 2-19$, typically accounted for 40% total products with relative mass ratios 16:18:11:8.8:6.4:4.8:3.8:3.3:3.4:3.3:2.7:2.6:2.6:2.3:2.8:2.3:2.1:2.1:1.7 for $n = 2-19$ (Figure 1). Isolated compounds ($n = 2-10$) were characterized by retention time, MS, 2D ^{1}H NMR spectroscopy, and a negative ninhydrin assay for free amine. The experiment was repeated at other concentrations by varying the volume of DMF.

(6) Electrospray MS: (a) Bruins, A. P.; Covey, T. R.; Henion, J. D. *Anal. Chem.* **1987**, *59*, 2642. (b) Iribane, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, *64*, 2287. A MALDI-TOF mass spectrum (Figure 9, Supporting Information) of the cyclooligomerization mixture showed *m*/*z* molecular ions for cycles $n = 11-19$.

(7) ^{1}H NMR chemical shift assignments were made on the basis of COSY, TOCSY, and presaturation experimental data for all cycles.

(8) Kessler, H. *Angew Chem., Int. Ed. Engl.* **1982**, *21*, 512−523.

(9) Dyson, H. J.; Wright, P. E. *Annu. Rev. Biophys. Chem.* **1991**, *20*, 519−538.

10.1021/ja983354n CCC: $18.00
Published on Web 03/09/1999

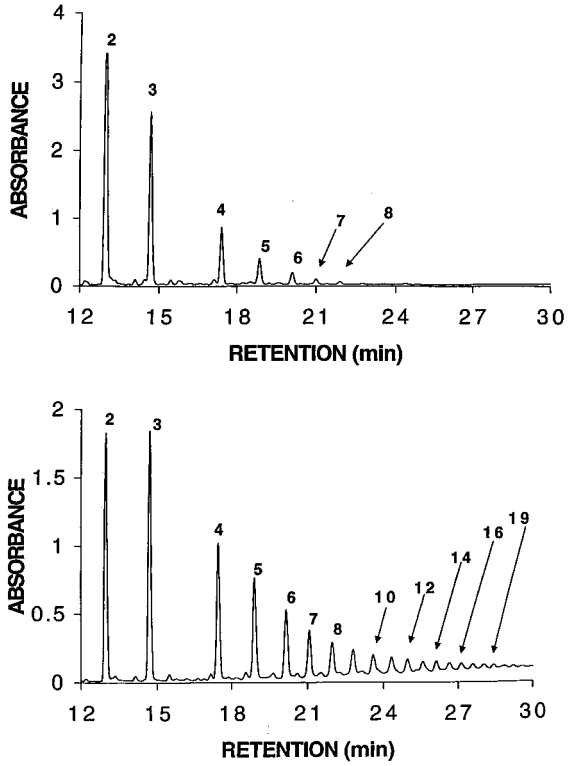


Figure 1. Reversed phase HPLC analysis of crude product mixtures (top: 0.01 M; bottom: 0.15 M) obtained directly from the cyclooligomerization of H-Ile-Ser-D-Val(Thz)-OH with BOP and DIPEA in DMF for 2 h at room temperature. Numbers beside peaks designate *n* for [-Ile-Ser-D-Val(Thz)]$_n$, as determined from mass spectra.

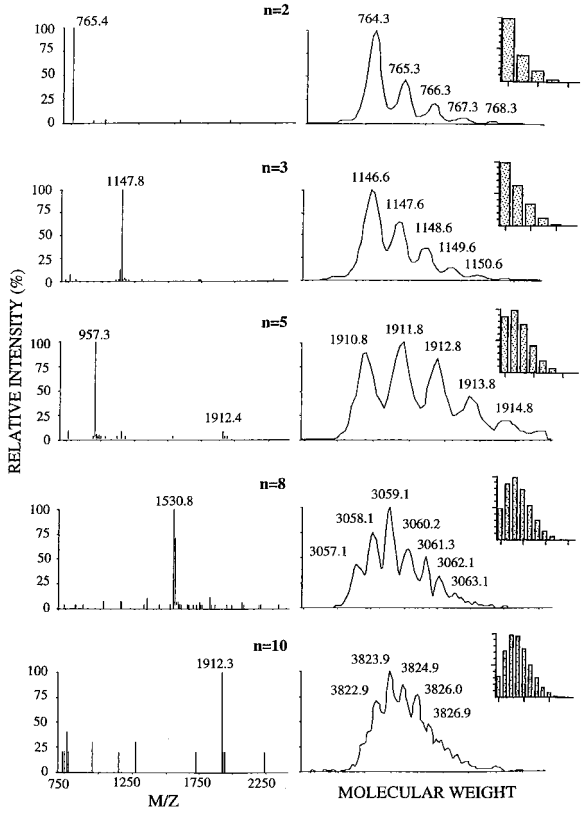


Figure 2. Electrospray mass spectra. (a) Left panel shows observed molecular ions for isolated compounds [Ile-Ser-D-Val(Thz)]$_n$, $n = 2$, 3, 5, 8, 10. (b) Right panel shows experimental masses and isotope distributions (theoretical isotope distributions shown as insets). MALDI-TOF mass spectra confirmed molecular weights (e.g., Figure 9, Supporting Information).

constants ($^3J_{NH-CH\alpha}$) for IleNH, ValNH, SerNH are 9, 8, 7.5 Hz respectively, consistent with local structure around Ile and Val residues. Deuterium exchange data in d_6-DMSO/CD_3OD (9:1) indicate that Ile-NH and Val-NH amide protons are protected from solvent exchange relative to the Ser-NH proton which exchanges more rapidly (Figure 6, Supporting Information). However ^{1}H NMR chemical shifts for the amide-NH resonances of [Ile-Ser-D-Val(Thz)]$_n$ in d_6-acetone (e.g., $n = 6$; Figure 7, Supporting Information) are somewhat temperature dependent ($\Delta\delta/T = 3$, 4, 6 ppb/deg; IleNH, SerNH, ValNH). Also cooling the samples to −90 °C (CD_3OD) caused slight broadening of the amide resonances indicating multi-site exchange. The probability of multiple conformations in the larger cycles ($n \geq 3$) makes structure interpretation complex. There may be local structure around each thiazole ring, but there is no unequivocal evidence of intramolecular hydrogen bonding.

Cyclooligomerization of H-Ile-Ser-D-Val-Ala-OH under identical conditions (0.15 M) gave only traces of cyclic dimer and trimer but no higher cyclooligomers were detected by rpHPLC or MS. This difference suggests that the thiazole encourages formation of higher cyclooligomers and that other turn-inducing constraints might promote similar cyclizations. Indeed we find that H-Ile-Ser-D-Val-Pro-OH and H-Ile-Ser-L-Val-Pro-OH, but not H-Ile-Ser-mamba-OH (mamba = *m*-aminomethyl benzoic acid), form cyclooligomers up to $n = 8$ under the same conditions (Figure 8, Supporting Information). Thus, cyclization competes effectively with oligomerization, even at peptide concentrations that favor intermolecular coupling, due to the presence of turn-inducing units which preorganize the peptide for cyclization. Ratios of cyclic oligomers to linear polymers were highest at 0.15 M tetrapeptide for: D-Val(Thz) (50%) > D-Val-Pro ≫ L-Val-Pro ≫ D-Val-Ala, mamba.

Models indicate that these macrocycles range in diameter from 10 to 50 Å and resemble large loop regions and cyclic peptide domains of proteins. Bioactive regions of proteins frequently involve discontinuous surfaces that span large areas and mimicry of these regions presents a great challenge.[3] Macrocycles such as these may guide development of conformationally restricted templates that span vast areas required for positioning attached peptide motifs to mimic protein surfaces. The functionalized serine side chain can be condensed onto the backbone to form additional oxazoline ring constraints[10] or derivatized to permit the use of "planar" macrocycles as large templates. Cyclooligomerizations of other functionalized turn-containing peptides may similarly generate combinatorial macrocycle libraries with conceivable uses as new antibiotics, new metal sequestrants, or as scaffolds for developing macromolecular devices and protein mimetics.

Acknowledgment. This work was supported in part by the Australian Research Council.

Supporting Information Available: Figures 3−9 show ^{1}H NMR spectra for isolated cycles [Ile-Ser-D-Val(Thz)]$_n$, $n = 3$, 5, 8, 10; VT-NMR data for $n = 2,3$; H−D exchange rates for $n = 2,3$; HPLC profiles for cyclooligomerization of 0.15 M solutions of H-Ile-Ser-D-Val-Pro-OH and H-Ile-Ser-mamba-OH; MALDI-TOF mass spectrum of mixture from Figure 1; and synthesis and characterization data for all of the macrocycles (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA983354N

(10) Sugiura, T.; Hamada, Y., Shioiri, T. *Tetrahedron Lett.* **1987**, *20*, 2251−54.

(11) Aguilar, E.; Meyers, A. I. *Tetrahedron Lett.* **1994**, 35, 2473. For other methods of synthesis of thiazole amino acids, see: Boden, C. D.; Pattenden, G.; Tao Ye *Synlett* **1995**, 417−9.