ELSEVIER

Carbohydrate Research 315 (1999) 334–338

CARBOHYDRATE RESEARCH

Note

Crystal structure of a fully protected β-*O*-galactosylated tripeptide

Marco Crisma *, Marina Gobbo, Claudio Toniolo, Raniero Rocchi

Biopolymer Research Centre, C.N.R., Department of Organic Chemistry, University of Padua, via Marzolo 1, I-35131 Padua, Italy

Received 20 October 1998; revised 5 January 1999; accepted 5 January 1999

Abstract

The crystal structure of the fully protected glycotripeptide *N*-benzyloxycarbonyl-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-L-threonyl-α-aminoisobutyryl-α-aminoisobutyric acid *tert*-butyl ester [Z-(β-D-GalAc$_4$)-L-Thr-Aib-Aib-O*t*Bu] has been determined by X-ray diffraction. The peptide backbone is fully extended at Thr(1), left-handed helical at Aib(2), while it is right-handed helical at Aib(3). The fully extended conformation at the N-terminal Thr residue is stabilized by an intramolecular H-bond involving the N-1–H and O-1=C-1 groups (intramolecularly H-bonded C_5 conformation). Two additional intramolecular H-bonds are observed, involving the peptide N–H groups of Aib(2) and Aib(3) as the donors, and the pyranosidic O-5 oxygen atom and the carbonyl oxygen atom of the acetoxy group on C-6 of the sugar, respectively, as the acceptors. Owing to the peptide–sugar H-bonds, the peptide backbone is forced to adopt a conformation dramatically different from the β-bend/3_{10}-helical conformation usually observed for Aib-rich peptides. The implications and limitations of these findings on the effect of *O*-glycosylation on the conformation of natural peptides are briefly outlined.

Keywords: Glycopeptides; X-ray diffraction analysis; α-Aminoisobutyric acid

A variety of experimental data indicate that C$^\alpha$-tetrasubstituted α-amino acids impart well-defined and predictable conformational constraints upon the peptide backbone. In particular, the prototypical α-aminoisobutyric acid (Aib) is a strong promoter and stabilizer of folded (β-bends) and 3_{10}/α-helical structures, as observed in the crystal structures of Aib-containing peptides composed of 3–20 residues [1,2].

To address the effect of *O*-glycosylation on the secondary structure of a peptide with well-defined conformational preferences, we synthesized a series of *O*-glycosylated Aib-rich peptides of general sequence Z-[(β-D-Gal)-L-Thr-Aib-Aib]$_n$-O*t*Bu, in analogy to a series of glycosylated polytripeptides [(β-D-Gal)-L-Thr-L-Ala-L-Ala]$_n$ ($n = 2–7$) previously studied in our laboratory [3] as models of antifreeze glycoproteins (AFGPs) of antarctic fishes. In the course of our investigation, we were able to grow a single crystal of the glycotripeptide Z-(β-D-GalAc$_4$)-L-Thr-Aib-Aib-O*t*Bu (**1**), the smallest term of the series, suitable for an X-ray diffraction analysis.

Abbreviations: Ac, acetyl; *p*BrBz, *para*-bromobenzoyl; *t*Bu, *tert*-butyl; Dab, α,γ-diaminobutyric acid; Dap, α,β-diaminopropionic acid; Z, benzyloxycarbonyl.

* Corresponding author. Tel.: +39-49-8275294; fax: +39-49-8275239.

E-mail address: biop02@chor.unipd.it (M. Crisma)

0008-6215/99/$ - see front matter
PII: S0008-6215(99)00010-5

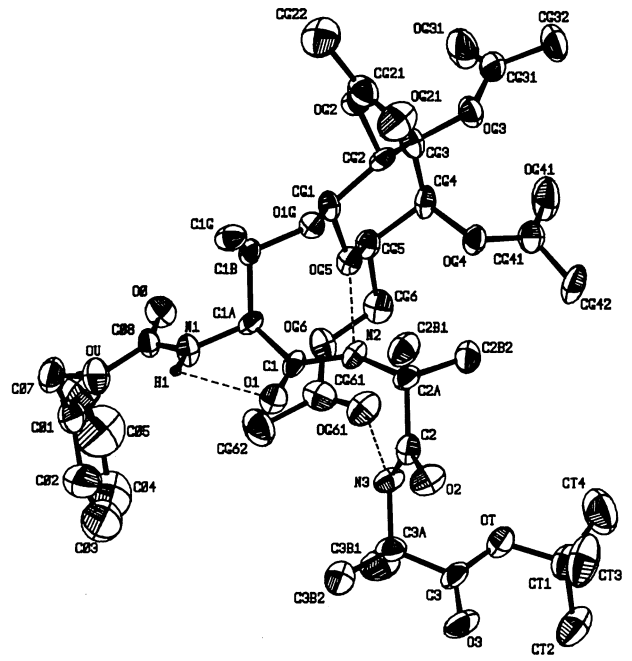
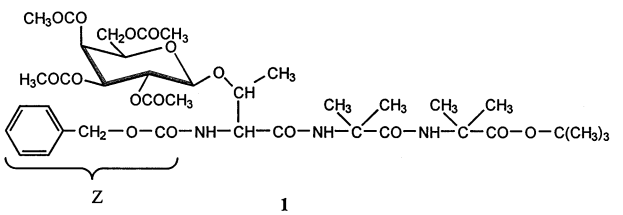


Fig. 1. ORTEP view of the X-ray diffraction structure of Z-(β-D-GalAc$_4$)-L-Thr-Aib-Aib-O*t*Bu with atom numbering. Intramolecular H-bonds are represented by dashed lines. Thermal ellipsoids are drawn at the 30% probability level.



1

In this communication, we describe the crystal structure of this compound which represents, to our knowledge, the first example of an *O*-glycopeptide studied by X-ray diffraction analysis. A conformational analysis in solution (by NMR and FTIR absorption) is underway on this as well on the other *O*-glycosylated, Aib-rich peptides synthesized. The results will be reported in due course. Previously, the X-ray diffraction structures of only two *O*-glycosylated amino acids, O^3-(β-D-xylopyranosyl)-L-serine [4] and O^3-(β-D-mannopyranosyl)-L-threonine [5], were reported.

The X-ray diffraction structure of **1**, with atom numbering, is illustrated in Fig. 1. Relevant torsion angles are listed in Table 1 (see supplementary material). Intra- and intermolecular H-bond parameters are listed in Table 2. Bond distances and bond angles (deposited) are in general agreement with previously reported values for the geometry of the benzyloxycarbonylamino moiety [6], the peptide unit [7,8], the ester group [9], and the glycopyranose ring [10].

All of the urethane, peptide and ester bonds are trans. However, significant deviations from the trans-planarity ($|\Delta\omega| > 10°$) are observed for the Thr(1)–Aib(2) (ω_1) peptide bond, and for the acetyl ester bond at position 2 in the sugar ring.

Table 1
Selected torsion angles (°) for Z-(β-D-GalAc$_4$)-L-Thr-Aib-Aib-O*t*Bu

C-01–C-07–O-U–C-08	75.3(8)	C-G5–O-G5–C-G1–O-1G	179.6(5)
C-07–O-U–C-08–N-1	−176.9(6)	O-1G–C-G1–C-G2–C-G3	171.7(5)
O-U–C-08–N-1–C-1A (ω_0)	174.4(5)	O-G5–C-G1–C-G2–C-G3	54.3(8)
C-08–N-1–C-1A–C-1 (ϕ_1)	−151.1(6)	C-G1–C-G2–C-G3–C-G4	−47.4(8)
N-1–C-1A–C-1B–O-1G ($\chi_1^{1,1}$)	−177.2(5)	C-G2–C-G3–C-G4–C-G5	48.5(9)
N-1–C-1A–C-1B–C-1G ($\chi_1^{1,2}$)	64.4(8)	C-G3–C-G4–C-G5–O-G5	−56.7(8)
N-1–C-1A–C-1–N-2 (ψ_1)	162.9(5)	C-G1–O-G5–C-G5–C-G4	65.8(7)
C-1A–C-1–N-2–C-2A (ω_1)	168.5(6)	C-G5–O-G5–C-G1–C-G2	−63.6(8)
C-1–N-2–C-2A–C-2 (ϕ_2)	63.3(9)	C-G1–C-G2–O-G2–C-G21	101.2(7)
N-2–C-2A–C-2–N-3 (ψ_2)	34.2(9)	C-G2–C-G3–O-G3–C-G31	−149.4(6)
C-2A–C-2–N-3–C-3A (ω_2)	176.9(6)	C-G3–C-G4–O-G4–C-G41	−103.0(7)
C-2–N-3–C-3A–C-3 (ϕ_3)	−50.1(10)	C-G4–C-G5–C-G6–O-G6	−179.6(6)
N-3–C-3A–C-3–O-T (ψ_3)	−45.5(9)	C-G2–O-G2–C-G21–O-G21	10.7(12)
C-3A–C-3–O-T–C-T1 (ω_3)	−173.4(6)	C-G2–O-G2–C-G21–C-G22	−168.3(6)
C-1B–O-1G–C-G1–C-G2 (Φ)	159.9(5)	C-G3–O-G3–C-G31–O-G31	5.9(12)
C-1B–O-1G–C-G1–O-G5	−81.2(6)	C-G3–O-G3–C-G31–C-G32	−177.2(6)
C-G1–O-1G–C-1B–C-1G (Ψ_1)	−152.6(6)	C-G4–O-G4–C-G41–O-G41	8.0(12)
C-G1–O-1G–C-1B–C-1A	82.1(7)	C-G4–O-G4–C-G41–C-G42	−171.8(7)
		C-G5–C-G6–O-G6–C-G61	−102.3(8)
		C-G6–O-G6–C-G61–C-G62	178.5(6)

Table 2
Intra- and intermolecular H-bond parameters for Z-(β-D-GalAc$_4$)-L-Thr-Aib-Aib-O*t*Bu

Donor (D–H)	Acceptor (A)	Distance (Å)		Angle (D–H···A) (°)	Symmetry equivalence of A
		D···A	H···A		
Intramolecular					
N-1–H	O-1	2.650(8)	2.35	101	*x*, *y*, *z*
N-2–H	O-G5	3.018(8)	2.21	156	*x*, *y*, *z*
N-3–H	O-G61	3.018(9)	2.18	165	*x*, *y*, *z*
Intermolecular					
N-1–H	O-G2	3.330(8)	2.51	159	$-1/2+x$, $1/2-y$, $-z$

The peptide backbone [11] is fully extended (ϕ_1, ψ_1) at Thr(1), left-handed helical (ϕ_2, ψ_2) at Aib(2), while right-handed helical ('ϕ_3', 'ψ_3') at Aib(3). The fully extended conformation at the N-terminal Thr residue is stabilized by an intramolecular H-bond involving the N-1–H and O-1=C-1 groups (intramolecularly H-bonded C_5 conformation) [12]. Two additional intramolecular H-bonds are observed, involving the peptide N-2–H and N-3–H groups as the donors, and the pyranosidic O-G5 oxygen atom and the O-G61 carbonyl oxygen atom of the acetoxy group at C-6 of the sugar, respectively, as the acceptors.

The substituted L-Thr side chain adopts the energetically unfavourable trans, gauche+ conformation [8,13] about the $\chi_1^{1,1}$ and $\chi_1^{1,2}$ torsion angles, respectively. It is reasonable to assume that the loss of energy of such an unusual conformation would be, at least in part, counterbalanced by formation of the (peptide) N-2–H···O-G5 (pyranose) intramolecular H-bond.

The β-D-galactopyranose ring is close to the $^4C_1$ chair conformation, with Cremer–Pople puckering parameters [14] $Q_T = 0.557(7)$ Å, $\Theta = 171.0(7)°$, and $\phi = 104(5)°$ (for the atom sequence C-G1–C-G2–C-G3–C-G4–C-G5–O-G5). The C-G4 and C-G1 atoms are displaced by $-0.651(7)$ Å and $+0.674(7)$ Å, respectively, from the average plane defined by the C-G2, C-G3, C-G5, and O-G5 atoms.

The conformation about the glycopyranosidic linkage is described by the values of 159.9(5) and $-152.6(6)°$ adopted by the torsion angles Φ (C-1B–O-1G–C-G1–C-G2) and Ψ_1 (C-1G–C-1B–O-1G–C-G1), respectively. Such values are consistent with a relevant exo-anomeric effect, being not far from the theoretical value of $\pm 180°$ [15].

The carbonyl oxygen atoms of the secondary acetate groups are nearly eclipsed to the axial H-atom of the corresponding ring C-atom, as previously reported for other peracetylated glycopyranose rings [10]. As a consequence, short intramolecular O···H contacts (in the range 2.30–2.32 Å) are observed. The disposition of the primary acetate moiety is defined by the trans orientation of the O-G6 and C-G4 atoms about the C-G5–C-G6 bond, and by the anticlinal disposition of the C-G61 and C-G5 atoms about the C-G6–O-G6 bond. This latter finding is at variance with the trans conformation most often found about the C(primary)–O(acetate) bond, which minimizes steric effects [16,17]. In the present case such discrepancy may be ascribed to the involvement of the carbonyl oxygen atom O-G61 in the H-bond with the peptide N-3–H group.

The crystal packing of Z-(β-D-GalAc$_4$)-L-Thr-Aib-Aib-O*t*Bu is characterized by a single, weak [18] intermolecular H-bond involving the urethane N-1–H group and the O-G2 oxygen atom of the acetoxy group on C-2 of the sugar of a symmetry-related molecule (Table 2), generating rows of molecules along the crystallographic two-fold screw axis parallel to the *a* direction. Such an interaction may account for the distortion from the trans planarity observed for this ester bond. Packing is then completed through van der Waals interactions.

Linear, terminally protected Aib-rich oligopeptides, made of three to six residues, usually fold into a single β-bend or multiple, consecutive β-bends depending upon main-chain length [1,2]. The crystal structures of Z-L-Dap(*p*BrBz)-(Aib)$_2$-NHCH$_3$ and Z-L-Dab(*p*BrBz)-(Aib)$_2$-NHCH$_3$ indicate that such folding may be preserved even when a side chain carries potential H-bonding donor and acceptor groups [19]. On the other hand, the results of the present work show that the oxygen atoms of the peracetylated sugar moiety of the (β-D-GalAc$_4$)-L-Thr residue effectively compete with the peptide carbonyl groups as acceptors of intramolecular H-bonds. As a result, this *O*-glycosylated residue dramatically alters the peptide backbone conformation. However, we suggest that caution should be exercised on whether this conclusion might also apply to non-peracetylated, *O*-glycosylated peptides, and hence on its biological relevance. In any case, it has to be noted that the intramolecular H-bond between the N-2–H group of the Aib residue next to the glycosylated Thr and the pyranosidic O-G5 oxygen atom observed in the present work may in principle occur also when the sugar hydroxyl groups are not protected.

Table 3
Crystal data and structure refinement for Z-(β-D-GalAc$_4$)-L-Thr-Aib-Aib-O*t*Bu

Molecular formula	$C_{38}H_{55}N_3O_{16}$
Molecular weight (amu)	809.85
Temperature (K)	293(2)
Wavelength (Å)	0.71073
Crystal system	orthorhombic
Space group	$P2_12_12_1$
Unit cell dimensions (Å)	$a = 10.462(2)$ $b = 15.037(2)$ $c = 27.154(3)$
Volume (Å^3)	4271.8(11)
Z (molecules per cell)	4
D_{calcd} (g cm^{-3})	1.259
Absorption coefficient (mm^{-1})	0.098
$F(000)$	1728
Crystal size (mm)	$0.40 \times 0.20 \times 0.15$
θ Range for data collection (°)	2.0–22.5
Index ranges	$0 \le h \le 13$; $0 \le k \le 19$; $0 \le l \le 35$
Reflections collected	3171
Independent reflections	3168
Reflections with $I \ge 2\sigma(I)$	1487
Refinement method	full-matrix-block least-squares on F^2
Function minimized	$\Sigma w(F_o^2 - F_c^2)^2$, $w = 1/[\sigma^2(F_o^2) + (0.0643P)^2]$, where $P = (F_o^2 + 2F_c^2)/3$
Data/restraints/parameters	3164/0/502
Goodness-of-fit on F^2	0.884
Final R indices [$I \ge 2\sigma(I)$]	R_1 (on F) = 0.0472, wR_2 (on F^2) = 0.0973
R indices (all data)	R_1 (on F) = 0.1746, wR_2 (on F^2) = 0.1794
Largest difference peak and hole (e Å^{-3})	0.230 and −0.276

1. Experimental

X-ray diffraction structure determination.— Suitable single crystals of Z-(β-D-GalAc$_4$)-L-Thr-Aib-Aib-O*t*Bu were grown from ethanolic solution by slow evaporation. The crystal data and a summary of experimental details are given in Table 3. Data collection was performed on a Philips PW1100 four-circle diffractometer, with graphite monochromated Mo Kα radiation, in the θ–2θ scan mode. The crystal did not significantly diffract above $\theta = 22.5°$ (0.93 Å resolution). Cell parameters were determined by least-squares refinement of the angular settings of 25 carefully centred reflections with $\theta = 7$–$12°$. The structure was solved by direct methods using the SHELXS-86 [20] program, and refined by full-matrix-block least-squares on F^2 using all data (except four low-angle reflections which were in part cut by the beamstop) and the SHELXL-93 [21] program, allowing the positional parameters and the anisotropic displacement parameters of the non-H atoms to refine at alternate cycles. All non-H atoms were refined anisotropically. The phenyl ring of the benzyloxycarbonyl N-protecting group was constrained to the idealized geometry. The H-atoms were calculated at idealized positions, and during the refinement they were allowed to ride on their parent atom, with U_{iso} set equal to 1.2 (or 1.5 for methyl groups) times the U_{eq} of the carrying atom. Geometrical calculations were performed with the PARST [22] program.

2. Supplementary material

Tables of atomic coordinates, anisotropic displacement parameters, bond lengths, and bond angles have been deposited with the Cambridge Crystallographic Data Centre. These tables may be obtained, on request, from the Director, Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, CB2 1EZ, UK.

References

[1] C. Toniolo, E. Benedetti, *Macromolecules*, 24 (1991) 4004–4009.
[2] I. L. Karle, P. Balaram, *Biochemistry*, 29 (1990) 6747–6756.
[3] F. Filira, L. Biondi, B. Scolaro, M. T. Foffani, S. Mammi, E. Peggion, R. Rocchi, *Int. J. Biol. Macromol.*, 12 (1990) 41–48.
[4] L. T. J. Delbaere, B. Kamenar, K. Prout, *Acta Crystallogr., Sect. B*, 31 (1975) 862–865.
[5] N. Darbon, Y. Oddon, J. M. Lacombe, A. A. Pavia, *Carbohydr. Res.*, 130 (1984) 55–64.
[6] E. Benedetti, C. Pedone, C. Toniolo, M. Dudek, G. Némethy, H.A. Scheraga, *Int. J. Pept. Protein Res.*, 21 (1983) 163–181.
[7] E. Benedetti, in B. Weinstein (Ed.), *Chemistry and Biochemistry of Amino Acids, Peptides and Proteins*, Vol. 6, Dekker, New York, 1982, pp. 105–184.
[8] T. Ashida, Y. Tsunogae, I. Tanaka, T. Yamane, *Acta Crystallogr., Sect. B*, 43 (1987) 212–218.
[9] W.B. Schweizer, J.D. Dunitz, *Helv. Chim. Acta*, 65 (1982) 1547–1554.
[10] J. Ohanessian, F. Longchambon, F. Arene, *Acta Crystallogr., Sect. B*, 34 (1978) 3666–3671.
[11] IUPAC–IUB Commission on Biochemical Nomenclature, *J. Mol. Biol.*, 52 (1970) 1–17.
[12] C. Toniolo, E. Benedetti, in P. Balaram, S. Ramaseshan (Eds.), *Molecular Conformation and Biological Interactions*, Indian Academy of Sciences, Bangalore, India, 1991, pp. 511–521.
[13] E. Benedetti, G. Morelli, G. Némethy, H. A. Scheraga, *Int. J. Pept. Protein Res.*, 22 (1983) 1–15.
[14] D. Cremer, J.A. Pople, *J. Am. Chem. Soc.*, 97 (1975) 1354–1358.
[15] R. U. Lemieux, S. Koto, D. Voisin, *ACS Symp. Ser.*, 87 (1979) 17–29.
[16] F. Leung, R.H. Marchessault, *Can. J. Chem.*, 51 (1973) 1215–1222.
[17] J.D. Olivier, L.C. Strickland, *Acta Crystallogr., Sect. C*, 40 (1984) 820–824.
[18] C.H. Görbitz, *Acta Crystallogr., Sect. B*, 45 (1989) 390–395.
[19] M. Crisma, G. Valle, F. Formaggio, A. Bianco, C. Toniolo, *J. Chem. Soc., Perkin Trans. 2*, (1993) 987–981.
[20] G.M. Sheldrick, SHELXS-86: *Program for the Solution of Crystal Structures*, University of Göttingen, Göttingen, Germany, 1986.
[21] G.M. Sheldrick, SHELXL-93: *Program for Crystal Structure Refinement*, University of Göttingen, Göttingen, Germany, 1993.
[22] M. Nardelli, *J. Appl. Crystallogr.*, 28 (1995) 659.