Journal of Peptide Science
J. Peptide Sci. **6**: 290–299 (2000)

Carbopeptides: Chemoselective Ligation of Peptide Aldehydes to an Aminooxy-functionalized D-galactose Template

JESPER BRASK and KNUD J. JENSEN*

Department of Organic Chemistry, Technical University of Denmark, DK-2800 Lyngby, Denmark

Received 22 October 1999
Accepted 23 December 1999

Abstract: Multifunctional, topological template molecules such as linear and cyclic peptides have been used for the attachment of peptide strands to form novel protein models of, for example, 4-α-helix bundles. The concept of carbohydrates as templates for *de novo* design of potential protein models has been previously described and these novel chimeric compounds were termed *carbopeptides*. Here, a second generation strategy in which carbopeptides are synthesized by chemoselective ligation of a peptide aldehyde to an aminooxy-functionalized α-D-galactopyranoside is described. This template was prepared by per-*O*-acylation of methyl α-D-galactopyranoside with *N*,*N*-Boc_2-aminooxyacetic acid to form a tetra-functionalized template, followed by treatment with TFA-CH_2Cl_2 to release the aminooxy functionality. The peptide aldehydes Fmoc-Ser-Gly-Gly-H and H-Ala-Leu-Ala-Lys-Leu-Gly-Gly-H were synthesized by a BAL strategy. Four identical copies of peptide aldehyde were smoothly attached to the template by chemoselective ligation to form a 2.1 and a 2.9 kDa carbopeptide, respectively. Copyright © 2000 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: carbopeptide; *de novo* design; template-assembled synthetic protein; chemoselective oxime ligation; *C*-terminal peptide aldehyde; carbohydrate

INTRODUCTION

Design and synthesis of novel, artificial macromolecules with three-dimensional structures or biological functions similar to natural proteins, represents one of the most challenging goals of bio-organic chemistry. A good general understanding of the different factors governing protein folding and stability has been reached; even 'rules' or guidelines have been set forth for *de novo* design of common motifs in proteins such as the 4-α-helix bundle comprising of four α-helical and amphiphilic peptide segments [1–3]. However, when this understanding has been put to test by design of wholly new structures, ensembles of multiple conformations rather than unique structures have often been found [4].

To bypass the problematic folding step of linear polypeptides, Mutter and coworkers have suggested a template-assembled synthetic proteins (TASP) concept which relies on a multifunctional topological scaffold or template, which enhances the spatial accommodation of the peptides. The templates have included an octapeptide sequence, designed to form a short stretch of antiparallel β-sheet [5], a decapeptide sequence, cyclized by a disulfide bridge between a *C*- and a *N*-terminal Cys residue [6], as well as a cyclic decapeptide [7]. Other research groups have applied the TASP concept to

Abbreviations: Aoa, aminooxyacetyl; APCI, atmospheric pressure chemical ionization; BAL, backbone amide linker; D-Gal*p*, D-galactopyranose; DIPCDI, *N*,*N*′-diisopropylcarbodiimide; DIEA, *N*,*N*-diisopropylethylamine; *o*-PALdehyde, 5-(2-formyl-3,5-dimethoxyphenoxy)valeric acid; TASP, template-assembled synthetic protein; Amino acid symbols denote the L-configuration.

* Correspondence to: Department of Organic Chemistry, Building 201, Kemitorvet, Technical University of Denmark, DK-2800 Lyngby, Denmark; e-mail: okkj@pop.dtu.dk

non-peptide templates. Sasaki and Kaiser [8] reported the synthesis of 'Helichrome', a 4-α-helix bundle TASP based on a coproporphyrin template, DeGrado and coworkers [9] synthesized 'Tetraphilin', a 4-α-helix bundle proton channel, on a tetraphenylporphyrin template, and recently Sherman and coworkers [10] have reported the assembly of a 4-α-helix bundle on a cyclic tetra-thiophenol cavitand.

The development of techniques for chemoselective ligation of unprotected peptide segments has significantly extended the reach of synthetic peptide chemistry. Common to these methods is the chemoselective reaction of two mutually reactive functionalities, one on each segment. Some methods rely on the formation a non-natural (i.e. other than amide) bond at the site of ligation, this includes oxime and hydrazone ligations, developed by Rose [11], in which aminooxy or hydrazine nucleophiles react with aldehydes or ketones. Due to the α-effect [12] of the neighboring heteroatom, aminooxy derivatives are weak bases but reactive nucleophiles towards carbonyl groups at an acidic pH range of 4–5.5 where basic side-chain nucleophiles are protonated. With aldehydes and ketones oximes are formed which are stable under neutral to mildly acidic conditions [13]. Other chemoselective reactions are thioether and thioester ligations, based on reactions between a thiol or thioacid nucleophile and a bromoacetyl moiety. This has been further extended by Kent and coworkers into so-called 'native' chemoselective ligation in which a thiol capture reaction between a thioester and a *N*-terminal Cys residue results in a native amide bond [14,15]. Furthermore, Liu and Tam have demonstrated how a *N*-terminal Cys can be used for chemical ligation to a glycolaldehyde moiety forming a thiazolidine ring mimicking a Pro residue at the site of ligation [16].

A number of 4-α-helix bundle TASPs have been assembled via these chemical ligation strategies. Dawson and Kent functionalized Mutter's linear peptide template with bromoacetyl groups in a thioester ligation to peptides with a *C*-terminal Gly-$^{\alpha}$COSH residue [17]. Mutter's cyclic decapeptide template was applied by Rau and Haehnel in their recent synthesis of MOP cytochrome *b* models via thioether ligation. Peptides modified with a bromoacetyl moiety at either the *N*-terminal or at a *C*-terminal Lys residue were ligated to four Cys residues in the template sequence [18,19]. The oxime ligation strategy was applied by Tuchscherer and coworkers to a non-cyclic antiparallel β-sheet type template, functionalized with *N*-Aloc- and *N*-Boc-protected aminooxyacetic acid (Aoa-OH) [20]. An amino acid derivative with a diethyl acetal protected aldehyde moiety was incorporated at the *C*-terminal of one peptide sequence, and at the *N*-terminal of another, and an antiparallel 4-α-helix bundle was assembled in solution after acetal deprotection.

Carbohydrates are promising candidates for templates for the display of functional groups due to their inherent multifunctionality, the relative rigidity of ring forms, ease of regioselective chemical manipulations, and access to stereoisomers of mono- and disaccharides [21,22]. In previous reports the concept of carbohydrates as templates for *de novo* design of proteins was described (Jensen KJ, Barany G. Carbopeptides: carbohydrates as potential templates for *de novo* design of protein models, in press) [23]. Depending on the size, the names *carbopeptides* and *carboproteins* were suggested for the members of this novel class of chimeric protein models assembled on a carbohydrate template. Furthermore, the concept was demonstrated in a first-generation solid-phase strategy, in which D-galactopyranose (D-Gal*p*) was tetra-functionalized with Fmoc-βAla-OH and anchored to a solid support through a hydroxymethylbenzoyl aglycon linker (Jensen and Barany, in press) [23]. Simultaneous sequential peptide chain elongation on all four 'arms' of the support bound template proceeded in a reliable manner to yield a *carbopeptide* comprising the first segments of a 4-α-helix bundle.

Here an extension of the carbopeptides concept is presented to allow chemoselective ligation of *C*-terminal peptide aldehydes to an aminooxy-funtionalized methyl α-D-Gal*p* template by oxime formation in solution.

MATERIALS AND METHODS

General Procedures

DMF was kept over 4 Å molecular sieves. Pyridine was redistilled from CaH_2. High-loading aminomethylated PS resin, all amino acids, and HBTU were obtained from NovaBiochem, while HOBt was from Quantum Richelieu. Solid-phase reactions were carried out manually in polypropylene syringes with polyethylene filters. Melting points were measured in open capillary tubes. Merck TLC Aluminium Sheets Silica Gel 60 F_{254} plates were used for TLC. Spots were visualized by UV light, staining

with 10% sulfuric acid (only carbohydrate derivatives), or staining with a solution of 2% $Ce(SO_4)_2$ and 5% $H_3[P(Mo_3O_{10})_4]$ in 2 M H_2SO_4. A HP 8452A diode array spectrophotometer was used in Fmoc quantifications. For HPLC the following solvents were used: H_2O (A); CH_3CN (B); 0.1% TFA in H_2O (C); 0.1% TFA in CH_3CN (D). All programs used a linear gradient between time points and a flow rate of 1.2 ml/min (analytical), 20 ml/min (semi-prep., increasing from 0 ml/min over the first 5 min, over 1 min in condition 2), or 30 ml/min (prep., increasing from 0 ml/min over the first 5 min, over 1 min in condition 3 to 20 ml/min). Analytical HPLC was performed on a Waters system equipped with a 600E pump and a 996 PDA detector on a 3.9 × 150 mm Nova-Pak C18 4 μm 60 Å column. Semi-prep. and prep. HPLC was performed on a Waters system with a Delta 600 pump and a 996 PDA detector with one of two columns, semi-prep. HPLC on a 25 × 100 mm Prep Nova-Pak HR C18 6 μm 60 Å, prep. HPLC on a stack of three 40 × 100 mm Prep Nova-Pak HR C18 6 μm 60 Å units.

Analytical condition 1. Initial 100% A; at 1 min 100% A; at 3 min 80% A, 20% B; at 20 min 5% A, 95% B; at 25 min 5% A, 95% B.

Analytical condition 2. Initial: 100% C; at 1 min 100% C; at 15 min 50% C, 50% D; at 20 min 5% C, 95% D; at 25 min 5% C, 95% D.

Semi-prep. condition 1. Initial 100% A; at 5 min 100% A; at 15 min 80% A, 20% B; at 70 min 5% A, 95% B; at 75 min 5% A, 95% B.

Semi-prep. condition 2. Initial 100% C; at 1 min 100% C; at 50 min 50% C, 50% D; at 70 min 5% C, 95% D; at 75 min 5% C, 95% D.

Prep. condition 1. Initial 100% A; at 5 min 100% A; at 15 min 50% A, 50% B; at 60 min 100% B; at 75 min 100% B.

Prep. condition 2. Initial 100% A; at 5 min 100% A; at 15 min 80% A, 20% B; at 60 min 50% A, 50% B; at 70 min 5% A, 95% B; at 75 min 5% A, 95% B.

Prep. condition 3. As semi-prep. condition 2.

High resolution FAB MS were recorded by Dr Gustav Bojesen, University of Copenhagen, MALDI-TOF MS were recorded by Dr Frederik Beck, Pantheco A/S, while the ES MS spectrum of carbopeptide **12** was recorded by Dr Henrik Pedersen, H. Lundbeck A/S. All other MS spectra were recorded by Flemming Jensen, Phytera Inc, on an atmospheric pressure chemical ionization (APCI) LC-MS system. ^{1}H NMR spectra were recorded at 200 MHz on a Bruker AC-200 or at 500 MHz on a Varian Inova 500 spectrometer. ^{13}C NMR spectra were recorded at 50.3 MHz on a Brucker AC-200 or at 125.7 MHz on a Varia Inova 500 spectrometer.

Acylation of Methyl α-D-Gal*p* with N^{β}-Boc-Aoa-OH

Methyl α-D-Gal*p* monohydrate (**5**) (32 mg, 0.15 mmol) and N^{β}-Boc-Aoa-OH (**1**) (367 mg, 1.92 mmol) were dissolved in pyridine–CH_2Cl_2 (1:1, 2 ml), to which *N*,*N*′-diisopropylcarbodiimide (DIPCDI) (297 μl, 1.92 mmol) and a catalytic amount of DMAP were added. After reaction for 18 h at room temperature the suspension was concentrated to near dryness, redissolved in CH_2Cl_2, and filtered to remove *N*,*N*′-diisopropylurea. Vacuum liquid chromatography, eluting with EtOAc–hexane (1:3) and EtOAc–hexane (1:1), isolated a spot (TLC, EtOAc–hexane, 6:4, R_f 0.68). Concentration gave a colorless syrup, 96 mg. APCI LC-MS showed several peaks, dominated by the signals from methyl (N^{β}-Boc-Aoa)$_5$-α-D-Gal*p* (calc. for $C_{42}H_{69}N_5O_{26}$: 1059.42, found: *m*/*z* 1061 $[M+H]^+$, 961 $[M\text{-}Boc+2H]^+$, 861 $[M\text{-}2Boc+3H]^+$, 761 $[M\text{-}3Boc+4H]^+$, 661 $[M\text{-}4Boc+5H]^+$, 561 $[M\text{-}5Boc+6H]^+$) and from methyl (N^{β}-Boc-Aoa)$_6$-α-D-Gal*p* (calc. for $C_{49}H_{80}N_6O_{30}$ 1232.49, found *m*/*z* 1234 $[M+H]^+$, 1134 $[M\text{-}Boc+2H]^+$, 1034 $[M\text{-}2Boc+3H]^+$, 934 $[M\text{-}3Boc+4H]^+$, 834 $[M\text{-}4Boc+5H]^+$, 734 $[M\text{-}5Boc+6H]^+$, 634 $[M\text{-}6Boc+7H]^+$).

N^{β}-Boc-Aoa-OBn (2)

N^{β}-Boc-Aoa-OH (**1**) (6.8 g, 36 mmol) was dissolved in CH_3OH-H_2O (10:1, 44 ml). 20% (w/v) Cs_2CO_3 in water was added until pH 7 (~35 ml). Next, the solution was evaporated to dryness, the residue suspended in DMF (~20 ml), and again concentrated to dryness. This procedure was repeated to give the dry Cs-salt. Finally, the white powder was suspended in dry DMF (50 ml) and benzyl bromide (4.7 ml, 39 mmol) was added. After 18 h the clear solution was concentrated to near dryness and partitioned between water (50 ml) and Et_2O (2 × 50 ml). The ether fractions were pooled, dried ($MgSO_4$), and concentrated to yield 9.3 g of a white solid. The crude product was purified by trituration with pentane (30 ml) to yield 8.1 g, 80%, of title product, m.p. 80–82°C. ^{1}H NMR ($CDCl_3$, 200 MHz), δ: 1.48 (s, 9H, CH_3), 4.49 (s, 2H, CH_2 Aoa), 5.22 (s, 2H, benzylic), 7.38 (m, 5H, arom.), 7.76 (s, 1H, NH).

N,N-(Boc)$_2$-Aoa-OBn (3)

The benzyl ester (**2**) (562 mg, 2.0 mmol) was dissolved in CH_3CN (4 ml). Boc_2O (480 mg, 2.2 mmol) was added, followed by DMAP (24 mg, 0.2 mmol). After 10 min the evolution of CO_2 gas ceased and after 30 min TLC (toluene–CH_3CN, 4:1) showed no sign of starting material (R_f(**2**) 0.54, R_f(**3**) 0.62). The mixture was transferred to a separation funnel with EtOAc (10 ml) and extracted with 1 M aqueous $KHSO_4$ (5 ml), 1 M aqueous Na_2CO_3 (5 ml), and brine (5 ml), dried ($MgSO_4$) and concentrated to afford a faintly yellow oil, yield 728 mg, 96%, of title product. ^{1}H NMR ($CDCl_3$, 200 MHz), δ: 1.53 (s, 18H, CH_3), 4.57 (s, 2H, CH_2 Aoa), 5.23 (s, 2H, benzylic), 7.37 (m, 5H, arom.).

N,N-(Boc)$_2$-Aoa-OH (4)

The di-Boc protected benzyl ester (**3**) (2.34 g, 6.1 mmol) was dissolved in CH_3OH (20 ml), Pd/C (10%) catalyst (91 mg, 4% w/w) was added, and the mixture was hydrogenated at 1 atm. After 18 h TLC (toluene–CH_3CN, 4:1) showed only one spot with R_f 0.1. The catalyst was filtered off through a glass fiber filter and the colorless solution was evaporated to dryness. The residue was redissolved in EtOAc (1 ml). Addition of hexane (6 ml) caused some product to crystallize and precipitate over 30 min. The mixture was then left to evaporate and the resulting white solid was triturated with hexane ($\sim$10 ml) to yield 1.72 g, 96%, of the title product, m.p. 81–83°C. High-resolution FAB MS, calc. for $C_{12}H_{21}NO_7$: 291.1318. Found: m/z 292.1399 $[M+H]^+$. APCI LC-MS, found: m/z 581.4 [2M-H]$^-$, 481.2 [2M-Boc]$^-$, 290.0 [M-H]$^-$, 216.1 [M-OCH$_2$COOH]$^-$, 190.0 [M-Boc]$^-$. ^{1}H NMR (DMSO-d_6, 200 MHz), δ: 1.46 (s, 18H, CH_3), 4.44 (s, 2H, CH_2 Aoa). ^{13}C NMR (DMSO-d_6, 50.3 MHz), δ: 27.6 (CH_3), 72.1 (CH_2 Aoa), 83.5 ($C(CH_3)_3$), 149.6 (CO Boc), 168.7 (COOH).

Methyl 2,3,4,6-tetra-*O*-(*N,N*-(Boc)$_2$-Aoa)-α-D-Gal*p* (6)

Methyl α-D-Gal*p* monohydrate (**5**) (61 mg, 0.29 mmol) and *N,N*-(Boc)$_2$-Aoa-OH (**4**) (500 mg, 1.72 mmol) were dissolved in pyridine–CH_2Cl_2 (1:1, 5 ml) and stirred for 1 h with 3 Å molecular sieves. Then DMAP (21 mg, 0.17 mmol) and DIPCDI (266 µl, 1.72 mmol) were added. After 30 min, TLC (EtOAc–hexane, 1:1) showed several products, and additional DIPCDI (266 µl) was added. After further 30 min TLC (EtOAc–hexane, 1:1) showed only one spot (R_f 0.75). Molecular sieves were filtered off and the suspension containing *N,N*′-diisopropylurea was concentrated to dryness, redissolved in CH_3CN (3 ml), and purified by prep. HPLC (condition 1). The fractions containing the product, t_R 62 min, were identified from the UV absorption at 215 nm, pooled and concentrated to yield 312 mg, 84%, of title product. Dissolving the colorless syrup in a small amount of CH_2Cl_2, followed by rapid evaporation under high vacuum, gave a semi-crystalline foam. High-resolution FAB MS, calc. for $C_{55}H_{90}N_4O_{30}$: 1286.5640. Found: m/z 1309.5564 $[M+Na]^+$. APCI LC-MS, found: m/z 1346.2 $[M+H_2O+CH_3CN+H]^+$, 1305.1 $[M+H_2O+H]^+$, 1205.1 [M-Boc + H_2O + 2H]$^+$, 1104.9 [M-2Boc + H_2O + 3H]$^+$, 1005.0 [M-3Boc + H_2O + 4H]$^+$, 887.9 [M-4Boc + 5H]$^+$, 787.7 [M-5Boc + 6H]$^+$, 687.5 [M-6Boc + 7H]$^+$, 587.4 [M-7Boc + 8H]$^+$, 487.3 [M-8Boc + 9H]$^+$. ^{1}H NMR ($CDCl_3$, 500 MHz), δ: 1.52 (s, 72H, CH_3), 3.41 (s, 3H, OCH_3), 4.19 (dd, J 4.7 Hz, J 10.2 Hz, 1H, H-6a), 4.26 (m, 1H, H-5), 4.29 (dd, J 7.3 Hz, J 10.2 Hz, 1H, H-6b), 4.42–4.65 (m, 8H, CH_2 Aoa), 5.11 (d, J 3.8 Hz, 1H, H-1), 5.12 (dd, J 3.8 Hz, J 13.2 Hz, 1H, H-2), 5.51 (dd, J 3.4 Hz, J 13.2 Hz, 1H, H-3), 5.52 (m, 1H, H-4). ^{13}C NMR ($CDCl_3$, 125.7 MHz), δ: 28.0 (CH_3), 55.6 (OCH_3), 62.3 (C-6), 65.9 (C-5), 67.6 (C-2/3/4), 68.7 (C-2/3/4), 69.0 (C-2/3/4), 71.8–72.2 (CH_2 Aoa), 84.2–84.4 ($C(CH_3)_3$), 96.7 (C-1), 149.8–150.1 (CO Boc), 166.1–166.7 (CO Aoa).

Methyl 2,3,4,6-tetra-*O*-Aoa-α-D-Gal*p*, tetra TFA salt (7)

The protected template **6** (37 mg, 29 µmol) was dissolved in CH_2Cl_2-TFA (1:1, 2 ml). After stirring for 30 min at room temperature, it was concentrated to dryness, redissolved in a small amount of water ($\sim$5 ml) and lyophilized to afford 27 mg, 100%, of title product as a fine white, very hygroscopic powder. APCI LC-MS, calc. for $C_{15}H_{26}N_4O_{14}$: 486.14. Found: m/z 487.2 $[M+H]^+$. ^{1}H NMR (CD_3OD, 500 MHz), δ: 3.45 (s, 3H, OCH_3), 4.30–4.60 (m, 11H, H-5, H-6a, H-6b, CH_2 Aoa), 5.08 (d, J 3.8 Hz, 1H, H-1), 5.28 (dd, J 3.8 Hz, J 10.7 Hz, 1H, H-2), 5.55 (dd, J 3.4 Hz, J 10.7 Hz, 1H, H-3), 5.65 (dd, J 1.3 Hz, J 3.4 Hz, 1H, H-4).

Fmoc-Ser-Gly-Gly-H (8), and H-Ala-Leu-Ala-Lys-Leu-Gly-Gly-H (9)

High-loading aminomethylated PS resin was placed in a syringe, washed with DMF–*N,N*-diisopropylethylamine (DIEA) (19:1, 1 ×), DMF (3 ×), CH_2Cl_2 (3 ×), and DMF (3 ×). 5-(2-Formyl-3,5-dimethoxyphenoxy)valeric acid (*o*-PALdehyde) (2

equivalents), HBTU (2 equivalents), and HOBt (2 equivalents), were dissolved in DMF. DIEA (3 equivalents) was added and after stirring for 5 min, the solution was transferred to the resin with additional DMF. After 18 h on a shaker, the resin was washed with DMF (5 ×) and CH_2Cl_2 (5 ×), and dried overnight *in vacuo*. A subsequent Kaiser ninhydrin test was negative. The resin was capped with Ac_2O–CH_2Cl_2 (1:3) for 1 h, and then washed with CH_2Cl_2 (5 ×) and DMF (5 ×). Aminoacetaldehyde dimethyl acetal (10 equivalents) and $NaBH_3CN$ (10 equivalents) were dissolved in DMF and the solution was transferred to the PALdehyde–PS resin and placed on a shaker for 18 h. The following washes included DMF (5 ×) and CH_2Cl_2 (5 ×). Fmoc-Gly-OH (10 equivalents) was dissolved in CH_2Cl_2–DMF (9:1) and DIPCDI (5 equivalents) was added. Heavy precipitation resulted in the stirred solution after 3 min and after 15 min the white slurry was transferred to the resin. After 2 h the resin was washed with DMF (5 ×) and CH_2Cl_2 (5 ×). Anhydride coupling and washes were then repeated once. Finally, the resin was capped with Ac_2O–CH_2Cl_2 (1:3) for 1 h, washed with DMF (5 ×) and CH_2Cl_2 (5 ×), and dried overnight *in vacuo*. The loading on the resin was quantified from the absorption of dibenzofulvene at 290 nm after deprotection of a 5 mg sample. In the synthesis of peptide aldehyde **8** on a 1.20 mmol/g resin the loading of the first amino acid was quantified to 0.30 mmol/g. In the synthesis of peptide aldehyde **9** on a 1.00 mmol/g resin, the result was 0.40 mmol/g. Chain elongation proceeded with Fmoc-deprotection in piperidine–DMF (1:4) and coupling of Fmoc-protected amino acids with HBTU, following a protocol for automated Fmoc solid-phase peptide synthesis [24]. Protected amino acid side chains and the aldehyde moiety were deprotected and the product was released from the resin by treatment with TFA–H_2O (19:1) over 2 h. The resin was washed with additional TFA and the combined TFA solutions were concentrated to near dryness under a flow of air. The syrupy residue containing peptide aldehyde **8** was redissolved in CH_3CN and purified by prep. HPLC (condition 2). Retention time of the product containing peak was t_R 45 min. The relevant fractions were pooled, concentrated, and lyophilized to yield a white powder, 20 mg, 45%, of peptide aldehyde **8**. High-resolution FAB MS, calc. for $C_{22}H_{23}N_3O_6$: 425.1587. Found: m/z 426.1661 $[M+H]^+$. APCI LC-MS, found: m/z 466.0 $[M+H_2O+Na]^+$, 448.1 $[M+Na]^+$, 426.1 $[M+H]^+$. After cleavage from the handle, peptide aldehyde **9** was precipitated with cold Et_2O, centrifuged at 3000 rpm for 10 min, decanted, redissolved in H_2O, and purified by prep. HPLC (condition 3). Retention time was t_R 28 min. The product was obtained as a white powder after being lyophilized, 112 mg, 19%, of peptide aldehyde **9** (assuming salt formation with 2 × TFA). APCI LC-MS, calc. for $C_{28}H_{52}N_8O_7$: 612.4. Found 595.5 $[M-H_2O+H]^+$.

Methyl 2,3,4,6-tetra-*O*-Aoa-α-D-Gal*p*, tetra-(2,4-dimethoxybenzaldehyde) oxime (10)

Deprotected template **7** (5.0 mg, 5.3 μmol) was dissolved in CH_3CN–buffer (2:1, 1.5 ml), with the buffer being aqueous HOAc/NaOAc (0.50 M each) adjusted to pH 4 with concentrated HCl. 2,4-Dimethoxybenzaldehyde (5.3 mg, 31.9 μmol) was added and the mixture stirred while the reaction progress was followed by withdrawing samples for HPLC analysis (condition 1). Retention times were t_R(2,4-dimethoxybenzaldehyde) 13.28 min and t_R(**10**) 23.68 min. After 3 h the reaction was complete. The reaction mixture was then purified by semi-prep. HPLC (condition 1). The fractions around t_R(**10**) 58 min were concentrated to yield a solid white product, 5.2 mg, 91%, of title product. APCI LC-MS, calc. for $C_{51}H_{58}N_4O_{22}$: 1078.35. Found: m/z 1101.6 $[M+Na]^+$, 1079.6 $[M+H]^+$. ^1H NMR ($CDCl_3$, 500 MHz), δ: 3.32 (s, 3H, OCH_3 C-1), 3.73–3.96 (m, 24H, OCH_3 arom.), 4.19–4.31 (m, 3H, H-5, H-6a, H-6b), 4.54-4.80 (m, 8H, CH_2 Aoa), 5.05 (d, *J* 3.4 Hz, 1H, H-1), 5.23 (dd, *J* 3.4 Hz, *J* 10.9 Hz, 1H, H-2), 5.52 (dd, *J* 3.4 Hz, *J* 10.9 Hz, 1H, H-3), 5.62 Hz (m, 1H, H-4), 6.38–6.54 (m, 8H, arom.), 7.65–7.80 (m, 4H, arom.), 8.30–8.53 (m, 4H, oxime).

Methyl 2,3,4,6-tetra-*O*-Aoa-α-D-Gal*p*, tetra-(Fmoc-Ser-Gly-Gly-H) oxime (11)

The deprotected template **7** (6.6 mg, 7.0 μmol) was dissolved in the previously used aqueous buffer (0.5 ml). The peptide aldehyde **8** (18 mg, 42 μmol) was suspended in CH_3CN (1.5 ml), and the template solution was added. The suspension was left stirring while the reaction was followed by HPLC (condition 1). Retention times were t_R(**8**) 12.89 min and t_R(**11**) 19.49 min. Samples were filtered through a syringe filter prior to injection. After 72 h, the reaction mixture which still contained precipitate was evaporated to near dryness, dissolved in CH_3CN–DMF (5:1, 1.2 ml), and purified by semi-prep. HPLC (condition 1). The fractions around t_R 45 min were pooled, concentrated, and lyophilized to a white

solid in 13.8 mg, 93%, yield of title carbopeptide. MALDI-TOF MS, calc. for $C_{103}H_{110}N_{16}O_{34}$: 2114.74 (monoisotopic), 2116.06 (average). Found: m/z 2154.2 $[M + K]^+$, 2138.0 $[M + Na]^+$.

Methyl 2,3,4,6-tetra-*O*-Aoa-α-D-Gal*p*, tetra-(H-Ala-Leu-Ala-Lys-Leu-Gly-Gly-H) oxime (12)

Peptide aldehyde **9** (31 mg, 37 μmol, assuming TFA salt formation with two primary amines) and deprotected template **7** (5.8 mg, 6.2 μmol) were dissolved in the aqueous buffer (2.0 ml). The reaction was followed by HPLC (condition 2). Retention times were t_R(**9**) 12.58 min and t_R(**12**) 15.03 min. After 3 h the chromatograms indicated no further development and after 6 h the solution was purified by the semi-prep. HPLC (condition 2). The fractions around t_R 28 min were pooled, concentrated, and lyophilized to yield 15 mg white powder, 64% (assuming TFA salt formation with all eight primary amines), of title carbopeptide. MALDI-TOF MS, calc. for $C_{127}H_{226}N_{36}O_{38}$: 2863.69 (monoisotopic), 2865.37 (average). Found: m/z 2867.5 $[M + H]^+$. ES MS, found 956.1 $[M + 3H]^{3+}$, 717.3 $[M + 4H]^{4+}$, 574.1 $[M + 5H]^{5+}$.

RESULTS AND DISCUSSION

The new strategy developed included chemoselective ligation of *C*-terminal peptide aldehydes to an aminooxy-functionalized D-Gal*p* template. Recently, a new method was developed by one of the authors for the synthesis of *C*-terminal peptide aldehydes using a backbone amide linker (BAL) strategy [25]. The present authors envisioned accessing the template by per-*O*-acylation of commercially available methyl α-D-Gal*p* (**5**) with *tert*-butyloxycarbonyl (Boc) protected Aoa-OH followed by removal of Boc groups with TFA–CH_2Cl_2.

Preparation of Template

Starting from commercial available methyl α-D-Gal*p* (**5**), per-*O*-acylation with commercially available N^{β}-Boc-Aoa-OH (**1**) mediated by *N*,*N*′-diisopropylcarbodiimide (DIPCDI) in pyridine–CH_2Cl_2 (1:1) in the presence of *N*,*N*-dimethylaminopyridine (DMAP) yielded a mixture of compounds. LC-MS analysis showed the major products to be over-acylated (N^{β}-Boc-Aoa)$_5$-α-D-Gal*p* and (N^{β}-Boc-Aoa)$_6$-α-D-Gal*p* arising from *N*-acylation of the desired tetra-*O*-acyl template. Fragmentation products corresponding to the loss of one to five, and one to six, Boc-groups, were also identified. To prevent this side-reaction, *N*,*N*-(Boc)$_2$-Aoa-OH (**4**) was synthesized from **1** following precedents by Gunnarsson and Ragnarsson [26] for the *N*,*N*-di-protection of amino acids. First, **1** was converted to the corresponding benzyl ester derivative by neutralizing an aqueous solution of the acid with Cs_2CO_3, followed by reaction of the dry salt with benzyl bromide in DMF at room temperature. The second Boc group was added in a DMAP catalysed reaction with Boc_2O in acetonitrile at room temperature. Finally, the carboxylic acid moiety was deprotected by Pd/C-catalysed hydrogenation at atmospheric pressure and room temperature in methanol to afford *N*,*N*-(Boc)$_2$-Aoa-OH (**4**) in 74% overall yield from **1** (Figure 1). The di-Boc-protected acid **4** (6 equivalents) was smoothly coupled to methyl α-D-Gal*p* monohydrate (**5**) with DIPCDI (6 equivalents) in pyridine–CH_2Cl_2 (1:1) in the presence of DMAP and 3 Å molecular sieves at room temperature. Further addition of DIPCDI (6 equivalents) was required to drive the reaction to completeness. After 1 h total reaction time, methyl 2,3,4,6-tetra-*O*-(*N*,*N*-(Boc)$_2$-Aoa)-α-D-Gal*p* (**6**) was isolated in 84% yield. Deprotection with TFA–CH_2Cl_2 (1:1) proceeded quantitatively at room temperature to afford the desired template methyl 2,3,4,6-tetra-*O*-Aoa-α-D-Gal*p* (**7**) (Figure 2).

Preparation of *C*-terminal Peptide Aldehydes

A BAL strategy was used for solid-phase synthesis of Fmoc-protected *C*-terminal tripeptide aldehyde, N^{α}-Fmoc-Ser-Gly-Gly-H (**8**), and *C*-terminal heptapeptide aldehyde, H-Ala-Leu-Ala-Lys-Leu-Gly-Gly-H (**9**). First, *o*-PALdehyde was coupled to a high-loading aminomethylated polystyrene resin with HBTU in the presence of HOBt and DIEA. Next, reductive amination with aminoacetaldehyde dimethyl acetal and sodium cyanoborohydride in DMF–HOAc (99:1) afforded H-(BAL-PS)N-$CH_2CH(OCH_3)_2$. The incorporated nitrogen then became the backbone nitrogen of the *C*-terminal glycinal residue in the peptide assembled by chain elongation (Figure 3). Double coupling with

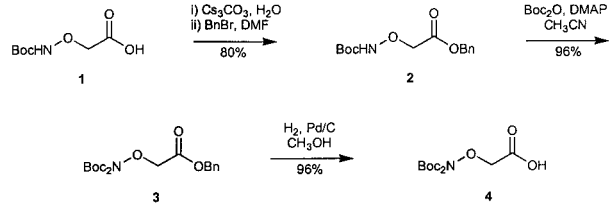

Figure 1 Preparation of *N*,*N*-(Boc)$_2$-Aoa-OH (**4**).

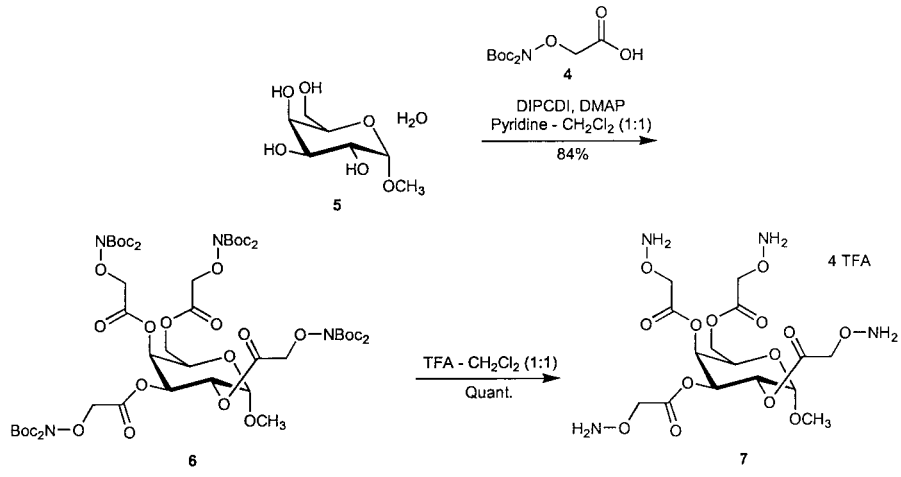

Figure 2 Preparation of methyl 2,3,4,6-tetra-*O*-Aoa-α-D-Gal*p* (**7**).

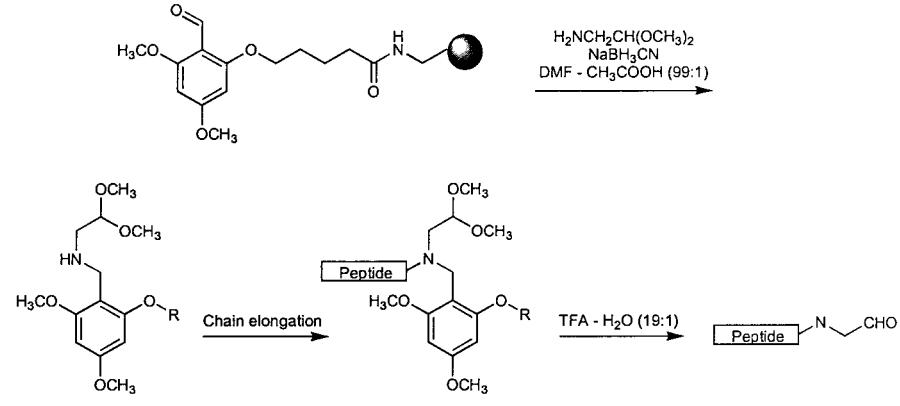

Figure 3 Preparation of peptide aldehydes by a BAL strategy.

(Fmoc-Gly)$_2$O in CH_2Cl_2–DMF (9:1) provided acylation of the secondary amine. The chain assembly was completed by sequential Fmoc-removal with piperidine–DMF (1:4) and coupling of Fmoc-protected amino acids with HBTU in the presence of HOBt and DIEA. Treatment with TFA–H_2O (19:1) for 2 h at room temperature released the peptide aldehyde with concomitant deprotection of the aldehyde and Ser or Lys side-chain moieties. After preparative HPLC, N^{α}-Fmoc-Ser-Gly-Gly-H (**8**) was obtained in 45% yield while H-Ala-Leu-Ala-Lys-Leu-Gly-Gly-H (**9**) was recovered in 19% yield.

Preparation of Carbopeptides by Oxime Ligation

In a model study, 2,4-dimethoxybenzaldehyde was coupled to methyl 2,3,4,6-tetra-*O*-Aoa-α-D-Gal*p* (**7**) at room temperature. Chemoselective oxime ligations of unprotected peptides have generally been performed in weakly acidic aqueous solutions [11,20,27,28]. However, due to the solubility properties of the aldehyde, and to facilitate HPLC-monitoring of the reaction progress, a solvent system of CH_3CN–buffer (2:1), with the buffer being aqueous acetate adjusted to pH 4, was chosen for this ligation. With 6 equivalents (50% excess) of the aldehyde, the reaction was complete within 2 h (Figure 4). The ligation product **10** was purified to give 91% isolated yield and characterized by APCI LC-MS and ^{1}H NMR. Next, N^{α}-Fmoc-Ser-Gly-Gly-H (**8**) (6 equivalents) was ligated to the template **7** in CH_3CN–buffer (3:1). The peptide aldehyde was only partially soluble in this solvent system and the reaction mixture remained a suspension. As in the model study, the reaction was essentially quantitative, though slower (Figure 6). After 72 h the desired 2.1 kDa tetra-substituted carbopeptide (**11**) was isolated by semi-prep. HPLC in 93% yield. It is likely that the prolonged reaction time is caused by the limited

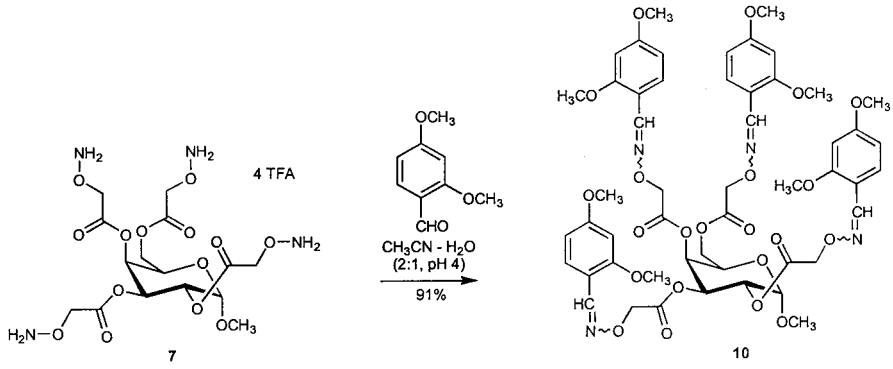

Figure 4 Preparation of methyl 2,3,4,6-tetra-*O*-Aoa-α-D-Gal*p*, tetra-(2,4-dimethoxybenzaldehyde) oxime (**10**). The wavy bonds indicate that both *cis* and *trans* forms are present.

solubility of the peptide aldehyde in the chosen solvent and that the use of, for example, DMF or DMSO as cosolvent, could reduce the reaction time. Furthermore, Shao and Tam have demonstrated that these solvents in general accelerate oxime ligation reactions [13]. Finally, H-Ala-Leu-Ala-Lys-Leu-Gly-Gly-H (**9**) (6 equivalents) was coupled to the template **7** in neat buffer. Both reagents were fully soluble and after 6 h the product, a 2.9 kDa carbopeptide (**12**), was isolated by semi-prep. HPLC in 64% yield (Figure 5).

NMR Studies

The authors speculated that the carbohydrate moieties of carbopeptides and -proteins could serve as structural markers for conformational changes caused by aggregating peptide chains, i.e. distortion of the pyranoside $^4C_1$ conformation[1] would reflect conformational changes in the peptide chains. $^3J_{H,H}$ coupling constants for the pyranoside ring in Boc-protected template **6** indicated the expected $^4C_1$ conformation ($^3J_{1,2}$ 3.8 Hz; $^3J_{2,3}$ 13.2 Hz). The four Aoa methylene groups showed some chemical shift dispersity. Coupling constants for deprotected template **7** also indicated a $^4C_1$ conformation ($^3J_{1,2}$ 3.8 Hz; $^3J_{2,3}$ 10.7 Hz). However, upon standing in CD_3OD for several hours additional peaks were observed in the ^1H NMR spectrum, possibly arising from loss of structural integrity of the template due to acyl transfer. This apparent slow change of the free template did not influence the chemoselective ligation (*vide supra*). In case of an intramolecular *N-O* acyl transfer to form Aoa_2-dimers, the modified template would only allow ligation at *three* positions. The dimethoxybenzylidene ligation product **10** showed the expected set of pyranoside protons indicative of a $^4C_1$ conformation ($^3J_{1,2}$ 3.4 Hz, $^3J_{2,3}$ 10.9 Hz). Due to the relatively short chain length of

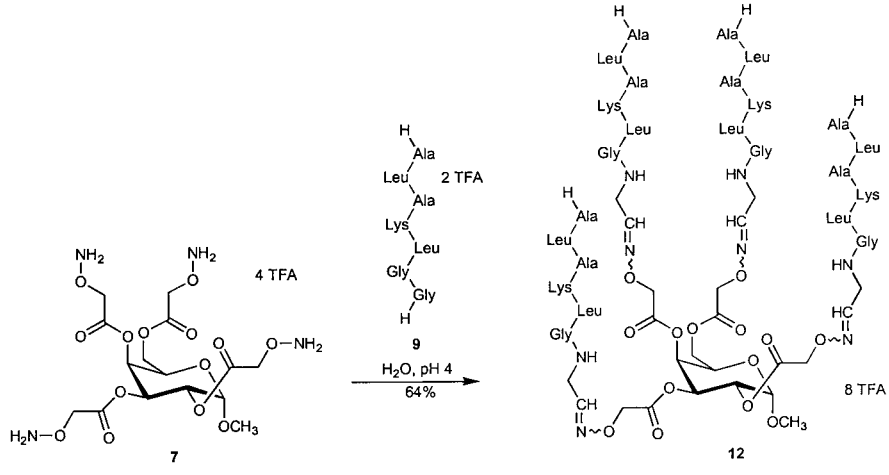

Figure 5 Preparation of methyl 2,3,4,6-tetra-*O*-Aoa-α-D-Gal*p*, tetra-(H-Ala-Leu-Ala-Lys-Leu-Gly-Gly-H) oxime (**12**).

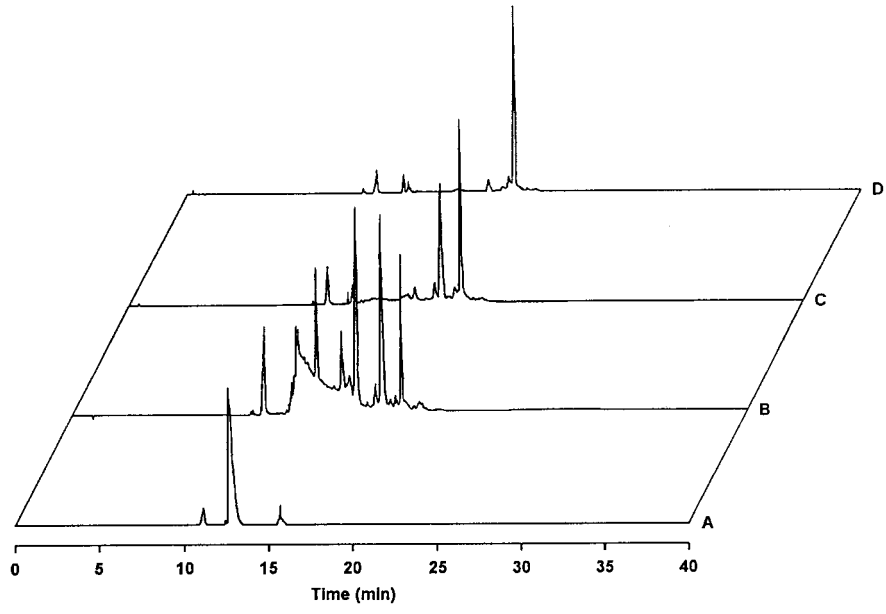

Figure 6 HPLC chromatograms showing formation of carbopeptide **11** by ligation of peptide aldehyde **8** to template **7**. A, initial; B, after 1 min; C, after 24 h; D, after 72 h. Retention times were t_R(**8**) 12.89 min and t_R(**11**) 19.49 min. Note that the *z*-axes of the chromatograms, showing the absorption at 265 nm, are not to scale.

six residues per 'arm', little secondary structure was expected in **12**. However, in H_2O–D_2O (9:1) and DMSO-d_6 solutions carbopeptide **12** showed promising chemical shift dispersion in the ^1H spectrum.

CONCLUSIONS

A new convergent strategy has been developed for the synthesis of carbopeptides in which *C*-terminal peptide aldehydes were chemoselectively ligated to an aminooxy-functionalized methyl α-D-Gal*p* to yield a novel class of protein models assembled on a carbohydrate template. Novel *N*,*N*-$(Boc)_2$-Aoa-OH was used in the per-*O*-acylation of methyl α-D-Gal*p* to give methyl 2,3,4,6-tetra-*O*-Aoa-α-D-Gal*p* after treatment with TFA–CH_2Cl_2 to cleave the Boc protecting groups. Four copies of *C*-terminal peptide aldehydes synthesized via a BAL strategy were simultaneously attached to the template by oxime formation in near-quantitative reactions. ^1H NMR studies of the template and the ligated product indicated a preference for the expected $^4C_1$ conformation. Future perspectives will include preparation and chemoselective ligation of longer amphiphilic α-helical *C*-terminal peptide aldehydes for *de novo* synthesis of 4-α-helix bundle models and the characterization of final ligated products by circular dichroism and ^1H NMR spectroscopy.

Acknowledgements

The authors thank Flemming Jensen (Phytera Inc), Dr Frederik Beck (Pantheco A/S), Dr Gustav Bojesen (University of Copenhagen), and Dr Henrik Pedersen (H. Lundbeck A/S) for providing mass spectra, and express gratitude to the Alfred Benzon Foundation (KJJ), the Danish Natural Science Research Council (KJJ), the Lundbeck Foundation (KJJ), and Novo Nordisk Research Foundation (stipend to JB) for financial support.

NOTES

1. $^4C_1$ denotes a chair conformation of the pyranoside ring in which C-4 and C-1, respectively, are above and below a plane defined by C-2, C-3, C-5, and O-5. See Reference [29].

REFERENCES

1. DeGrado WF, Wasserman ZR, Lear JD. Protein design, a minimalist approach. *Science* 1989; **243**: 622–628.

2. Betz SF, Bryson JW, DeGrado WF. Native-like and structually characterized designed α-helical bundles. *Curr. Opin. Struct. Biol.* 1995; **5**: 457–463.
3. Betz SF, DeGrado WF. Controlling topology and native-like behavior of *de novo*-designed peptides: design and characterization of antiparallel four-stranded coiled coils. *Biochemistry* 1996; **35**: 6955–6962.
4. Betz SF, Raleigh DP, DeGrado WF. *De-novo* protein design — from molten globules to native-like states. *Curr. Opin. Struct. Biol.* 1993; **3**: 601–610.
5. Mutter M, Altmann E, Altmann K-H, Hersperger R, Koziej P, Nebel K, Tuchscherer G, Vuilleumier S, Gremlich H-U, Müller K. The construction of new proteins. Part III. Artificial folding units by assembly of amphiphilic secondary structures on a template. *Helv. Chim. Acta* 1988; **71**: 835–847.
6. Mutter M, Tuchscherer G, Miller C, Altmann K-H, Carey RI, Wyss DF, Labhardt AM, Rivier JE. Template-assembled synthetic proteins with four-helix-bundle topology. Total chemical synthesis and conformational studies. *J. Am. Chem. Soc.* 1992; **114**: 1463–1470.
7. Dumy P, Eggleston IM, Cervigni S, Sila U, Sun X, Mutter M. A convenient synthesis of cyclic peptides as regioselectively addressable functionalized templates (RAFT). *Tetrahedron Lett.* 1995; **36**: 1255–1258.
8. Sasaki T, Kaiser ET. Helichrome: synthesis and enzymatic activity of a designed hemeprotein. *J. Am. Chem. Soc.* 1989; **111**: 380–381.
9. Åkerfeldt KS, Kim RM, Camac D, Groves JT, Lear JD, DeGrado WF. Tetraphilin: a four-helix proton channel built on a tetraphenylporphyrin framework. *J. Am. Chem. Soc.* 1992; **114**: 9656–9657.
10. Wallhorn D, Mezo AR, Sherman JC. Cavitand templated *de novo* four-helix bundlesPoster # 356. *The 16th American Peptide Symposium*, Minneapolis, MN, 1999.
11. Rose K. Facile synthesis of homogeneous artificial proteins. *J. Am. Chem. Soc.* 1994; **116**: 30–33.
12. Hudson RF. Störungstheoretische Behandlung der Chemischen Reaktivität. *Angew. Chem.* 1973; **85**: 63–84.
13. Shao J, Tam JP. Unprotected peptides as building blocks for the synthesis of peptide dendrimers with oxime, hydrazone, and thiazolidine linkages. *J. Am. Chem. Soc.* 1995; **117**: 3893–3899.
14. Dawson PE, Muir TW, Clark-Lewis I, Kent SBH. Synthesis of proteins by native chemical ligation. *Science* 1994; **266**: 776–779.
15. Canne LE, Bark SJ, Kent SBH. Extending the applicability of native chemical ligation. *J. Am. Chem. Soc.* 1996; **118**: 5891–5896.
16. Liu C-F, Tam JP. Chemical ligation approach to form a peptide bond between unprotected peptide segments. Concept and model study. *J. Am. Chem. Soc.* 1994; **116**: 4149–4153.
17. Dawson PE, Kent SBH. Convenient total synthesis of a 4-helix TASP molecule by chemoselective ligation. *J. Am. Chem. Soc.* 1993; **115**: 7263–7266.
18. Rau HK, Haehnel W. Design, synthesis, and properties of a novel cytochrome *b* model. *J. Am. Chem. Soc.* 1998; **120**: 468–476.
19. Rau HK, DeJonge N, Haehnel W. Modular synthesis of *de novo*-designed metalloproteins for light-induced electron transfer. *Proc. Natl. Acad. Sci. USA* 1998; **95**: 11526–11531.
20. Nyanguile O, Mutter M, Tuchscherer G. Synthesis of antiparallel 4α-helix bundle TASP by chemoselective ligation. *Lett. Pept. Sci.* 1994; **1**: 9–16.
21. Hirschmann R, Nicolaou KC, Pietranico S, Salvino J, Leahy EM, Sprengeler PA, Furst G, Smith AB, Strader CD, Cascieri MA, Candelore MR, Donaldson C, Vale W, Maechler L. Nonpeptidal peptidomimetics with a β-D-glucose scaffolding. A partial somatostatin agonist bearing a close structural relationship to a potent, selective substance P antagonist. *J. Am. Chem. Soc.* 1992; **114**: 9217–9218.
22. Hirschmann R, Nicolaou KC, Pietranico S, Leahy EM, Salvino J, Arison B, Cichy MA, Spoors PG, Shakespeare WC, Sprengeler PA, Hamley P, Smith AB, Reisine T, Raynor K, Maechler L, Donaldson C, Vale W, Freidinger RM, Cascieri MR, Strader CD. *De novo* design and synthesis of somatostatin non-peptide peptidomimetics utilizing β-D-glucose as a novel scaffolding. *J. Am. Chem. Soc.* 1993; **115**: 12550–12568.
23. Jensen KJ, Barany G. Carbohydrates as templates for *de novo* design of proteins. In *Peptides 1994. Proceedings of the Twenty-Third European Peptide Symposium, Braga, Portugal*, Maia HLS (ed.). Escom: Leiden, 1995; 443–444.
24. Fields GB, Tian Z, Barany G. Principles and practice of solid-phase peptide synthesis. In *Synthetic Peptides. A User's Guide. UWBC Biotechnical Resource Series*, Grant GA (ed.). Freeman and Co: New York, 1992.
25. Jensen KJ, Alsina J, Songster MF, Vágner J, Albericio F, Barany G. Backbone amide linker (BAL) strategy for solid-phase synthesis of *C*-terminal-modified and cyclic peptides. *J. Am. Chem. Soc.* 1998; **120**: 5441–5452.
26. Gunnarsson K, Ragnarsson U. Preparation and properties of N^{α}-di-*tert*-butoxycarbonyl amino acids. Applicability in the synthesis of Leu-enkephalin. *Acta Chem. Scand.* 1990; **44**: 944–951.
27. Tuchscherer G. Template assembled synthetic proteins: condensation of a multifunctional peptide to a topological template via chemoselective ligation. *Tetrahedron Lett.* 1993; **52**: 8419–8422.
28. Canne LE, Ferré-D'Amaré AR, Burley SK, Kent SBH. Total chemical synthesis of a unique transcription factor-related protein: cMyc-Max. *J. Am. Chem. Soc.* 1995; **117**: 2998–3007.
29. Collins PM, Ferrier RJ. *Monosaccharides. Their Chemistry and Their Roles in Natural Products*. John Wiley & Sons: Chichester, 1995.

