Bioorganic & Medicinal Chemistry Letters, Vol.3, No.12, pp. 2515-2520, 1993
Printed in Great Britain

0960-894X/93 $6.00 + .00

ENZYMATIC SYNTHESIS OF DISACCHARIDE-SERINE AND PEPTIDE CONJUGATES

Sylvie BAY, Abdelkadder NAMANE, Danièle CANTACUZENE

Institut Pasteur, Département de Biochimie et Génétique Moléculaire, Unité de Chimie Organique, associée au CNRS, 28 Rue du Dr. Roux, 75724 Paris cedex 15 (France)

(*Received in Belgium* 29 *June* 1993; *accepted* 1 *September* 1993)

Abstract : *We describe enzymatic transglycosylations between an appropriate glycosyl donor and galactosyl (or glucosyl)-serine and -peptide conjugates to obtain diglycosyl-serine or -peptide derivatives. The reactions are catalyzed by β-galactosidase (from E. coli or from Aspergillus oryzae) and β-glucosidase (from Almonds). The enzymatic reactions give, preferentially, β(1→6) linked diglycosyl-serine (or -peptide) conjugates. However, in the case of the digalactosyl derivatives, β(1→3) linkages are mainly observed. By changing the source of the enzyme (E. coli or Aspergillus oryzae) the regioselectivity can be reversed for these digalactosyl derivatives. Deprotection of the aminoacid of the diglycosyl-peptides under mild conditions is also described.*

The important role played by glycoproteins in biological systems has increased the interest and demand for a variety of glycopeptides as model compounds[1,2]. Although much progress has been made in the classical chemical methods[3,4], numerous protection and deprotection steps are required for a regioselective synthesis of glycopeptides. In addition, stereospecific reactions are often difficult and the use of glycosidases[5] (or glycosyltransferases[6]) could be, in some cases, an alternative to the chemical approach.

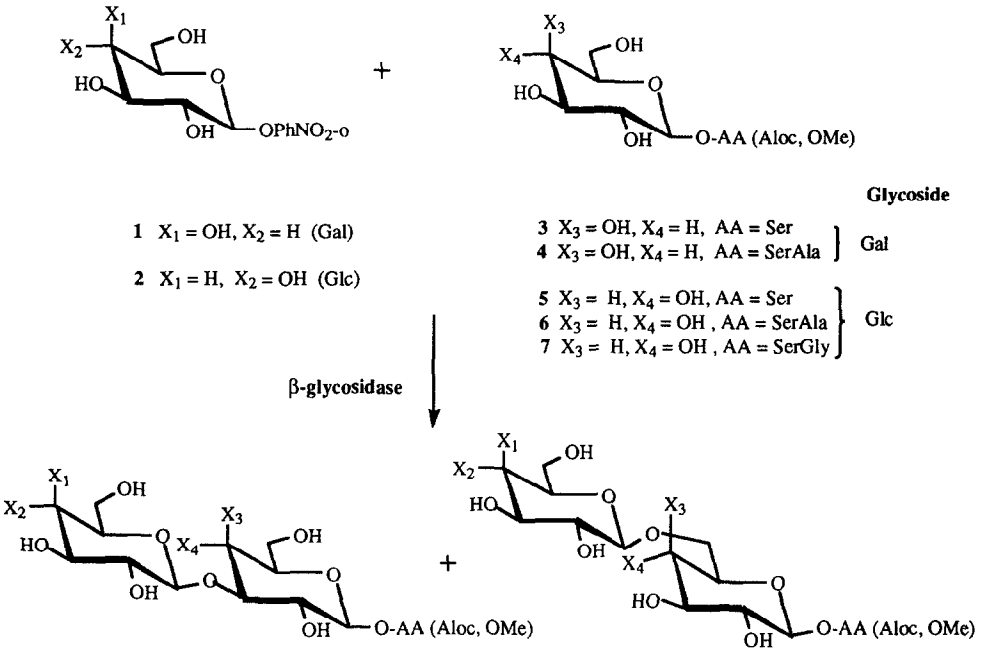
As part of a programm to develop enzymatic synthesis of glycopeptides we present here some of the results we obtained in transglycosylations catalyzed by β–galactosidase and β-glucosidase using glycosyl-serine (or glycosyl-dipeptides) derivatives as acceptors.

In a previous paper we showed that the transgalactosidation from *o*-nitrophenyl-β-D-galactopyranoside (ONPGal) to β-galactosyl-serine (or -dipeptide) conjugates could be achieved with β-galactosidase from *E. coli* as catalyst[7,8]. The reactions were stereospecific and highly regioselective since the β(1→3) digalactosyl-peptide regioisomers were obtained as the major products (ratio β(1→3) : β(1→6) = 9 : 1). Furthermore some of the derivatives (nature of the dipeptide, essentially) were obtained in good yields (up to 50% for **10**, table 1) which is rather unusual for this type of enzymatic condensation.

The reactions that we describe here were performed with *o*-nitrophenyl-β-D-galactopyranoside or *o*-nitrophenyl-β-D-glucopyranoside as the glycosyl donor depending on the enzyme used. By an appropriate combination of the donor and the acceptor four types of disaccharide-peptide conjugates can be obtained.

Digalactosyl-peptide derivatives have been described before. The formation of galactosyl-glucosyl, diglucosyl and glucosyl-galactosyl-serine (or -peptide) conjugates are discussed in this paper.

The transglycosylations catalyzed by β-galactosidase from *E. coli* were performed with *o*-nitrophenyl-β-D-galactopyranoside **1** (ONPGal) as the donor and β-glucosyl-serine **5** (or β–glucosyl-dipeptides **6** and **7**) as the acceptor. The amino acid and the peptide were protected by an allyloxycarbonyl group on the amine and by a methyl ester on the acid.



Diglycoside β-(1-3)			Diglycoside β-(1-6)
Gal-Gal	**8**	$X_1 = X_3 = OH, X_2 = X_4 = H$, AA = Ser	**9**
	10	$X_1 = X_3 = OH, X_2 = X_4 = H$, AA = SerAla	**11**
Gal-Glc	**12**	$X_1 = X_4 = OH, X_2 = X_3 = H$, AA = Ser	**13**
	14	$X_1 = X_4 = OH, X_2 = X_3 = H$, AA = SerAla	**15**
	16	$X_1 = X_4 = OH, X_2 = X_3 = H$, AA = SerGly	**17**
Glc-Glc	**18**	$X_1 = X_3 = H, X_2 = X_4 = OH$, AA = Ser	**19**
	20	$X_1 = X_3 = H, X_2 = X_4 = OH$, AA = SerAla	**21**
Glc-Gal	**22**	$X_1 = X_4 = H, X_2 = X_3 = OH$, AA = Ser	**23**

The reactions of ONPGal with β-glucosyl-serine **5** and *E. coli* β-galactosidase give mainly the β-Gal-(1→6)-β-Glc-OSer regioisomer **13** (ratio β(1→6): β(1→3) = 5 : 1), as opposed to the reactions with the

homologous β-galactosyl-serine **3** as acceptor[8] (28% of the β(1→3) regioisomer **8** and 3% of the β(1→6) isomer **9** : ratio **9** : **8** = 0.1: 1) (Table 1). The products (purified by chromatography on silicagel) have been identified unambigously by ^{13}C N.m.r. spectroscopy by comparison with literature data on different β-linked diglycosides[9,10].

In the reactions of ONPGal with β-galactosyl-serine **3** we observed that the regioselectivity could be reversed if β-galactosidase from *Aspergillus oryzae* was used as catalyst since with this enzyme preparation the β(1→6) isomer is formed as the major product (ratio **9** : **8** = 2.5 : 1).

Table 1 : Enzymatic transglycosylations from activated galactose (or glucose) and some glycosyl-serine or -peptide derivatives

DOR + HOA ⟶ DOA + ROH

DOR (Mole/l)	HOA[a]	HOA/ DOR	Enzyme (Units)[b,c]	Products[d] β(1→3)	Products[d] β(1→6)
βGal-OPhNO$_2$-*o* **1**					
0.36	βGal-OSer **3**	5	βGal. (*E.Coli*), 84U	28% **8**[e]	3% **9**[e]
0.18	"	5	βGal. (*A.Oryzae*), 42U	11% **8**	27% **9**
0.24	βGal-OSerAla **4**	5	βGal. (*E.Coli*), 42U	50% **10**[e]	2% **11**[e]
			βGal. (*A. Oryzae*), 42U	5% **10**	26% **11**
0.24[f]	βGal-OSerGly[f]		βGal. (*E.Coli*), 42U	27%[e]	1%[e]
0.24	βGlc-OSer **5**	5	βGal. (*E.Coli*), 42U	7% **12**	33% **13**
0.36	"	5	βGal. (*A.Oryzae*), 42U	<1% **12**	15% **13**
0.24	βGlc-OSerAla **6**	5	βGal. (*E.Coli*), 42U	2% **14**	20% **15**
	βGlc-OSerGly **7**	5	βGal. (*E.Coli*), 42U	2% **16**	14% **17**
βGlc-OPhNO$_2$-*o* **2**					
0.25	βGlc-OSer **5**	5	β Glucosidase, 15U	4% **18**	19% **19**
	βGlc-OSerAla **6**	5	β Glucosidase, 15U	2% **20**	4% **21**
βGlc-OPhCH$_2$OH-*o* (Salicin)	0.25 βGal-OSer	5	β Glucosidase, 45U	2% **8**	5% **9**

a) The serine residue is protected by an allyloxycarbonyl group on the amine; the acid function of serine or alanine is protected by a methyl ester; b) The units are given for experiments performed on 0.3 mmole of donor; c) the reactions are run at pH 7.8 in 0.03M sodium phosphate buffer (mM $MgCl_2$, 5mM DTT) with β-galactosidase from *E. coli* and at pH 5 in 0.1M sodium acetate buffer with β-galactosidase from *Aspergillus oryzae*; d) after purification on silica gel; e) ref. 8; f) in this case the amine of the dipeptide is protected by a t-butyloxycarbonyl group (ref. 8)

These results appear interesting since changes in regioselectivity are not very common; they have been observed, so far, when changing the anomeric configuration of the glycosyl acceptor[11,12]. Also, different selectivities have been obtained in the reaction of lactose with N-acetylgalactosamine by using β-galactosidase either from *E. coli* [β(1→6) linkage] or from *bovine testes* [β(1→3) linkage][13,14]. However, in the β-galactosidase catalyzed synthesis of nitrophenyl-disaccharides recently reported by Sauerbrei and Thiem, no such high difference in regioselectivity was observed when changing the source of the enzyme (*E. coli* or *Aspergillus oryzae*)[15].

Attempts to reverse the selectivity in the case of β-glucosyl-serine **5** as acceptor were unsuccessful and lower yields are observed with the β-galactosidase from *Aspergillus oryzae* (Table 1).

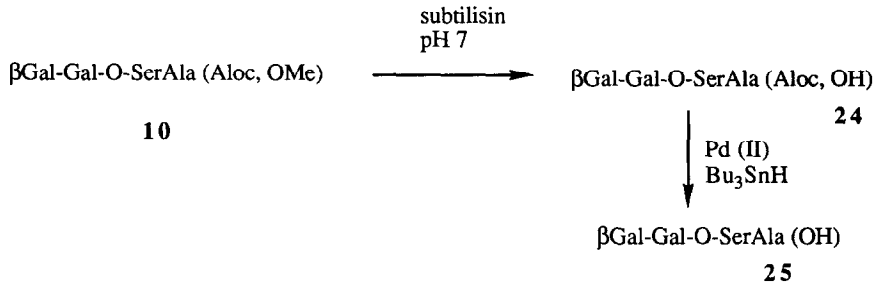
Reactions with the β–glucosyl-dipeptide derivative **6** were performed since high yields of galactosidations were observed when this same serine-alanine dipeptide was β-linked to galactose (50% yield of the disaccharide **10**). In fact we observed the formation of the expected β-Gal-(1→6)-β-Glc-OSer-Ala derivative **15** but with a yield somewhat lower than for diglycosyl-serine **13** (20% as compared to 33%). Reactions with the β–glucosyl-serylglycine derivative **7** give also the same regioselectivity preference but the yield of the β-(1→6) regioisomer **17** (14%) is even lower than for the homologous serylalanine derivative **15** (20%).

Transglycosylations catalysed by β-glucosidase from *Almonds* using *o*-nitrophenyl-β-D-glucopyranoside **2** as donor and a glucosyl-serine derivative as acceptor give also the β(1→6) regioisomer as the major product. However very poor yields of condensation are observed, except for the diglucosyl-serine derivative **19** (19% yield). As compared to the amount of data found in the literature for the reactions with β-galactosidases[16] only few examples have been reported for the transglycosylation with glucosidases[17,18,19]. These enzymes appear to be more specific for the acceptor than galactosidases.

With 2-(hydroxymethyl)phenyl β-D-glucopyranoside (salicin) as donor and galactosyl-serine **3** as acceptor we observed the formation of the digalactosyl-serine derivatives **8** and **9** ; the formation of the expected disaccharides **22** and **23** was not detected. It is not clear whether the enzyme preparation is contaminated by traces of β-galactosidase or whether the enzyme can accept galactose as substrate; such unusual behavior of β-glucosidase[20] or β-galactosidase have already been observed[15].

Full deprotection of the diglycosyl-serine or -peptide derivatives was undertaken. Due to the extreme lability of these glycosidic bonds in basic medium, an enzymatic deprotection of the ester of the aminoacid was considered. Two substrates were submitted to the deprotection by subtilisin: the β-digalactosylserylalanine **10** and the β–galactosyl-glucosylserine **13**.

In a previous paper we showed that subtilisin was very useful for the deprotection of galactosyl-dipeptide esters under mild conditions[7]. The hydrolysis of the β-(1→3)-digalactosyl-dipeptide ester **10** and of the β-(1→6)-galactosyl-glucosyl-serine derivative **13** works also well if subtilisin is used as catalyst under neutral pH and the yield of the expected acids is nearly quantitative.



Removal of the allyloxycarbonyl group is performed by a palladium-catalyzed hydrostannolytic cleavage with tributyltinhydride in the presence of acetic acid as described in the litterature[22]. However, for unknown reasons, the amount of Bu_3SnH needed for complete deprotection is quite high (sixteen fold excess) as already observed for the galactosyl-peptide derivatives[21].

By a combination of glycosyl donors and acceptors we were able to synthetize diglycosyl-peptide conjugates by using the appropriate glycosidases. The reactions are stereospecific and highly regioselective. The best results were obtained in the formation of β-(1→3)digalactosyl-serylalanine derivatives since a 50% yield was obtained with β-galactosidase from *E.Coli* as catalyst. The regioselectivity could be reversed with the *Aspergillus oryzae* β-galactosidase and, in this case, the β-(1→6) regioisomer is obtained. However, although several reports have been made on the use of *Aspergillus oryzae* β -galactosidase for transglycosylation reactions[23-26], the best yields were always obtained, with our substrates, by using β-galactosidase from *E. coli*.

References

1 Montreuil, J. *Adv. Carbohydr. Chem. Biochem.* **1980**, *37*, 157-223-379
2 Schwarz, R.T.; Datema, R. *Adv. Carbohydr. Chem. Biochem.* **1982**, *40*, 287
3 Kunz, H. *Angew. Chem. Int. Ed. Engl.* **1987**, *26*, 294-308
4 a) Paulsen, H. *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 823-839; b) Paulsen, H.; Brenken, M. *Liebigs Ann. Chem.* **1988**, 649-654
5 a) Drueckhammer, Hennen, W.J.; Pederson, P.L.; Barbas, C.F.; Gautheron, C.M.; Krach, T.; Wong, C.-H. *Synthesis*, **1991**, 499; b) Nilsson, K.G.I. TIBTECH, **1988**, *6*, 256-264; c) Ichikawa, Y.; Look, G.L; Wong, C.-H. Anal. Biochem., **1992**, *202*, 215-238
6 a) Augé, C.; Gautheron, C.; Pora, H. *Carbohydr. Res.*,**1989**, *193*, 288-293; b) David, S.; Augé, C.; Gautheron, C. *Adv. Carbohydr. Chem. Biochem.*, **1991**, *49*, 175-237; c) Thiem, J.; Wiemann, T. *Angew. Chem. Int. Ed. Engl.*, **1990**, *29*, 80-82
7 Attal, S.; Bay, S.; Cantacuzene, D. *Tetrahedron*, **1992**, *48*, 9251-9260
8 Bay, S.; Namane, A.; Cantacuzene D. *Carbohydr. Res.* (in press)
9 Jain, R.K.; Kohaya, K.; Abbas, S.A.; Matta, K.L. *Carbohydr. Res.* **1988**, *182*, 290-296
10 Thomas, R.L.; Rutan, J.F.; Abbas S.A.; Matta, K.L. *Carbohydr. Res.* **1989**, *189*, 21-30
11 Nilsson, K.G.I. *Carbohydr. Res.* **1987**, *167*, 95-103
12 Nilsson, K.G.I. *Carbohydr. Res.* **1988**, *180*, 53-59
13 Hedbys, L.; Larsson P.-O.; Mosbach, K.; Svensson, S. *Biochem. Biophys. Res. Commun.*, **1984**, *123*, 8-15
14 Hedbys, L.; Johansson, E.; Mosbach, K.; Larsson, P.-O; Gunnarsson, A.; Svensson, S. *Carbohydr. Res.*, **1989**, *186*, 217-223
15 Sauerbrei, B.; Thiem, J. *Tetrahedron Lett.*, **1992**, *33* , 201-204
16 Toone, E.J.; Simon, E.S.; Bednarski, M.D.; Whitesides G.M. *Tetrahedron* **1989**, *45*, 5365-5422
17 Vic, G.; Thomas, D. *Tetrahedron Lett.*. **1992**, *33* , 4567-4570
18 Turner, N.J.; Webberley, M.C. *J. Chem. Soc., Chem. Commun.*, **1991**, 1349-1350

19 Vulfson, E.N.; Patel, R.; Law, B.A. *Biotechnol. Lett.* **1990**, *12*, 397-402
20 Helferich, B.; Kleinschmidt, T. *Z. Physiol. Chem.*, **1967**, *348*, 753-758
21 Cantacuzene, D., Attal, S.; Bay, S. *BioMed. Chem. Lett.* **1991**, *1*, 197-200
22 Dangles, O.; Guibé, F.; Balavoine, G.; Lavielle,S.; Marquet, A.*J. Org. Chem.*, **1987**, *52*, 4984-4993
23 Ooi, Y.; Hashimoto,T.; Mitsuo, N.; Sato, T. *Chem. Pharm. Bull.* **1985**, *33*, 1808-1814
24 Fortun, Y.; Colas, B. *Biotechnol. Lett.*, **1991**, *13*, 863-866
25 Petit, J.-M; Paquet, F.; Beau, J.-M. *Tetrahedron Lett.* **1991**, *32*, 6125-6128
26 Kren, V.; Sedmera, P.; Havlicek, V.; Fiserova, A. *Tetrahedron Lett.*, **1992**, *33*, 7233-7236
27 Characteristics are given for some representative products:

12: ^{1}H n.m.r. (D_2O, 300MHz) δ 3.48 (H-2), 3.58 (H-2'), 3.67 (H-3'), 3.74 (H-3, H-4), 3.79 (CH_3), 3.68-3.9 (H-5, 2H-6,H-5', 2H-6'), 3.91 (H-4', $J_{3',4'}$ 3.3Hz), 3.98-4.33 (CH_2 ser, J 3.7Hz, 4.8 Hz, 10.7 Hz), 4.5 (H-1, $J_{1,2}$ 8Hz), 4.46, CHα), 4.6 (CH_2 allyl), 4.65 (H-1' $J_{1',2'}$ 7.6 Hz) 5.28 (CH_2 allyl), 5.98 (CH= allyl); ^{13}C n.m.r. δ 53.82 (COO*CH*$_3$), 54.92 CHα), 61.22-61.67 (C-6, C-6'), 66.79 (CH_2 allyl), 68.74-69.19 (C-4, C-4'), 69.71 (CH_2 ser, 71.83 (C-2'), 73.16-73.26 C-2, C-3'), 75.95-76.13 (C-5, C-5'), 84.89 (C-3), 102.82-103.9 (C-1,C-1'), 117.92 (CH_2 allyl), 133.12 (CH= allyl), 158.67 (OCONH), 173.02 (*C*OOCH$_3$). F.a.b. m.s. (M+H$^+$) 528. [α$_D$] -16° (c 0.52, water).

13: ^{1}H n.m.r. (D_2O, 300MHz) δ 3.26 (H-2), 3.47 (H-3), 3.53 (H-2'), 3.6 (H-4), 3.65 (H-3'), 3.77 (CH_3), 3.48-3.8 (H-5, H-5', 2H-6'), 3.83 (dd, H-6, $J_{6a,6b}$ 11.8 Hz; $J_{5,6a}$ 5.3 Hz)), 3.9 (H-4', $J_{3',4'}$ 3.2 Hz), 3.95 (dd, CH_2 ser, J 3.7Hz, 10.7 Hz), 4.19 (dd, H-6, $J_{5,6b}$ 1.8 Hz), 4.3 (CH_2 ser, J 4.9 Hz), 4.42 (H-1', $J_{1',2'}$ 8.1 Hz), 4.45 (H-1, $J_{1,2}$ 8.5 Hz), 4.54 (CHα), 4.59 (CH_2 allyl), 5.28 (CH_2 allyl), 5.93 (CH= allyl); ^{13}C n.m.r. δ 54.10 (COO*CH*$_3$), 55.15 CHα), 61.85 (C-6'), 67.04 (CH_2 allyl), 69.19-70.11 (CH_2 ser, C-6) 69.49-70.11 (C-4, C-4'), 71.60 (C-2'), 73.53-73.79 (C-2, C-3'), 75.86-76.01-76.33 (C-3, C-5, C-5'), 103.40-104.32 (C-1, C-1'), 118.20 (CH_2 allyl), 133.38 (CH=allyl), 158.94 (OCONH), 173.22 (*C*OOCH$_3$). m.p. 103°C; F.a.b. m.s. (M+H$^+$) 528. [α$_D$] -13° (c 1.8, water).

15: ^{1}H n.m.r. (D_2O, 300MHz) δ 1.4 (CH_3, ala), 3.31 (H-2), 3.51 (H-3), 3.53 (H-2'), 3.56 (H-2'), 3.66 (H-3'), 3.77 (CH_3 ester), 3.48-3.82 (H-4, H-5, H-5', 2H-6'), 3.87 (dd, H-6, $J_{6a,6b}$ 11.8 Hz; $J_{5,6a}$ 5.6 Hz), 3.93 (H-4', $J_{3',4'}$ 3.2 Hz), 3.95-4.17 (dd, CH_2 ser, J 3.7Hz, 6.2 Hz, 10.7 Hz), 4.23 (dd, H-6, $J_{5,6b}$ 1.5 Hz), 4.43 (CHα ser), 4.46 (H-1', $J_{1',2'}$ 7.9 Hz), 4.47 (CHα ala), 4.49 (H-1, $J_{1,2}$ 8 Hz), 4.62 (CH_2 allyl), 5.31 CH_2 allyl), 5.98 (CH= allyl); ^{13}C n.m.r. δ 16.75 (CH_3 ala), 49.59 (CHα ala), 53.71 (COO*CH*$_3$), 55.67 CHα ser), 61.74 (C-6'), 67.0 (CH_2 allyl), 69.25-69.72 (CH_2 ser, C-6) 69.40-70.13 (C-4, C-4'), 71.52 (C-2'), 73.43-73.73 (C-2, C-3'), 75.77-75.91-76.22 (C-3, C-5, C-5'), 103.06-104.27 (C-1, C-1'), 118.23 (CH_2 allyl), 133.26 (CH=allyl), 158.60 (OCONH), 172.53 (*C*OOCH$_3$), 175.51 (CONH). m.p. 107°C; F.a.b. m.s. (M+H$^+$) 599. [α$_D$] -27° (c 1.5, water).

19: ^{1}H n.m.r. (D_2O, 300MHz) δ 3.27 (H-2), 3.3 (H-2'), 3.45 (H-5'), 3.47 (H-3, H-4, H-4'), 3.5 (H-3'), 3.6 (H-5), 3.71 (H-6', $J_{6'a,6'b}$ 12.4 Hz, $J_{5',6'a}$ 5.6 Hz), 3.79 (CH_3 ester), 3.85 (dd, H-6, $J_{6a,6b}$ 11.8 Hz; $J_{5,6a}$ 5.6 Hz), 3.92 (H-6' $J_{5',6'b}$ 2.1 Hz), 3.97 (CH_2 ser, J 3.7Hz, 10.7 Hz), 4.19 (dd, H-6, $J_{5,6b}$ 1.7 Hz), 4.32 (CH_2 ser, J 4.8 Hz), 4.46 (H-1 $J_{1,2}$ 8 Hz), 4.5 (H-1', $J_{1',2'}$ 8 Hz) , 4.5(CHα), 4.6 (CH_2 allyl), 5.31 CH_2 allyl), 5.87 (CH= allyl); ^{13}C n.m.r. δ 53.81 (COO*CH*$_3$), 54.86 (CHα), 61.31 (C-6'), 66.74 (CH_2 allyl), 69.01-69.82 (CH_2 ser, C-6) 69.82-70.20 (C-4, C-4'), 73.49-73.64 (C-2, C-2'), 75.54-76.03-76.20-76.49 (C-3, C-5, C-3', C-5'), 103.10-103.45 (C-1, C-1'), 117.9 (CH_2 allyl), 133.09 (CH=allyl), 158.64 (OCONH), 172.93 (*C*OOCH$_3$). m.p. 104°C; F.a.b. m.s. (M+H$^+$) 528. [α$_D$] -21° (c 1.2, water).

25: ^{1}H n.m.r. (D_2O, 300MHz) δ 1.32 (CH_3 ala), 3.56 (H-2'), 3.63 (H-3'), 3.69 (H-2), 3.54-3.8 (H-5, 2H-6, H-5', 2H-6'), 3.8 (H-3, $J_{2,3}$ 9.8 Hz, $J_{3,4}$ 3 Hz), 3.87 (H-4', $J_{3',4'}$ 2.6 Hz), 4.1 (CHα ala), 4.15 (H-4), 4.09-4.24 (CHα ser, CH_2 ser), 4.5 (H-1, $J_{1,2}$ 7.7 Hz), 4.57 (H-1', $J_{1',2'}$ 7.3 Hz); ^{13}C n.m.r. δ 17.96 (CH_3 ala), 52.26 (CHα ala), 54.01 (CHα ser), 61.77 (C-6, C-6'), 68.33 (CH_2 ser), 69.19, 69.37 (C-4, C-4'), 70.55 , 71.83 (C-2, C-2'), 73.31 (C-3'), 75.76, 75.87 (C-5, C-5'), 82.86 (C-3), 103.01, 105.13 (C-1, C-1'), 167.16 (CONH), 180.38 (COOH). m.p. 181°C; F.a.b. m.s. (M+H$^+$) 501. [α$_D$] +9° (c 1.4, water).