Pergamon

Bioorganic & Medicinal Chemistry Letters 11 (2001) 697–700

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

Carboproteins: A 4-α-Helix Bundle Protein Model Assembled on a D-Galactopyranoside Template

Jesper Brask and Knud J. Jensen*

Department of Organic Chemistry, Technical University of Denmark, Building 201, Kemitorvet, DK-2800 Kgs. Lyngby, Denmark

Received 7 November 2000; revised 14 December 2000; accepted 10 January 2001

Abstract—We have recently introduced the concept of monosaccharides as templates for de novo design of protein models and described the synthesis of a model 'carbopeptide'. Here, we report the synthesis of a 64 amino acid (AA) 'carboprotein' by chemoselective ligation of a C-terminal hexadecapeptide aldehyde to a tetra-aminooxy functionalized methyl α-D-galactopyranoside (D-Gal*p*) template. Biophysical characterizations by CD spectroscopy and NMR amide H–D exchange experiments indicated that the four-stranded carboprotein forms a 4-α-helix bundle structure. © 2001 Elsevier Science Ltd. All rights reserved.

Introduction

De novo protein design has emerged as a valuable tool to critically test our understanding of protein folding and structure.[1] A design approach relying on the anchoring of peptide strands to a molecular template has been proposed by Mutter and co-workers[2] and later explored also by others, for example, the groups of Sasaki,[3] DeGrado,[4] Ghadiri,[5] and Goodman.[6] The main advantage of this totally synthetic strategy is a bypassing of the problematic folding step of linear polypeptides, thereby allowing more freedom in the sequence design for the construction of small structures with native-like folding. The 4-α-helix bundle, consisting of four amphiphilic helices forming a hydrophobic core, has been a target for a number of de novo designed proteins. In recent contributions, templated 4-α-helix bundle constructs have been synthesized on peptide templates[7,8] or aromatic derivatives.[9,10]

In previous reports, we introduced carbohydrates as templates for de novo design of protein models based on their multifunctionality, the relative rigidity of pyranose rings, the availability of epimeric forms, and their well-described chemistry.[11,12] Several groups have used the anchoring of peptide chains through the modified *O*-6 of cyclodextrin scaffolds as a means to improve the pharmacological properties of peptides,[13,14] to aid in drug targeting,[15] or to allow molecular sensing.[16] However, using the primary and secondary hydroxyls of mono- or disaccharides should provide a more flexible control of the directionality and distances between anchoring points for the peptide chains in de novo design of protein models. We employed a methyl α-D-Gal*p* derivative in which the non-anomeric hydroxyls are in or above a plane defined by *C*-2, *C*-3, *C*-5, and *O*-5, and suggested the terms 'carbopeptide' or 'carboprotein' for the chimeric constructs, depending on their size and degree of folding. Aminooxyacetyl (Aoa) functionalization of the template allowed a convergent strategy for carbopeptide synthesis, in which C-terminal peptide aldehydes were coupled by oxime[17] formation.[12] Here we report an application of this strategy for the synthesis of a single-stranded 16 AA carbopeptide (**1**), a four-stranded 48 AA carbopeptide (**2**), and a four-stranded 64 AA carboprotein (**3**). We further present CD and NMR biophysical studies of the three structures to evaluate the stabilization provided by the carbohydrate templates.

Synthesis

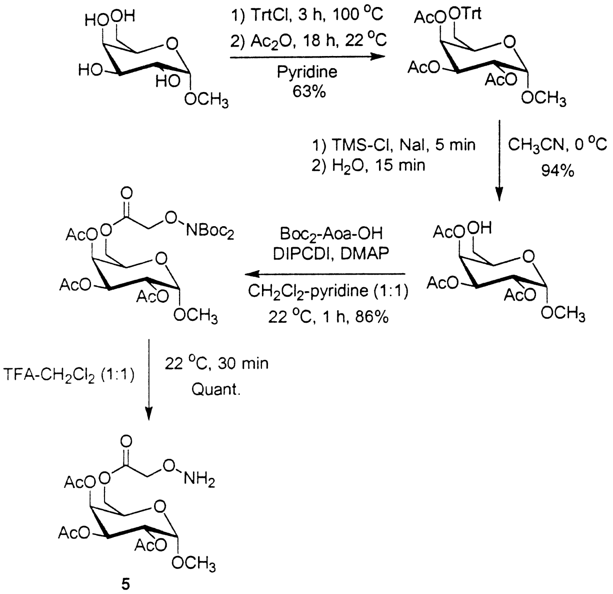
The previously reported[12] methyl 2,3,4,6-tetra-*O*-Aoa-α-D-Gal*p* (**4**) was used as a template for the four-stranded carbopeptide **2** and carboprotein **3**. A new template, methyl 6-*O*-Aoa-2,3,4-tri-*O*-acetyl-α-D-Gal*p* (**5**), was required for the single-stranded carbopeptide **1** (Scheme 1). Methyl α-D-Gal*p* was tritylated at *O*-6 and acetylated at *O*-2, *O*-3, and *O*-4 in a two-step–one-pot pro-

---

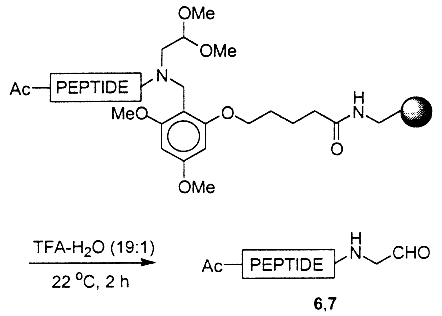
*Corresponding author. Tel.: +45-4525-2147; fax: +45-4593-3968; e-mail: okkj@pop.dtu.dk

PII: S0960-894X(01)00035-X

cedure. A potential problem in the synthesis is acetyl migration to *O*-6 under conditions of acidic trityl deprotection. However, deprotection with in situ formed TMS-I[18] prevented this problem as less than 2% migration was observed by NMR. The desired template **5** was then obtained after DIPCDI/DMAP acylation with Boc_2-Aoa-OH, followed by Boc removal with TFA–CH_2Cl_2.



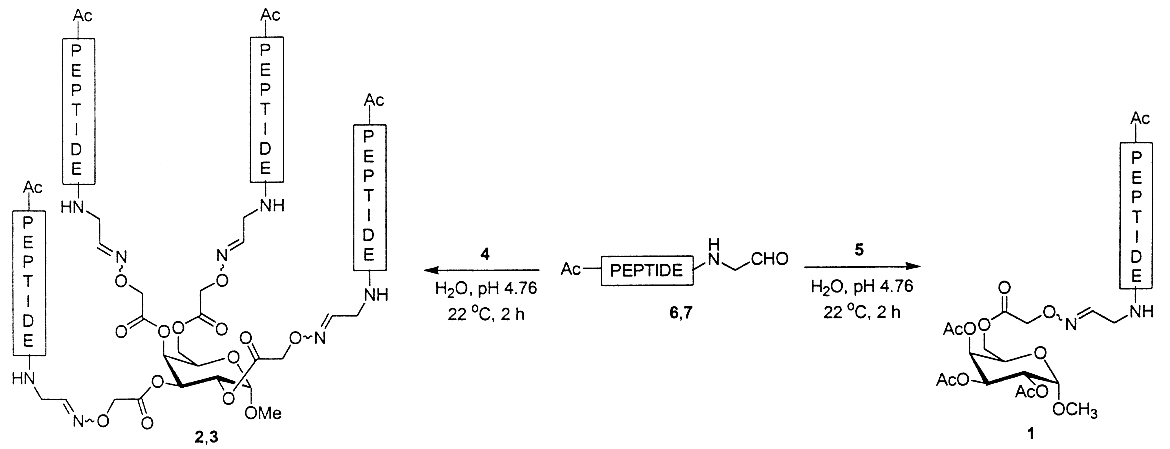
Scheme 1. Synthesis of methyl 6-*O*-Aoa-2,3,4-tri-*O*-acetyl-α-D-Gal*p* (**5**).



Scheme 2. Synthesis of C-terminal peptide aldehydes by a BAL strategy.

The sequence of the peptide aldehyde building blocks was based on an amphiphilic peptide used by Mutter et al.[19] in the synthesis of a 4-α-helix bundle template-assembled synthetic protein (TASP), T_4-($4\alpha_{15}$-Ac). This should allow a comparison between Mutter's cyclic peptide and D-Gal*p* derivatives as templates. Both structures **1** and **3** were assembled from the amphiphilic C-terminal hexadecapeptide aldehyde Ac-Glu-Ala-Leu-Glu-Lys-Ala-Leu-Lys-Glu-Ala-Leu-Ala-Lys-Leu-Gly-Gly-H (**6**). For compound **2**, the truncated dodecapeptide aldehyde Ac-Lys-Ala-Leu-Lys-Glu-Ala-Leu-Ala-Lys-Leu-Gly-Gly-H (**7**) was used. The peptide aldehydes were synthesized by a BAL strategy, starting with reductive amination of *o*-PALdehyde on 0.34 mmol/g Champion I (PEG-AMPS) resin with aminoacetaldehyde dimethyl acetal.[20–22] After coupling of the second residue as (Fmoc-Gly)$_2$O, the remaining synthesis was performed on an automated continuous-flow peptide synthesizer with 4 equiv of HBTU/HOBt activated Fmoc amino acid in 30 min couplings. Release of the peptide, dimethyl acetal hydrolysis, and deprotection of side-chain functionalities occurred by treatment with TFA/H_2O (19:1) for 2 h (Scheme 2). Purification of the two peptide aldehydes by prep. HPLC and lyophilization provided **6** in 21% yield and **7** in 15% yield based on an Fmoc quantification of the resin-bound dipeptide acetal.[23] In both cases ES–MS analysis showed dominating signals corresponding to $[M+3H]^{3+}$ and $[M–H_2O+2H]^{2+}$. While the former signal most likely originated from protonation of the three Lys residues in the sequences, we speculate that the latter signal resulted from imine formation between one of these Lys residues and the aldehyde functionality.

Chemoselective oxime ligation to form single-stranded carbopeptide **1** was achieved simply by dissolving hexadecapeptide aldehyde **6** and 2 equiv of mono-functionalized template **5** in 0.1 M aqueous acetate buffer at pH 4.76 (Scheme 3). After 2 h, **1** was isolated by prep. HPLC in 67% yield after lyophilization.[23] By a similar



Scheme 3. Synthesis of templated structures by chemoselective oxime ligation of C-terminal peptide aldehydes to Aoa functionalized methyl α-D-Gal*p* templates.

procedure, but now using 1.5 equiv of **6** for each aminooxy functionality in tetra-functionalized template **4**, carboprotein **3** was prepared in 76% yield, and with the same excess of dodecapeptide aldehyde **7** relative to template **4**, carbopeptide **2** was obtained in 67% yield. Although analytical HPLC of the purified products showed somewhat broad peaks, ES–MS indicated pure compounds with the multiple protonated species $[M+xH]^{x+}$ as the only significant signals. The putative derivatives of **2** and **3** missing one or two peptide strands were not observed. ^{1}H NMR (500 MHz, D_2O) showed for all three structures an approximate 1:1 ratio of *Z* and *E* oxime isomers according to the C*H*=N triplet signal from the *Z* isomer at 6.8 ppm and the *E* isomer at 7.5 ppm. The presence of *E*/*Z* oxime isomers could, in principle, give two^4 isomers of each of **2** and **3**.

Biophysical Characterization

The CD spectra of the three templated structures **1**, **2**, and **3** were recorded at 22 °C in 10 mM phosphate buffer at pH 7.0.[24] All three compounds showed some degree of α-helicity (Fig. 1, left). Based on $[\theta]_{222}$ and the formula proposed by Chen and co-workers,[25] the content of α-helix was calculated to be 45% in single-stranded carbopeptide **1**, 29% in the short-chained four-stranded **2**, and 67% in the 64 AA four-stranded carboprotein **3**. Dilution studies, in which the CD spectrum of single-stranded **1** was recorded at 0.91–91 μM concentrations indicated weak concentration dependence, with increasing concentration leading to increasing helicity. In contrast, the CD spectra of four-stranded **3** were concentration independent in the concentration range 2.7–27 μM, i.e., no secondary structural changes were observed. The CD spectrum of carboprotein **3** resembles the reported spectrum for Mutter's TASP T_4-($4\alpha_{15}$-Ac).[19] However, for the shorter TASP T_4-($4\alpha_{11}$), with a peptide sequence corresponding[26] to our **2**, Mutter and co-workers[19] reported a CD spectrum characteristic of a random coil, with a single minimum around 198 nm.

Denaturation experiments with guanidinium chloride (GuHCl) were carried out to compare the stabilities of single-stranded carbopeptide **1** and the corresponding four-stranded carboprotein **3**. A pronounced effect of inter-helical stabilization was observed, as **3** in contrast to **1** displayed a concerted denaturation curve with a midpoint at 5.1 M GuHCl (Fig. 1, right). From the data for carboprotein **3** the free energy of folding in water was calculated according to the linear extrapolation method, $\Delta G = \Delta G_{H_2O} + m[\text{GuHCl}]$.[27,28] The folding free energy was extrapolated to $\Delta G_{H_2O} = -6.9\,\text{kcal}\,\text{mol}^{-1}$, while the slope, reflecting the cooperativity of folding, was $m = 1.4\,\text{kcal}\,\text{mol}^{-1}\,M^{-1}$. These results are in accordance with data obtained for other templated 4-α-helix bundles by Sasaki and Kaiser[3] ($-4.4\,\text{kcal}\,\text{mol}^{-1}$; $0.84\,\text{kcal}\,\text{mol}^{-1}\,M^{-1}$) and Ghadiri and co-workers[5] ($-5.6\,\text{kcal}\,\text{mol}^{-1}$; $1.0\,\text{kcal}\,\text{mol}^{-1}\,M^{-1}$), but are not as high as the extraordinary result recently obtained by Mezo and Sherman[10] with a thiophenyl cavitand template ($-22.9\,\text{kcal}\,\text{mol}^{-1}$; $3.1\,\text{kcal}\,\text{mol}^{-1}\,M^{-1}$). Mutter and co-workers[19] reported a GuHCl denaturation curve for T_4-($4\alpha_{15}$) with a midpoint of 3 M, considerably less than we observe for the N-terminally acetylated carboprotein **3**.[26]

The structures of four-stranded **3** and the single-stranded reference **1** were further compared by NMR H–D exchange experiments to evaluate exposure of amide protons to solvent. The lyophilized samples were dissolved in D_2O, and the time course of exchange was followed by ^{1}H NMR spectroscopy. For carboprotein **3**, containing 64 amides, the first spectrum recorded after 5 min showed 35 exchangeable protons.[29] After 7 h, a total of four NH remained, while the last trace of exchangeable protons disappeared only after 42 h. For carbopeptide **1**, containing 16 amides, the spectrum after 5 min showed six exchangeable protons, and none after only 30 min. These results indicate a significant shielding of amide protons from the solvent in carboprotein **3**.

In conclusion, carboprotein **3** shows the characteristics of a 4-α-helix bundle, according to CD spectroscopy and H–D exchange experiments. Denaturation experiments indicate a higher stability for **3** than reported for TASP T_4-($4\alpha_{15}$), with which it shares the peptide sequence, and the shorter carbopeptide **2** shows a higher degree of α-helicity than reported for the corresponding TASP T_4-($4\alpha_{11}$).[19,26] Thus, on the basis of these initial

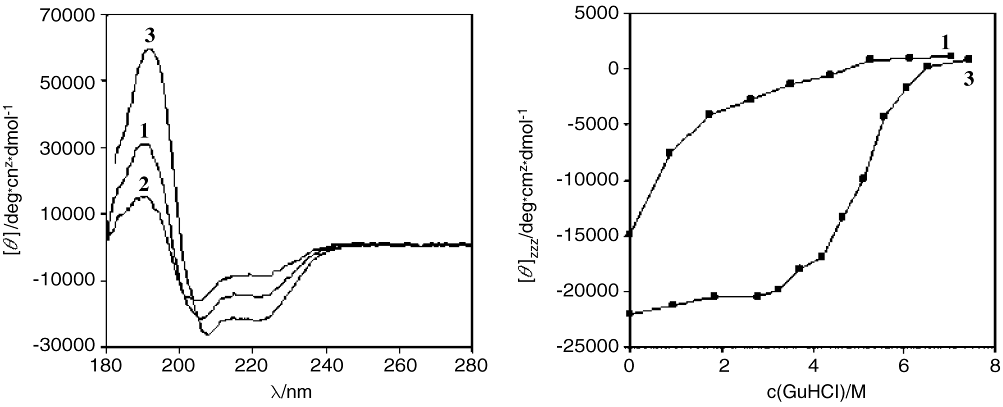


Figure 1. Left: CD spectra of **1**, **2**, and **3**. Recorded with 91 μM of **1**, 25 μM of **2**, and 27 μM of **3**. Right: GuHCl titration of **1** and **3**. Recorded with 91 μM of **1**, and 27 μM of **3**. Compound numbers are written on the spectra.

results it appears that the D-Gal*p* template and/or its linker region might be more apt in inducing 4-α-helix bundle formation than a cyclic peptide. Further investigations are required to establish whether the structure resembles that of a molten globule or in fact is native-like.

Acknowledgements

The authors thank Ms. Karen Jørgensen, University of Copenhagen, for providing the CD spectra, Dr. John Nielsen, Technical University of Denmark, for the ES–MS spectra, Professor Arne Holm, Royal Veterinary and Agricultural University, for donating a MilliGen 9050 PepSynthesizer, and Dr. Derek Hudson, Solid Phase Sciences, for a gift of Champion resins. Furthermore, we are grateful to the Alfred Benzon Foundation, the Lundbeck Foundation, the Torkil Holm Foundation, and the Danish Natural Science Research Council for financial support to KJJ.

References

1. DeGrado, W. F.; Summa, C. M.; Pavone, V.; Nastri, F.; Lombardi, A. *Annu. Rev. Biochem.* **1999**, *68*, 779.
2. Mutter, M.; Vuilleumier, S. *Angew. Chem., Int. Ed. Engl.* **1989**, *28*, 535.
3. Sasaki, T.; Kaiser, E. T. *Biopolymers* **1990**, *29*, 79.
4. Åkerfeldt, K. S.; Kim, R. M.; Camac, D.; Groves, J. T.; Lear, J. D.; DeGrado, W. F. *J. Am. Chem. Soc.* **1992**, *114*, 9656.
5. Ghadiri, M. R.; Soares, C.; Choi, C. *J. Am. Chem. Soc.* **1992**, *114*, 4000.
6. Goodman, M.; Feng, Y.; Melacini, G.; Taulane, J. P. *J. Am. Chem. Soc.* **1996**, *118*, 5156.
7. Tuchscherer, G.; Grell, D.; Mathieu, M.; Mutter, M. *J. Peptide Res.* **1999**, *54*, 185.
8. Rau, H. K.; DeJonge, N.; Haehnel, W. *Angew. Chem., Int. Ed.* **2000**, *39*, 250.
9. Wong, A. K.; Jacobsen, M. P.; Winzor, D. J.; Fairlie, D. P. *J. Am. Chem. Soc.* **1998**, *120*, 3836.
10. Mezo, A. R.; Sherman, J. C. *J. Am. Chem. Soc.* **1999**, *121*, 8983.
11. Jensen, K. J.; Barany, G. *J. Peptide Res.* **2000**, *56*, 3.
12. Brask, J.; Jensen, K. J. *J. Peptide Sci.* **2000**, *6*, 290.
13. Hristova-Kazmierski, M. K.; Horan, P.; Davis, P.; Yamamura, H. I.; Kramer, T.; Horvath, R.; Kazmierski, W. M.; Porreca, F.; Hruby, V. J. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 831.
14. Schaschke, N.; Musiol, H.-J.; Assfalg-Machleidt, I.; Machleidt, W.; Moroder, L. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 2507.
15. Péan, C.; Créminon, C.; Wijkhuisen, A.; Grassi, J.; Guenot, P.; Jéhan, P.; Dalbiez, J.-P.; Perly, B.; Djedaïni-Pilard, F. *J. Chem. Soc., Perkin Trans. 2* **2000**, 853.
16. Matsumura, S.; Sakamoto, S.; Ueno, A.; Mihara, H. *Chem. Eur. J.* **2000**, *6*, 1781.
17. Rose, K. *J. Am. Chem. Soc.* **1994**, *116*, 30.
18. Klemer, A.; Bieber, M.; Wilbers, H. *Liebigs Ann. Chem.* **1983**, 1416
19. 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.
20. Jensen, K. J.; Alsina, J.; Songster, M. F.; Vágner, J.; Albericio, F.; Barany, G. *J. Am. Chem. Soc.* **1998**, *120*, 5441.
21. Alsina, J.; Jensen, K. J.; Albericio, F.; Barany, G. *Chem. Eur. J.* **1999**, *5*, 2787.
22. Guillaumie, F.; Kappel, J. C.; Kelly, N. M.; Barany, G.; Jensen, K. J. *Tetrahedron Lett.* **2000**, *41*, 6131.
23. Yields were determined gravimetrically based on the assumption that all peptide Lys residues had formed TFA salts.
24. The mean residue ellipticity was calculated from the formula $[\theta]=\theta/(10\times l\times c\times n)$, where θ [mdeg] is the measured ellipticity, l [cm] is the path length, and n is the number of residues in each helix, not counting the C-terminal glycinal oxime. The concentration c [mol L^{-1}] was calculated from the mass of lyophilized carbopeptide or -protein used in preparation of a stock solution, assuming TFA salts as in ref 23.
25. Chen, Y.-H.; Yang, J. T.; Chau, K. H. *Biochemistry* **1974**, *13*, 3350.
26. TASPs T_4-($4\alpha_{11}$) and T_4-($4\alpha_{15}$) were not N-terminally acetylated. However, it was reported (ref 19) that *N*-acetylation, at least for T_4-($4\alpha_{15}$), only caused minor changes in the CD spectrum.
27. Pace, C. N. *Methods Enzymol.* **1986**, *131*, 266.
28. The calculations in the transition region 3.2 to 6.1 M GuHCl were based on the values $[\theta]_{222}=-20,000$ deg cm^2 dmol^{-1} for folded protein and $[\theta]_{222}=0$ deg cm^2 dmol^{-1} for denatured protein.
29. The number of remaining exchangeable protons was calculated from the amide integral, 7.6–8.6 ppm in the case of **3**, normalized to the template methoxy signal at 3.4 ppm.