Accepted Manuscript

Synthetic multivalent ligands for cholera & cholera-like toxins: Protected cyclic neoglycopeptides

Vajinder Kumar, Narender Yadav, K.P. Ravindranathan Kartha

PII: S0008-6215(16)30181-1

DOI: 10.1016/j.carres.2016.05.011

Reference: CAR 7205

To appear in: *Carbohydrate Research*

Received Date: 13 April 2016

Revised Date: 24 May 2016

Accepted Date: 25 May 2016

Please cite this article as: V. Kumar, N. Yadav, K.P.R. Kartha, Synthetic multivalent ligands for cholera & cholera-like toxins: Protected cyclic neoglycopeptides, *Carbohydrate Research* (2016), doi: 10.1016/j.carres.2016.05.011.

Synthetic multivalent ligands for cholera & cholera-like toxins: Protected cyclic neoglycopeptides

Vajinder Kumar, Narender Yadav and K. P. Ravindranathan Kartha*

Department of Medicinal Chemistry, National Institute of Pharmaceutical Education and Research (NIPER), S.A.S. Nagar, Punjab-160062, India Fax: 91-172-2214692

E-mail: rkartha@niper.ac.in

Abstract

Synthesis of a set of novel glycopeptide analogues as potential cholera/cholera-like toxin inhibitors in their protected form is described. They include di-, tri-, tetra- and pentavalent scaffolds. The synthetic steps were achieved using a combination of solvent-free mechanochemical as well as the conventional solution-phase reactions. During the conventional DIC-HOBt-mediated peptide coupling followed for the preparation of certain glycopeptide analogues an interesting *in situ* Fmoc deprotection was observed which has been demonstrated to hold potential for synthesiszing glycopeptides/neoglycopeptides with extended polyamide chains.

Keywords: Glycoconjugates; Neoglycopeptides; Diarrhoea; Cholera toxin; Multivalent inhibitors

Highlights

- Synthesis of a set of protected neoglycopeptides are reported

- They include di-, tri-, tetra- & pentavalent scaffolds bearing upto 10 sugar units
- Intended to derive potential inhibitors for cholera toxin/other analogous proteins
- Sugars are attached by glycosylation to cyclic/acyclic backbone structures

Introduction

Gastrointestinal infections have foremost impact in the developing world, where diarrhoeal diseases account for approximately 1.5 million deaths each year.[1] Though it can be caused by a vast number of bacteria, viruses, and parasites, the three important members of this class of toxins have been the pertussis toxin, the cholera toxin (CT) and the shiga toxin (ST). The important agent responsible for the diarrhoeal infection is the AB_5 class of toxins found in the microorganism.[2] They possess a symmetric architecture and consist of a subunit A and a pentamer of B subunits. While the catalytic activity of the toxins is due to the A subunit, specific binding to the cell surface and the delivery of the toxin is controlled by the B subunits. Therefore although the complete AB_5 holotoxin is considered as required for their toxic effects it is the B subunit that is important in designing inhibitors for the toxin.[3] The crystal structures for these toxins have already been reported. Our specific interest was in the cholera toxin for which the cell surface ganglioside GM1 is known to be the natural receptor.[4,5]

The mechanism of action of the AB_5 toxins offers several possibilities for targeting the infection,[6] and additionally, as the inhibitors do not need to cross any barrier (blood barrier), it as such does not necessarily place any constraints on the ligand size. In the recent past several strategies have been adopted for finding effective inhibitors for the AB_5 toxins; while some target on the individual binding sites, the others are intended at designing multivalent ligands

against the B subunit pentamer.[7] Monovalent ligands can be designed as mimics of the natural ligand preferably with simpler structural complexity and improved receptor-binding ability. Taking cue from the symmetry of the pentameric B subunit attempts have also been to synthesize polyvalent ligands with potentially enhanced receptor-affinity. Thus the *starfish* ligand designed and synthesized by Kitov *et al.* led to a sandwich of SLT-ligand complex formed upon binding, with the ligand sandwiched between the two B_5 units of the AB_5 toxin.[7] Importantly, the ligand was 10^7-fold more potent than the monomeric Pk trisaccharide.[7-10] Thus the advantage of designing multivalent ligands for such protein structures as the B_5 unit of the AB_5 class of toxins was clearly demonstrated.

The *starfish* model consisted of five units of a pair of cross-linked Pk trisaccharide units anchored on to a central glucose core through rather long flexible linkers. Although the ligands were very powerful and were the best known until then, the flexibility of the linker would have caused some loss of binding (energy) arising out of the entropy loss due to rotational and translational freedom of the ligands. Therefore a more rigid platform for anchoring the carbohydrate ligand units would be a rational choice. Zhang *et al.* synthesised a small library of cyclic peptides of increasing ring size as the anchor for the carbohydrate ligand. The carbohydrate unit, namely, galactose, the minimum structural component required for binding with the toxin, was then anchored in sufficient numbers on to the central cyclic peptide core through flexible spacer units to achieve the desired pentavalency.[11] In comparison to the monomeric ligand, the multivalent glycopeptides were indeed proved stronger inhibitors of the toxin. Thus, it could be seen that if we were to decrease the entropy cost of binding by presenting the sugar units using a platform that is somewhat more rigid, it could lead to an increase in the binding affinity. Hence, the interest was to design a structure with reduced flexibility (enhanced

rigidity) to enable the entropy cost reduce significantly. The sugar (galactose) units shall be placed at such distances to each other as to enable them span the distance between two successive binding sites on the pentameric B subunit. For CT, this distance was estimated as approximately 23 Å,[12] and was roughly the same in all the AB_5 toxins. The distance of 23 Å, it can be seen, could be spanned/provided by a linear unit of fifteen C-C bonds, which for instance, can be had from a 1,12-diaminododecane moiety when introduced between two amino acid residues to which are attached the galactose unit (or alternatively, if desired, a galabiose/globotriose unit) each and an amino substituent placed on a phenyl ring attached to the desired sugar in the form of a glycoside can be used for attaching the sugar to the spacer group. Thus, the synthesis of a set of linear and cyclic glycopeptide analogues (in their protected form) as potential inhibitors of CT (see Fig 1 for a cartoonic representation) carried out based on these considerations are summarized here.

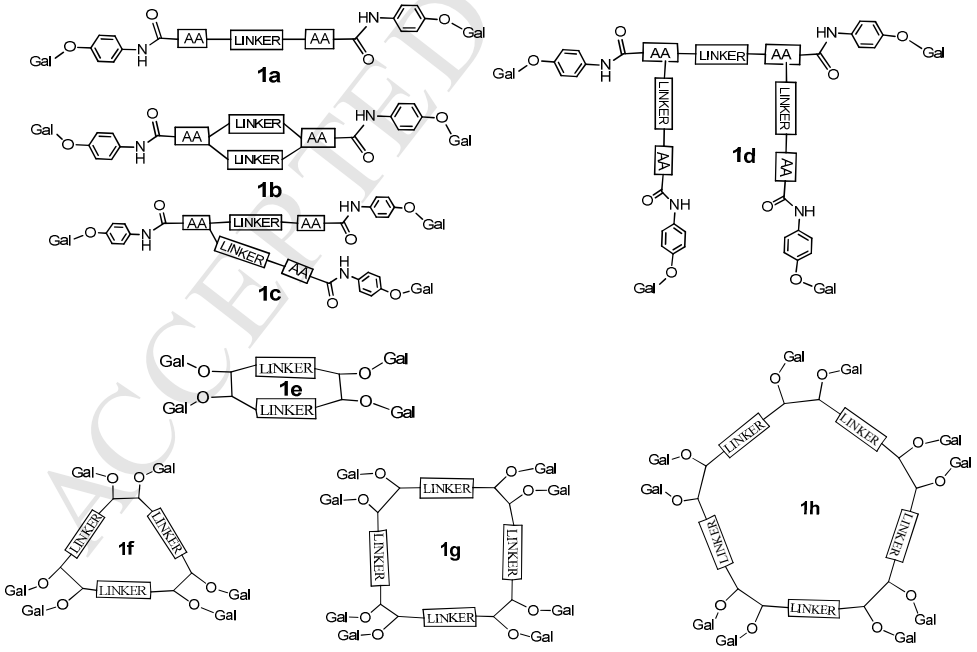
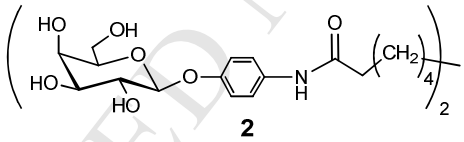


Figure 1. A cartoonic representation of the proposed glycoconjugates as potential polyvalent protein inhibitors

Results and discussion

Synthesis of cross-linked galactosides as potential divalent ligands for CT. Recently we reported the synthesis of a cross-linked diamide **2** by a solvent-free mechanochemical method.[13] When this molecule was examined as a ligand for CT by molecular docking experiments it was revealed that the length of the ligand fell insufficient for bridging the space between two adjacent lobes of the AB_5 protein to enable an effective binding of the two galactose residues on to the respective lobes. Insertion of an amino acid unit could address this deficiency. Thus the synthesis of cross-linked galactosides (analogous to **2**, Fig 2) with an extended linker-unit became necessary.



2

Figure 2. Structure of a previously prepared cross-linked galactopyranoside

Three different amino acids, namely, lysine, glutamic acid and alanine, were used for the linker-chain extension. The compounds **3-8** prepared under this category, along with the building blocks **9-18** used for their synthesis, are shown in Fig 3. The synthesis of the galactopyranoside derivative **9** was by a recently reported solvent-free mechanochemical method[14] using acetobromogalactose and *p*-nitrophenol as the reactants in the presence of K_2CO_3 followed by the reduction of the nitro group on the aglycon moiety by catalytic hydrogenation using H-Cube (flow chemistry). Compound **9** was then coupled to the lysine derivative **10** in the presence of

HOBt and DIC in DMF following a conventional protocol for amino acid coupling.[15] The N^2-Fmoc group on the anilide **11** obtained was then removed and the free amino group so generated was coupled to the dicarboxylic acid linker **12** under essentially the same conditions as for **11**[16] to afford the protected cross-linked galactoside **3**. In a similar manner the cross-linked galactosides **4** and **5** were also

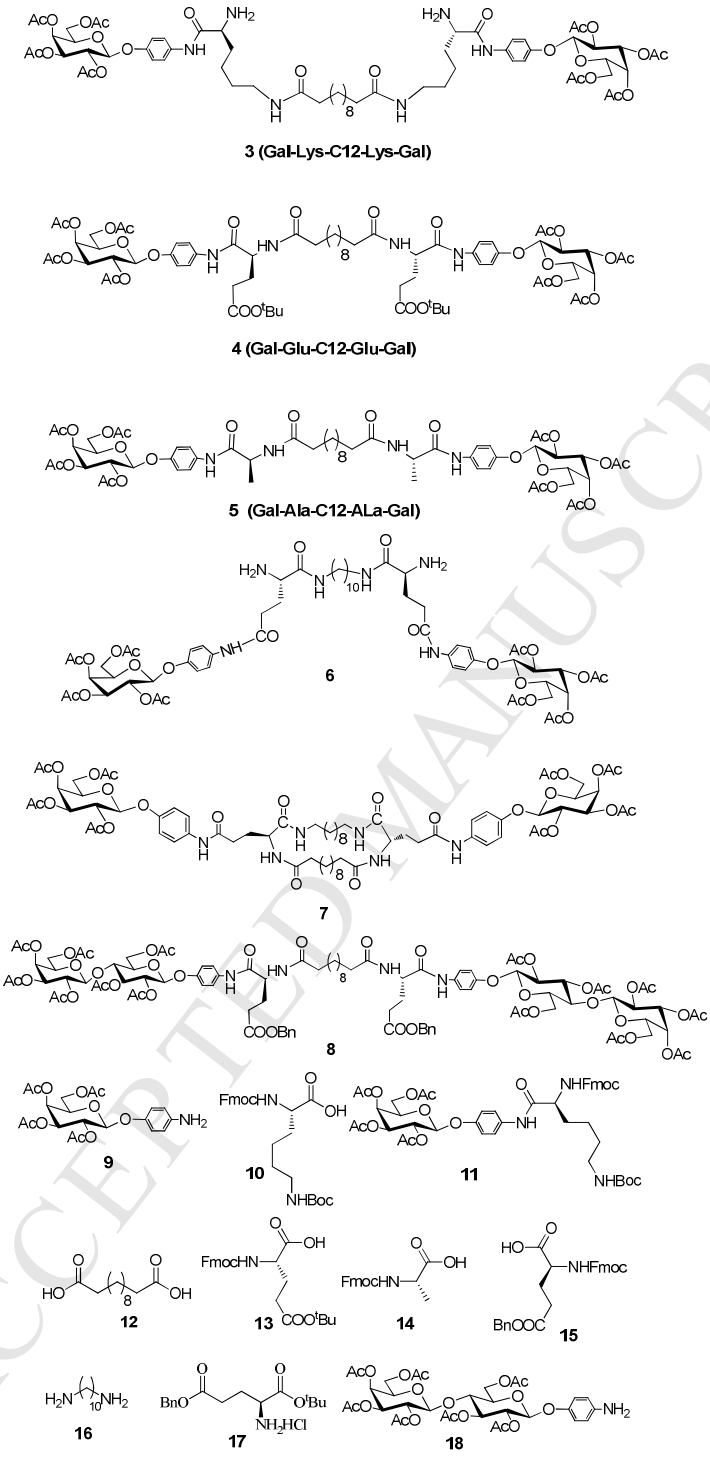


Figure 3. Various divalent compounds synthesized in the protected form and the building blocks used for their synthesis

prepared using the required building blocks **13** and **14** respectively in place of **10**. Likewise, using the partially protected glutamic acid derivative **15** and the diamino linker **16**, the cross-linked analogous galactoside **6** was also synthesized. Coupling of the dicarboxylic acid linker **12** with **6** gave rise to the cross-linked galactoside **7** in which the two galactose residues are appended to a central core (that forms a loop-like structure) *via* the *p*-aminophenyl units as can be seen from Fig 3 (structure **7**). As an additional related structure the lactose-linked compound **8** was also synthesized in a similar manner. Thus, when the dicarboxylic acid **12** was coupled to the partially protected glutamic acid building block **17** and the resulting cross-linked glutamic acid derivative, after deprotection of the *t*-butyl ester groups, was coupled to the lactoside **18**,[16] yielded the cross-linked lactoside **8**.

Synthesis of tri- and tetravalent ligands for CT. The methods employed for the synthesis of compounds, **19** and **20** (Fig 4), were essentially the same as those described in the foregoing section. The building blocks (**21** and **22**)[16] used for the synthesis have also been shown in Fig 4.

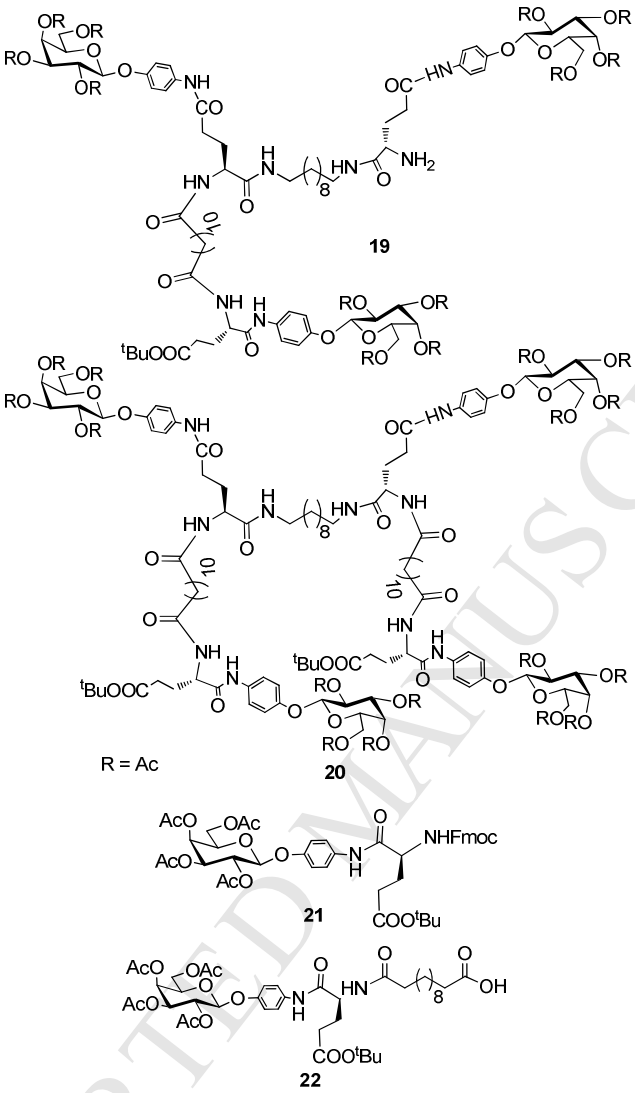


Figure 4. Synthetic tri- and tetravalent neoglycopeptide ligands

The glycosylated amino acid **21**[16] was sourced from the galactoside **9** and the partially protected amino acid **13** by coupling them in the presence of DIC and HOBt as mentioned earlier. Following the removal of the Fmoc group in the coupled product **21**, the amine was further coupled to **12** in a manner similar to the method adopted for the preparation of **4** from **9** and **13** but in the presence of excess of the dicarboxylic acid linker **12** in order to ensure preponderance of the mono-coupled product **22**.[16] The free carboxylic unit of **22** was then used to perform the

coupling reaction with the free amino unit (s) on the divalent compound described above to yield the targeted tri- and tetravalent compounds **19** and **20** (Fig 4). The separation of **19** and **20** from the reaction mixture could be very successfully carried out by gel permeation chromatography (GPC) using a column of HW-40S resin and employing DMF as the eluent. They were obtained in 30% and 38% yields, respectively.

Cases of in situ Fmoc-deprotection observed during amino acid coupling. In an alternative approach towards the tri- and tetravalent glycopeptides of the type **19** and **20** (Fig 4) described above, the aminophenyl galactoside **9** was first coupled to a lysine residue protected at its N^2 and N^6 positions by Fmoc (**23**) to yield the glycopeptide building block **24** (Fig 5).[16] Following the

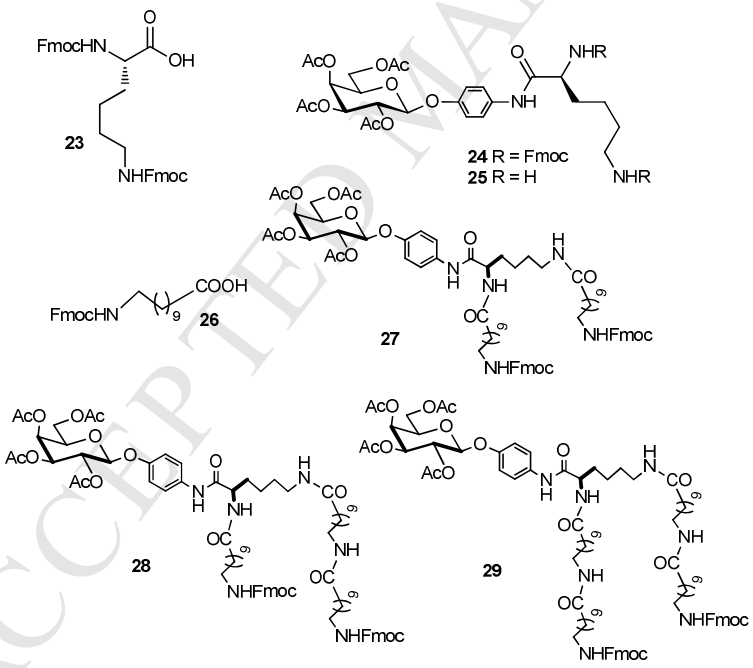


Figure 5. Unexpected by-products **28** and **29** in the preparation of the neoglycopeptide **27** using building blocks **23/26**

removal of the Fmoc groups in **24** by treatment with piperidine, when the diamine **25**[17] obtained was then coupled to the Fmoc-protected alkyl amino carboxylic acid **26**, besides the expected **27** (isolated in 40% yield), were obtained the glycopeptides **28** (isolated in 30% yield) and **29** (isolated in 10% yield) as by-products. All the three products could successfully be separated by GPC using a column of HW-40S resin as mentioned above. The compounds were characterized by spectroscopy (NMR and HRMS). Clearly, the by-products **28** and **29** have resulted from a selective *in situ* Fmoc deprotection of the initial coupling product and a subsequent round of coupling with **26**. Selective deprotection of the Fmoc group on either the terminal or the internal fatty alkyl chain of the initially formed **27** followed by coupling with the carboxylic acid **26** must yield the by-products **28** and **29**, respectively. Arguably, while this may have indeed been due to the traces of any residual piperidine that may have been left behind in the reaction mixture prior to the second round of coupling, is perhaps unprecedented (just having not been reported by any so far) in peptide coupling.

That this is not an isolated case was demonstrated by the formation of the by-products **32** and **33** during the preparation of the neoglycopeptide structure **31** from the building blocks **30** (and **26**) as depicted in Fig 6. Compound **30** was in turn obtained from the coupling of

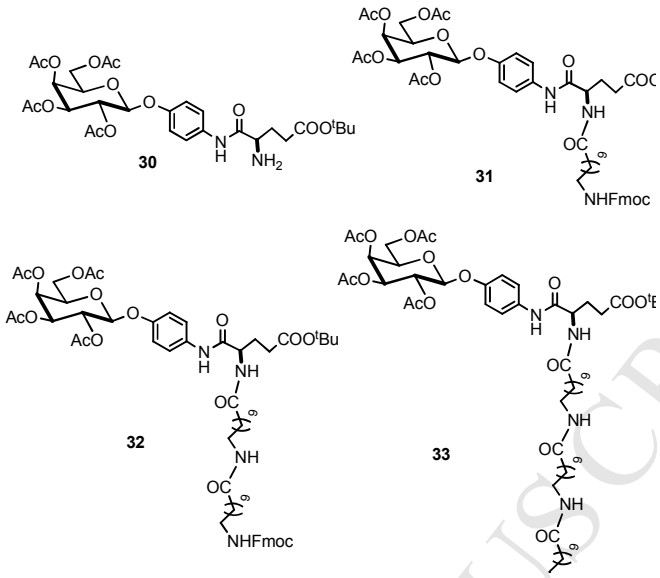


Figure 6. Unexpected by-products **32** and **33** in the preparation of the neoglycopeptide **31** using building blocks **30** and **26**

9 and **15** described earlier. Undoubtedly, subsequent to the formation of **31**, it undergoes the deprotection of the Fmoc group present in it followed by coupling of the newly generated amino group with the carboxylic acid unit of **26** present in the reaction mixture. A similar sequence of reactions when subjected to the newly formed **32** must now necessarily lead to the formation of **33**. Thus, this reaction clearly holds good potential for the synthesis of such structures as **33** with extended polyamide chain. These molecules, besides serving as ligands to be evaluated against polyvalent proteins such as CT, are also expected to possess interesting supramolecular self-assembling behaviour to be investigated by TEM.

An alternative approach towards the targeted polyvalent ligands explored was starting from diethyl tartrate (**34**) in which is available two hydroxyl groups that could be

conveniently glycosylated to afford a unit having two pendent galactoside residues. Thus, the commercially available diethyl tartrate **34** was reacted with phenyl 2,3,4,6-tetra-*O*-benzyl-1-thio-β-D-galactopyranoside **35** under solvent-free mechanochemical conditions promoted by In(III) triflate and the diglycosylated product **36**[17] was prepared in multi-gram quantities. After saponification of the ethyl ester functionality by aq. LiOH it was coupled to the diamine linker **37**, again in the presence of DIC and HOBt in DMF, and the multivalent ligands **38-41** formed (Fig 7) were conveniently obtained in pure form by separation by GPC on an HW-40 column using DMF as the eluent and were characterized by NMR and mass spectrometry (MALDI-TOF-MS) as usually for glycopeptides and their analogues. During the above reaction, formation of some of the analogous open chain compounds with free carboxylic and amino terminal groups were also formed, but, they were proved difficult to get in pure form for unambiguous characterization and therefore have not been shown here.

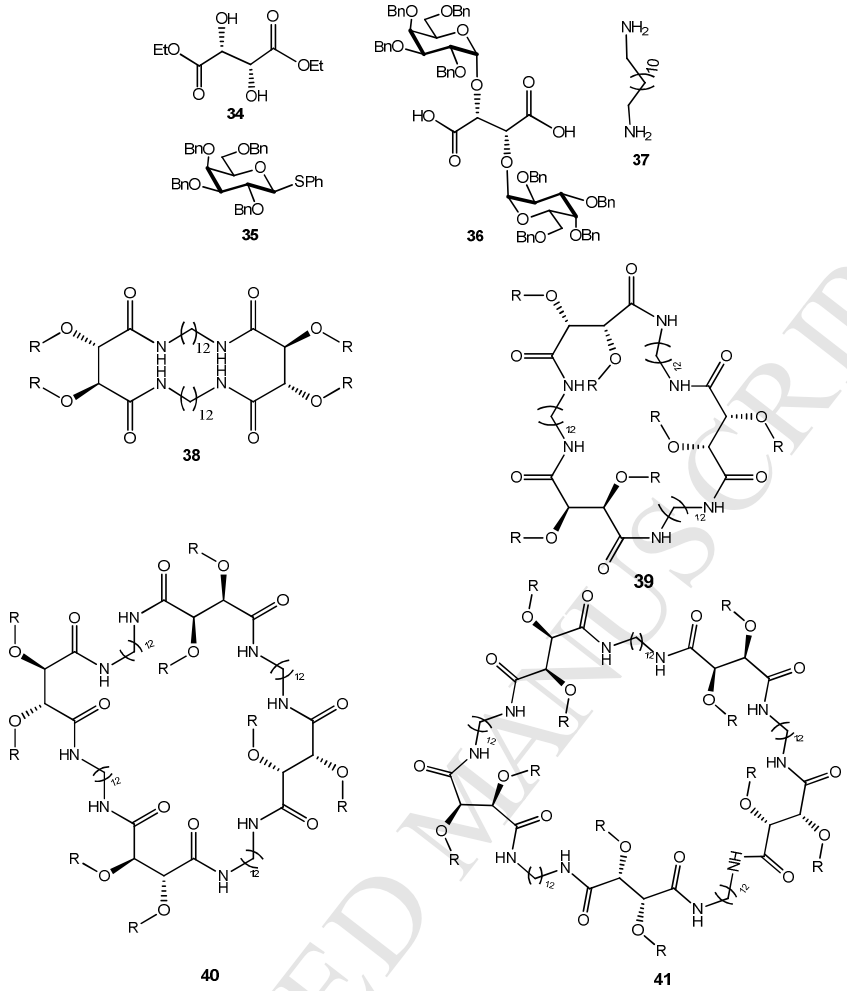


Figure 7. Synthetic multivalent neoglycopeptide ligands as potential CT inhibitors prepared in their protected form and the building blocks used for their preparation.

Experimental

General experimental methods

All the reagents used were as purchased without further purification. Solvents used for reactions were dried according to standard methods. Reactions were monitored by TLC, which was performed with 0.2 mm pre-coated silica gel 60 F254 aluminum sheets. Compounds were detected by dipping the TLC plates in an ethanolic solution of sulphuric acid (5% v/v)/alkaline

bromocresol dye/ninhydrin solution and thereafter heating them. Melting points were determined on a Büchi melting point apparatus. Specific rotations were recorded on a digital polarimeter at room temperature (approximately 20-25 °C). NMR spectra were recorded on 400 MHz spectrometer. ^{1}H NMR and ^{13}C NMR Spectra were referenced using either residual solvent signals, or tetramethylsilane in the respective deuterated solvents. Whenever needed ^{1}H-^{1}H COSY and ^{1}H-^{13}C HMQC were used additionally to confirm/assist in the NMR peak assignment. Splitting patterns are described as singlet (s), doublet (d), triplet (t), quartet (q), and broad (br); the value of chemical shifts (δ) are given in ppm and coupling constants (J) are reported in Hertz (Hz). Mass spectra were recorded on MALDI-TOF/TOF and HRMS (TOF) Spectrometers.

General experimental procedure

General amide coupling procedure: The compound having free carboxylic group (1 mmol) was dissolved in DMF (15 mL) and HOBT (1.2 mmol/carboxylic group) was added to the solution and was stirred for 15 min. The reaction mixture was then cooled to -10 ˚C. After 15 min, DIC (1.2 mmol/carboxylic group) was added to the reaction mixture and stirring was continued for another 30 min. The compound (to be coupled) bearing free amino group was then added to the flask and the stirring was allowed to continue at 4 ˚C until TLC showed completion of the reaction (up to 36 h at 4 ˚C). On completion of the reaction, the solvent was evaporated under reduced pressure and the crude product was purified by GPC (Gel Permeation Chromatography) on HW-40S resin using CH_2Cl_2–MeOH (1:1) (for compounds **3-8**, **19-21**, **27-29**, **31-33**) or DMF (for compounds **38-41**) as the eluent [(i) Using a chromatographic equipment: Column, 100 cm (h) × 2.5 cm (d); flow rate, 12.5 ml/min; fraction size 300 drops on Spectrum/Chrom CF-2 fraction collector from Spectrum Chromatography, Houston USA; or (ii) Manual: Column, 150 cm (h) × 3.0 cm (d); flow rate 1.0 ml/min by gravity; fraction size 6 ml].

Divalent compound, Gal-Lys-C12-Lys-Gal (3): Yield - 60% (0.650 g). Colourless solid: m. p. 199-204 °C; $[\alpha]_D$ -2.9° (*c* 0.2, $CHCl_3$); ^1H NMR (400 MHz, $CDCl_3$) δ 7.76 (d, 4H, Ph-H), 7.58 (bs, 4H, Ph-H), 7.49 (m, 4H, Ph-H), 7.38 (m, 4H, Ph-H), 7.27 (m, 4H, Ph-H), 6.94 (d, 4H, Ph-H), 5.97 (d, 1H, N-H), 5.90 (bs, 1H, N-H), 5.49-5.45 (m, 2H, H-2, H-4), 5.13-5.09 (dd, 1H, H-3), 4.96 (d, 1H, H-1), 4.41-4.30 (m, 3H, methylene proton of Fmoc, *tert*-C-H), 4.23-4.13 (m, 3H, H-6a, H-6b, C-H of Fmoc), 4.08-4.02 (m, 1H, H-5), 3.25 (bs or m, 2H, methylene protons of lysine), 2.20 (s, 3H, $COCH_3$), 2.15-2.12 (t, 2H, methylene protons of lysine), 2.09 (s, 3H, $COCH_3$), 2.05 (s, 3H, $COCH_3$), 2.02 (s, 3H, $COCH_3$), 1.49-1.39 (m, 4H, methylene protons), 1.30-.122 (m, 13H, methylene protons); $^{13}C\{^1H\}$ NMR (100 MHz, $CDCl_3$) 170.4, 170.3, 170.2, 169.5, 143.6, 141.3, 127.8, 127.1, 125.0, 121.5, 120.0, 117.5, 100.1, 71.0, 70.8, 68.7, 67.2, 66.9, 61.4, 47.1, 36.7, 33.8, 31.9, 29.7, 29.4, 28.9, 28.7, 25.5, 20.8, 20.7; MS (MALDI-TOF) $C_{94}H_{112}N_6O_{28}Na$, calculated m/z 1795.741 $(M+Na)^+$, found m/z 1796.483 $(M+Na)^+$.

Divalent compound, Gal-Glu-C12-Glu-Gal (4): Yield- 60% (0.880 g). Colourless solid: m. p. 103-115 °C; $[\alpha]_D$ -34.4° (*c* 0.4, $CHCl_3$); ^1H NMR (400 MHz, $CDCl_3$) δ 7.46 (d, 2H, Ph-H), 6.95 (d, 2H, Ph-H), 6.77 (d, 1H, N-H), 5.51-5.42 (m, 2H, H-2,H-4), 5.13-5.09 (dd, 1H, H-3), 4.98 (s, 1H, H-1), 4.60-4.50 (m, 2H, methylene of Fmoc), 4.26-4.11 (m, 3H, H-6a, H-6b, C-H), 4.06 (bt, 1H, H-5), 2.61-2.48 (m, 1H, one of methylene proton), 2.41-2.26 (m, 1H, one of methylene proton), 2.18 (s, 3H, $COCH_3$), 2.07 (s, 3H, $COCH_3$), 2.05 (s, 3H, $COCH_3$), 2.01 (s, 3H, $COCH_3$), 1.70-1.54 (m, 9H, methylene protons), 1.47 (s, 9H, CCH_3), 1.33-1.19 (m, 8H, methylene protons); $^{13}C\{^1H\}$ NMR (100 MHz, $CDCl_3$) 174.3, 173.7, 172.0, 170.4, 169.1, 158.7, 151.2, 143.5, 140.5, 134.1, 133.7, 130.0.8, 128.5, 124.7, 121.3,

120.0, 117.6, 100.1 (C-1), 81.4, 71.0, 70.8, 68.6, 67.2, 66.8, 61.4, 53.3, 48.0, 31.9, 30.32, 28.8, 28.1, 27.3, 25.4, 23.6, 20.7; MS (MALDI-TOF) $C_{70}H_{98}N_4O_{28}Na$, calculated m/z 1465.626 $(M+Na)^+$, found m/z 1465.740 $(M+Na)^+$.

Divalent compound, Gal-Ala-C12-Ala-Gal (5): Yield- 71% (0.870 g). Colourless solid: m. p. 101-115 °C; $[\alpha]_D$ -36.7° (*c* 0.4, $CHCl_3$); ^1H NMR (400 MHz, $CDCl_3$) δ 9.37 (bs, 1H, N-H), 7.46 (d, 2H, Ph-H), 7.01 (bs, 1H, N-H), 6.91 (d, 2H, Ph-H), 5.44-5.41 (m, 2H, H-2,H-4), 5.17-513 (dd, 1H, H-3), 4.99 (s, 1H, H-1), 4.77-4.73 (m, 1H, *tert* C-H), 4.18-4.06 (m, 3H, H-6a, H-6b, H-5), 2.20 (s, 3H, $COCH_3$), 2.05 (s, 3H, $COCH_3$), 2.02 (s, 3H, $COCH_3$), 2.01 (s, 3H, $COCH_3$), 1.61-1.51 (M, 2H, methylene protons), 1.44 (d, 3H, $C(CH_3)$, 1.26-1.15 (m, 8H, methylene protons); $^{13}C\{^1H\}$ NMR (100 MHz, $CDCl_3$) 174.0, 171.1, 170.4, 170.3, 170.2, 169.5, 153.5, 133.7, 121.3, 117.4, 100.0, 71.0, 70.7, 68.7, 66.9, 61.4, 49.5, 36.2, 29.7, 29.5, 25.5, 20.8, 20.7, 18.1; MS (HRMS-TOF) $C_{58}H_{78}N_4O_{24}Na$, calculated m/z 1237.4904 $(M+Na)^+$, found m/z 1237.4891 $(M+Na)^+$.

Divalent compound, Gal–Glu-C10-Glu-Gal (6): Yield- 68% (1.17 g). Colourless solid: m. p. 176-178 °C; $[\alpha]_D$ -5.59° (*c* 0.2, $CHCl_3$); ^1H NMR (400 MHz, $CDCl_3$) δ 7.76 (d, 2H, Ph-H), 7.57 (m, 2H, Ph-H), 7.48 (m, 4H, Ph-H), 7.38 (m, 2H, Ph-H), 6.94 (m, 2H, Ph-H), 6.50 (d, 1H, N-H), 5.43-5.40 (m, 2H, H-2,H-4), 5.10-507 (dd, 1H, H-3), 4.92 (s, 1H, H-1), 4.39-4.34 (m, 1H, *tert* C-H of Glu and *tert* C-H of Fmoc), 4.21-4.14 (m, 4H, methylene proton of Fmoc, H-6a, H-6b), 4.01-3.98 (m, 1H, H-5), 3.20-3.19 (m, 2H, methylene protons of Glu), 2.37-2.31 (m, 2H, methylene protons of Glu), 2.20-2.02 (m, 15H, 4 x $COCH_3$ and methylene protons of long chain), 1.54-1.36 (m, 2H, methylene protons of long chain), 1.35-1.23 (m,

7H, methylene protons of long chain); $^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$) 172.2 ($COOCH_3$), 171.2 (NHCO), 170.2,169.8, 169.6, 169.1, 156.17 (NHCO of Fmoc), 153.4, 143.8, 141.3, 127.7, 127.1, 125.0, 121.3, 120.0, 117.5, 100.1 (C-1), 70.9, 70.0, 68.7, 67.9, 67.0 (CH_2 of Fmoc), 48.9 (*tert* C-H), 47.2 (C-H of Fmoc), 47.1, 40.0, 32.9, 30.9, 29.1, 28.9, 20.6, 20.5; MS (HRMS-TOF) $C_{90}H_{104}N_6O_{28}Na$, calculated m/z 1739.6796 $(M+Na)^+$, found m/z 1740.6906 $(M+Na)^+$.

Divalent compound, Gal-Glu-Cyclic core-Glu-Gal (7): Yield- 8% (0.133 g). Oily liquid: $[\alpha]_D$ -0.5° (*c* 2.5, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.00 (bs, 1H, N-H), 7.40 (d, 2H, Ph-H), 6.87 (d, 2H, Ph-H), 6.82-6.76 (bm, 1H, N-H), 5.42-5.35 (m, 2H, H-2, H-4), 5.06-5.03 (dd, 1H, H-3), 4.91 (s, 1H, H-1), 4.55-4.46 (m, 1H, *tert* C-H), 4.17-4.05 (m, 3H, H-6a, H-6b), 4.03-3.95 (m, 1H, H-5), 3.11-3.03 (q, 2H, methylene protons), 2.50-2.40 (m, 1H, methylene protons of Glu), 2.33-2.22 (m, 1H, methylene protons of Glu), 2.18-2.03 (m, 6H, methylene protons and $COCH_3$), 2.00 (s, 3H, $COCH_3$), 1.98 (s, 3H, $COCH_3$), 1.94 (s, 3H, $COCH_3$), 1.60-1.46 (m, 2H, methylene protons), 1.38 (bs, 9H, methylene protons), 1.29 (T, 4H, methylene protons), 1.23-1.15 (m, 10H, methylene protons); $^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$) 174.2, 173.3, 170.4, 170.3, 170.2, 169.7, 169.5, 153.6, 133.5, 121.3, 117.5, 100.0, 81.3, 71.0, 70.8, 68.7, 66.9, 61.4, 53.2, 45.9, 36.4, 31.9, 29.7, 28.9, 28.1, 27.4, 25.4, 20.8, 20.7, 20.6 8.1; MS (MALDI-TOF) $C_{72}H_{102}N_6O_{26}Na$, calculated m/z 1466.6844 $(M+Na)^+$, found m/z 1466.6240 $(M+Na)^+$.

Divalent compound, Lac-Glu-C12-Glu-Lac (8): Yield- 50% (1.04 g). Colorless solid: m. p. 108.7-119.4 °C; $[\alpha]_D$ -12.7° (*c* 0.2, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.44 (d, J = 8.8 Hz,

2H, 2,6-Ph-H), 7.38-7.30 (m, 5H, Ph-H), 6.92 (d, *J* = 8.8 Hz, 2H, 3,5-Ph-H), 6.75 (d, 1H, N-H), 5.38 (d, *J* = 3.0 Hz, 1H, H-4'), 5.29 (t, *J* = 9.3 Hz, 1H, H-3), 5.18-5.09 (m, 4H, H-2, H-2', $PhCH_2$), 5.02-4.94 (m, 2H, H-3', H-1), 4.64-4.60 (m, 1H, *tert* C-H), 4.54-4.49 (d, 1H, H-1', H-6_a), 4.19-4.07 (m, 3H, H-6_b, H-6_a', H-6_b'), 3.93-3.87 (m, 2H, H-4, H-5'), 3.80-3.77 (m, 1H, H-5), 2.69-2.41 (m, 2H, methylene protons), 2.30-1,99 (m, 26H, methylene protons and $COCH_3$), 1.80-1.54 (m, 6H, methylene protons), 1.27-1.15 (m, 8H, methylene protons); $^{13}C\{^1H\}$ NMR (100 MHz $CDCl_3$) 174.2, 173.5, 170.4, 170.3, 170.1, 169.8, 169.6, 169.5, 169.1, 153.5, 135.6, 133.3, 128.6, 128.4, 128.3, 121.3, 117.6, 101.1, 99.3, 76.2, 72.8, 71.5, 71.4, 71.0, 70.7, 69.1, 66.8, 66.7, 62.1, 60.8, 53.1, 36.3, 30.6, 28.7, 27.2, 25.3, 20.8, 20.7, 20.6, 20.5; MS (HRMS-TOF) $C_{100}H_{126}N_4O_{44}Na$, calculated m/z 2109.7643 $(M+Na)^+$, found m/z 2109.7866 $(M+Na)^+$; MS (MALDI-TOF) found m/z 2111.181 $(M+Na)^+$.

Trivalent compound, 19: Yield- 25 % (0.520 g). Colourless solid: m. p. 123-134 °C; $[\alpha]_D$ -2.8° (*c* 0.4, $CHCl_3$); ^1H NMR (400 MHz, $CDCl_3$) δ 7.49 (d, 8H, Ph-H), 6.97 (d, 8H, Ph-H), 5.50-5.40 (m, 8H, H-2, H-4), 5.16-5.10 (dd, 4H, H-3), 5.02 (s, 4H, H-1), 4.68-4.37 (m, 3H, *tert* C-H), 4.26-3.93 (m, 13H, H-6a, H-6b, H-5, *tert* C-H), 3.90-3.80 (q, 2H, methylene protons), 3.30-3.10 (m, 5H, methylene protons), 2.52-2.31 (m, 9H, methylene protons of Glu), 2.30-2.00 (m, 60H, methylene protons and 16 x $COCH_3$), 1.89-1.54 (m, 27H, methylene protons), 1.45 (s, 18H, $C(CH_3)_3$), 1.35-1.17 (m, 45H, methylene protons); $^{13}C\{^1H\}$ NMR (100 MHz, $CDCl_3$) 174.1, 171.3, 170.4, 170.3, 169.5, 162.5, 141.4, 133.4, 121.5, 121.3, 117.5, 100.1, 81.2, 71.0, 70.8, 68.7, 66.9, 66.7, 61.4, 42.2, 39.5, 36.5, 31.8, 31.4, 30.0, 29.7, 29.5, 29.0, 28.1, 23.5 20.8, 20.7, 20.6; MS (MALDI-TOF)

$C_{101}H_{142}N_8O_{39}Na$, calculated m/z 2091.936 (M+H)+, 2113.927 $(M+Na)^+$, found m/z 2093.732, 2115.803 $(M+Na)^+$, 2131.765 $(M+K)^+$.

Tetravalent compound, 20: Yield- 30 % (0.880 g). Colourless solid: m. p. 116-124 °C; $[\alpha]_D$ -5.5° (*c* 0.3, $CHCl_3$); ^1H NMR (400 MHz, $CDCl_3$) δ 7.49 (d, 8H, Ph-H), 6.97 (d, 8H, Ph-H), 5.50-5.40 (m, 8H, H-2, H-4), 5.16-5.10 (dd, 4H, H-3), 5.02 (s, 4H, H-1), 4.68-4.37 (m, 3H, *tert* C-H), 4.26-3.93 (m, 13H, H-6a, H-6b, H-5, *tert* C-H), 3.90-3.80 (q, 2H, methylene protons), 3.30-3.10 (m, 5H, methylene protons), 2.52-2.31 (m, 9H, methylene protons of Glu), 2.30-2.00 (m, 60H, methylene protons and 16 x $COCH_3$), 1.89-1.54 (m, 27H, methylene protons), 1.45 (s, 18H, $C(CH_3)_3$), 1.35-1.17 (m, 45H, methylene protons); $^{13}C\{^1H\}$ NMR (100 MHz, $CDCl_3$) 174.1, 171.3, 170.4, 170.3, 169.5, 162.5, 141.4, 133.4, 121.5, 121.3, 117.5, 100.1, 81.2, 71.0, 70.8, 68.7, 66.9, 66.7, 61.4, 42.2, 39.5, 36.5, 31.8, 31.4, 30.0, 29.7, 29.5, 29.0, 28.1, 23.5 20.8, 20.7, 20.6; MS (MALDI-TOF) $C_{142}H_{100}N_{10}O_{54}Na$, calculated m/z 2934.131 $(M+Na)^+$, found m/z 2935.908 $(M+Na)^+$.

Compound 27 (N^2, N^6-Bis (12-(Fmocamino-dodecanoyl)-Lys-4-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyloxy)-anilide): Yield - 40% (0.570 g). Colorless solid; m. p. 92-110 °C; $[\alpha]_D$ -14.9° (*c* 0.2, $CHCl_3$); ^1H NMR (400 MHz, $CDCl_3$) δ 7.77 (m, 4H, Ph-H), 7.58 (m, 4H, Ph-H), 7.49 (d, 2H, Ph-H), 7.43-7.39 (m, 4H, Ph-H), 7.33-7.28 (m, 4H, Ph-H), 6.95 (d, 2H, Ph-H), 5.60 (bd, 1H, N-H), 5.50-5.45 (m, 2H, H-2, H-4), 5.14-5.10 (dd, 1H, H-3), 4.97 (d, 1H, H-1), 4.95-4.80 (m, 1H, methylene protons of Fmoc), 4.60-4.51 (m, 1H, methylene protons of Fmoc), 4.41-4.30 (m, 3H, methylene proton of Fmoc and *tert* C-H), 4.24-4.17 (m, 4H, H-6a, H-6b, 2 x C-H of Fmoc), 4.05-4.00 (m, 1H, H-5), 3.30-3.00 (m, 6H, methylene

protons), 2.28-2.24 (m, 2H, methylene protons), 2.17 (s, 3H, $COCH_3$), 2.08 (s, 3H, $COCH_3$), 2.06 (s, 3H, $COCH_3$), 2.02 (s, 3H, $COCH_3$), 1.97-1.85 (m, 2H, methylene protons), 1.85-1.70 (m, 2H, methylene protons), 1.62-1.41 (m, 4H, methylene protons), 1.27 (bs, 30H, methylene protons); $^{13}C\{^1H\}$ NMR (100 MHz, $CDCl_3$), 174.4, 173.7, 170.4, 170.3, 170.1, 169.4, 156.5, 153.5, 144.0, 141.3, 133.7, 129.4, 127.7, 127.1, 125.1, 121.3, 120.1, 120.0, 117.5, 100.1, 71.0, 70.8, 68.7, 66.9, 66.5, 61.4, 53.5, 47.3, 41.1, 38.2, 36.8, 36.4, 31.6, 30.0, 29.7, 29.4, 29.3, 26.7, 25.8, 22.3, 20.8, 20.7, 20.6; MS (ESI HRMS-TOF) $C_{80}H_{103}N_5O_{17}Na$, calculated m/z 1428.7247 $(M+Na)^+$, found m/z 1428.3935 $(M+Na)^+$.

Compound 28 (N^2-(12-(Fmocamino)dodecanoyl)-N^6-(12-(12-(Fmocamino)dodecanamido)-dodecanoyl)-Lys-4-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyloxy)-anilide): Yield - 27% (0.450 g). Colorless solid; m. p. 88-110 °C; $[\alpha]_D$ -5.8° (*c* 0.8, $CHCl_3$); ^1H NMR (400 MHz, $CDCl_3$) δ 7.77 (m, 4H, Ph-H), 7.64-7.59 (m, 4H, Ph-H), 7.49 (d, 2H, Ph-H), 7.44-7.38 (m, 4H, Ph-H), 7.33-7.28 (m, 4H, Ph-H), 6.94 (d, 2H, Ph-H), 5.60 (bd, 1H, N-H), 5.49-5.45 (m, 2H, H-2, H-4), 5.13-5.10 (dd, 1H, H-3), 4.97 (d, 1H, H-1), 4.95-4.80 (m, 1H, methylene protons of Fmoc), 4.60-4.51 (m, 1H, methylene protons of Fmoc), 4.44-4.38 (m, 3H, methylene proton of Fmoc and *tert* C-H), 4.23-4.17 (m, 4H, H-6a, H-6b, 2 x C-H of Fmoc), 4.06-4.03 (m, 1H, H-5), 3.29-3.01 (m, 8H, methylene protons), 2.27-2.13 (m, 3H, methylene protons), 2.10-2.02 (s, 8H, $COCH_3$ and methylene protons), 1.97 (s, 3H, $COCH_3$), 1.96 (s, 3H, $COCH_3$), 1.94 (s, 3H, $COCH_3$), 1.92-1.80 (m, 2H, methylene protons), 1.79-1.60 (m, 1H, methylene protons), 1.61-1.30 (m, 17H, methylene protons), 1.27 (bs, 40H, methylene protons); $^{13}C\{^1H\}$ NMR (100 MHz, $CDCl_3$), 174.3, 173.8, 173.2, 170.3, 170.2, 170.1, 169.4, 156.5, 153.5, 144.0, 141.3, 133.7, 129.4,

128.1, 127.8, 127.1, 127.0, 125.2, 125.0, 121.3, 120.1, 120.0, 117.5, 100.1, 71.0, 70.8, 68.7, 66.9, 66.5, 61.4, 53.5, 47.3, 41.1, 38.2, 36.9, 36.8, 36.4, 30.0, 29.7, 29.6, 29.5, 29.4, 29.3, 29.2, 26.8, 26.7, 25.8, 25.7, 25.6, 22.3, 20.7, 20.7, 20.6; MS (ESI HRMS-TOF) $C_{92}H_{126}N_6O_{18}Na$, calculated m/z 1625.9026 $(M+Na)^+$, found m/z 1625.8971 $(M+Na)^+$.

Compound 29 (N^2, N^6-Bis-(12-(12-(Fmocamino)dodecanamido)-dodecanoyl)-Lys-4-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyloxy)-anilide): Yield - 5-7% (0.100 g). Colorless solid; m. p. 85-110 °C; $[\alpha]_D$ -4.9° (*c* 0.8, $CHCl_3$); ^1H NMR (400 MHz, $CDCl_3$) δ 7.82-7.76 (m, 4H, Ph-H), 7.61-7.59 (m, 4H, Ph-H), 7.54 (d, 2H, Ph-H), 7.43-7.37 (m, 4H, Ph-H), 7.34-7.28 (m, 4H, Ph-H), 6.96 (m, 2H, Ph-H), 5.60 (bd, 1H, N-H), 5.50-5.44 (m, 2H, H-2, H-4), 5.14-5.10 (dd, 1H, H-3), 5.00 (d, 1H, H-1), 4.95-4.80 (m, 2H, methylene protons of Fmoc), 4.65-4.44 (m, 1H, methylene protons of Fmoc), 4.48-4.37 (m, 3H, methylene proton of Fmoc and *tert* C-H), 4.28-4.01 (m, 5H, H-6a, H-6b, 2 x C-H of Fmoc, H-5), 3.84-3.50 (m, 3H, methylene protons), 3.27-3.06 (m, 6H, methylene protons), 2.41-2.23 (m, 3H, methylene protons), 2.12-2.02 (s, 11H, $COCH_3$ and methylene protons), 2.08 (s, 3H, $COCH_3$), 2.06 (s, 3H, $COCH_3$), 2.03 (s, 3H, $COCH_3$), 1.88-1.74 (m, 3H, methylene protons), 1.68-1.56 (m, 6H, methylene protons), 1.56-1.37 (m, 9H, methylene protons), 1.37-1.13 (bs, 40H, methylene protons); $^{13}C\{^1H\}$ NMR (100 MHz, $CDCl_3$), 174.4, 173.7, 173.2, 170.4, 170.3, 170.1, 169.4, 156.7, 153.6, 144.1, 141.3, 133.6, 129.5, 128.2, 127.9, 127.2, 127.0, 125.3, 125.1, 121.3, 120.1, 120.0, 117.6, 100.1, 71.0, 70.8, 68.6, 66.7, 66.4, 61.5, 53.6, 47.3, 41.1, 38.2, 36.9, 36.7, 36.8, 36.4, 30.0, 29.7, 29.6, 29.5, 29.4, 29.3, 29.2, 26.9, 26.8, 26.7, 25.8, 25.9, 25.7, 25.6, 22.3, 20.7, 20.7, 20.6; MS (ESI HRMS-TOF) $C_{104}H_{49}N_7O_{19}Na$, calculated m/z 1823.0806 $(M+Na)^+$, found m/z 1823.023 $(M+Na)^+$.

Compound 31 (N^2-(12-(Fmocamino)dodecanoyl)-C^5-*tert*-butyl-Glut-4-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyloxy)-anilide): Yield - 35% (0.360 g). Colorless solid: m. p. 75-81 °C; $[\alpha]_D$ -5.9° (*c* 1, $CHCl_3$); ^1H NMR (400 MHz, $CDCl_3$) δ 7.75 (d, 2H, Ph-H), 7.59 (m, 2H, Ph-H), 7.46 (m, 4H, Ph-H), 7.39 (t, 2H, Ph-H), 7.31 (t, 4H, Ph-H), 6.94 (d, 2H, Ph-H), 6.74 (d, 1H, N-H), 5.58-5.41 (m, 2H, H-2, H-4), 5.12-5.08 (dd, 1H, H-3), 4.97 (s, 1H, H-1), 4.84 (bs, 1H, N-H), 4.62-4.48 (m, 1H, *tert* C-H), 4.46-4.31 (m, 2H, methylene protons of Fmoc), 4.30-4.11 (m, 3H, H-6a, H-6b, *tert* C-H), 4.10-4.00 (m, 1H, H-5), 3.26-3.07 (m, 3H, methylene protons), 2.61-2.44 (m, 1H, methylene protons of Glu), 2.41-2.26 (m, 1H, methylene protons of Glu), 2.25-2.09 (m, 8H, methylene protons and $COCH_3$), 2.07 (s, 3H, $COCH_3$), 2.05 (s, 3H, $COCH_3$), 2.01 (s, 3H, $COCH_3$), 1.67-1.54 (m, 4H, methylene protons), 1.54-1.35 (t, 6H, methylene protons), 1.36-1.16 (m, 10H, methylene protons); $^{13}C\{^1H\}$ NMR (100 MHz, $CDCl_3$) 174.0, 173.3, 173.1, 170.3, 170.2, 169.7, 169.4, 156.5, 153.5, 144.0, 141.34, 133.5, 127.6, 127.0, 125.0, 121.3, 119.5, 117.5, 100.0, 81.2, 71.0, 70.8, 68.7, 66.9, 66.5, 61.4, 53.2, 47.3, 41.1, 39.5, 36.8, 36.5, 31.9, 29.9, 29.7, 29.5, 29.4, 29.3, 29.2, 29.1, 28.1, 27.3, 26.9, 26.7, 25.8, 25.5 20.7, 20.6, 20.5; MS (HRMS-TOF) $C_{56}H_{73}N_3O_{16}K$, calculated m/z 1082.4628 $(M+K)^+$, found m/z 1083.6021 $(M+K)^+$.

Compound 32 (N^2-(12-(12-(Fmocamino)dodecanamido)-dodecanoyl)-C^5-*tert*-butyl-Glut-4-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyloxy)-anilide): Yield - 20% (0.250 g). Colorless solid: m. p. 73-81 °C; $[\alpha]_D$ -6.3° (*c* 1, $CHCl_3$); ^1H NMR (400 MHz, $CDCl_3$) δ 7.76 (d, 2H, Ph-H), 7.59 (m, 2H, Ph-H), 7.45 (m, 4H, Ph-H), 7.39 (t, 2H, Ph-H), 7.30 (t, 4H, Ph-H), 6.94 (d, 2H, Ph-H), 6.80 (d, 1H, N-H), 5.48-5.43 (m, 3H, H-2, H-4, N-H), 5.12-5.09 (dd,

1H, H-3), 4.96 (s, 1H, H-1), 4.89 (bs, 1H, N-H), 4.62-4.51 (m, 1H, *tert* C-H), 4.45-4.34 (m, 2H, methylene protons of Fmoc), 4.30-4.10 (m, 3H, H-6a, H-6b, *tert* C-H), 4.04 (t, 1H, H-5), 3.24-3.15 (m, 3H, methylene protons), 2.57-2.43 (m, 1H, methylene protons of Glu), 2.40-2.27 (m, 1H, methylene protons of Glu), 2.27-2.09 (m, 8H, methylene protons and $COCH_3$), 2.07 (s, 3H, $COCH_3$), 2.05 (s, 3H, $COCH_3$), 2.01 (s, 3H, $COCH_3$), 1.69-1.53 (m, 4H, methylene protons), 1.53-1.37 (t, 12H, methylene protons), 1.36-1.15 (m, 25H, methylene protons); $^{13}C\{^1H\}$ NMR (100 MHz, $CDCl_3$) 174.0, 173.3, 173.1, 170.3, 170.2, 169.7, 169.4, 156.5, 153.5, 144.0, 141.34, 133.5, 127.6, 127.0, 125.0, 121.3, 119.5, 117.5, 100.0, 81.2, 71.0, 70.8, 68.7, 66.9, 66.5, 61.4, 53.2, 47.3, 41.1, 39.5, 36.8, 36.5, 31.9, 29.9, 29.7, 29.5, 29.4, 29.3, 29.2, 29.1, 28.1, 27.3, 26.9, 26.7, 25.8, 25.5 20.7, 20.6, 20.5; MS (HRMS-TOF) $C_{68}H_{96}N_4O_{17}Na$, calculated m/z 1263.6668 $(M+Na)^+$, found m/z 1263.6660 $(M+Na)^+$.

Compound 33 (*N*^{2}-(12-(12-(12-(Fmocamino)-dodecanamido)-dodecanamido)-dodecano -yl)-*C*^{5}-*tert*-butyl-Glut-4-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyloxy)-anilide):

Yield - 5-8% (0.060 g). Colorless solid: m. p. 70-81 °C; $[\alpha]_D$ -6.4° (*c* 1, $CHCl_3$); ^1H NMR (400 MHz, $CDCl_3$) δ 7.76 (d, 2H, Ph-H), 7.59 (m, 2H, Ph-H), 7.46 (m, 4H, Ph-H), 7.40 (m, 2H, Ph-H), 7.30 (m, 4H, Ph-H), 6.95 (d, 2H, Ph-H), 5.53-5.40 (m, 3H, H-2, H-4, N-H), 5.12-5.08 (dd, 1H, H-3), 4.97 (s, 1H, H-1), 4.82 (bs, 1H, N-H), 4.59-4.47 (m, 1H, *tert* C-H), 4.46-4.34 (m, 2H, methylene protons of Fmoc), 4.30-4.10 (m, 3H, H-6a, H-6b, *tert* C-H), 4.10-3.99 (m, 1H, H-5), 3.26-3.11 (q, 4H, methylene protons), 2.63-2.50 (m, 1H, methylene protons of Glu), 2.41-2.26 (m, 1H, methylene protons of Glu), 2.25-2.10 (m, 6H, methylene protons and $COCH_3$), 2.07 (s, 3H, $COCH_3$), 2.06 (s, 3H, $COCH_3$), 2.01 (s, 3H, $COCH_3$),

1.70-1.54 (m, 2H, methylene protons), 1.54-1.35 (t, 4H, methylene protons), 1.35-1.17 (m, 8H, methylene protons); $^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$) 174.1, 173.4, 173.2, 170.4, 170.3, 169.7, 169.5, 156.6, 153.6, 144.1, 141.4, 133.4, 127.6, 127.2, 125.3, 121.2, 119.6, 117.5, 100.1, 81.3, 71.0, 70.8, 68.6, 66.9, 66.4, 61.4, 53.3, 47.7, 41.2, 39.5, 36.9, 36.6, 32.0, 30.0, 29.9, 29.8, 29.6, 29.5, 29.3, 29.2, 29.1, 28.1, 27.3, 26.9, 26.7, 25.8, 25.5 20.7, 20.6, 20.5; MS (HRMS-TOF) $C_{80}H_{119}N_5O_{18}Na$, calculated m/z 1460.8448 $(M+Na)^+$, found m/z 1460.8417 $(M+Na)^+$.

Cyclic compound 38: Yield- 15-20% (0.50 g). oily liquid: $[\alpha]_D$ -7.6° (*c* 1, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) and $^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$) spectra are same as that for cyclic pentavalent compound **41**, but, mass value was:MS (MALDI-TOF) $C_{168}H_{196}N_4O_{28}Na$, calculated m/z 2740.3934 $(M+Na)^+$, found m/z 2744.6539 $(M+Na)^+$.

Cyclic compound 39: Yield- 12% (0.420 g). oily liquid: $[\alpha]_D$ +31.9° (*c* 1, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) and $^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$) spectra are same as that for cyclic pentavalent compound **41**, but, mass value obtained was: MS (MALDI-TOF) $C_{253}H_{296}N_6O_{41}Na$, calculated m/z 4097.1159 $(M+Na)^+$, found m/z 4101.191 $(M+Na)^+$.

Cyclic compound 40: Yield- 5% (0.250 g). oily liquid: $[\alpha]_D$ -7.9° (*c* 1, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) and $^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$) Spectra are same as that for cyclic pentavalent compound **41**, but, mass value obtained was: MS (MALDI-TOF) $C_{337}H_{394}N_8O_{56}Na$, calculated m/z 5471.8126 $(M+Na)^+$, found m/z 5470.975 $(M+Na)^+$.

Cyclic compound 41: Yield- 1% (0.032 g). oily liquid: $[\alpha]_D$ +31.7° (*c* 1, $CHCl_3$); ^1H NMR (400 MHz, $CDCl_3$) δ 7.71-7.65 (m, 1H, N-H), 7.34-7.17 (m, 20H, Ph-H), 5.03 (d, 1H, J = 11.6 Hz, CH_2 of Benzyl), 4.86-4.77 (m, 1H, J = 11.2 Hz, H-2, CH_2 of Benzyl), 4.76-4.59 (m, 4H, CH_2 of Benzyl), 4.58-4.46 (m, 2H, H-1, CH_2 of Benzyl), 4.45-4.33 (d, 1H, CH_2 of Benzyl), 4.18-3.96 (m, 3H, H-3, H-4, CH_2 of Benzyl), 3.90-3.80 (m, 1H, H-6a), 3.68-3.52 (m, 2H, H-5, H-6b), 3.09-2.95 (m, 1H, methylene protons), 2.81-2.67 (m, 1H, methylene protons), 1.31-0.91 (m, 10H, methylene protons); $^{13}C\{^1H\}$ NMR (100 MHz, $CDCl_3$) 168.9 (NHCO), 138.7, 138.1, 128.5, 128.4, 128.3, 128.2, 128.1, 127.9, 127.8, 127.7, 127.6, 127.5, 127.4, 127.3, 98.8 (C-1), 80.0, 78.3, 75.8, 75.0, 74.7, 73.9, 73.5, 72.1, 69.8, 67.4, 39.3, 30.0, 29.9, 29.8, 29.6, 27.1; MS (MALDI-TOF) $C_{421}H_{492}N_{10}O_{70}Na$, calculated m/z 6835.4260 $(M+Na)^+$, found m/z 6838. 999 $(M+Na)^+$.

Conclusions

In conclusion, a set of novel galactose- and lactose-based multivalent molecules having amino acid-linked fatty acid chain as the linker unit between the carbohydrate moieties have been prepared in their protected form as potential CT inhibitors. A novel selective Fmoc deprotection observed during this work has the potential to be used as a method to prepare glycopeptides with extended polyamide chains, molecules that could possess interesting supramolecular self-assembled structures.

Acknowledgements

Vajinder Kumar sincerely acknowledges the award of a Research Fellowship by CSIR-UGC, New Delhi, India.

Notes and references

1 WHO Report: http://www.who.int/wer/2013/wer8831.pdf?ua=1, **2014**.

2 Fan, E.; Merrit, E. A.; Verlinde, C. L. M. J.; Hol, W. G. J. *Curr. Opin. Struct. Biol.*, **2000,** *10*, 680-686.

3 Merritt, E. A.; Hol, W. G. J. *Curr. Opin. Struct. Biol.*, **1995**, *5*, 165-171.

4 Hol, W. G. J.; Sixma, T. K.; Merritt, E. A. *Bacterial Toxins and Virulence Factors in Disease*, **1995**, *8*, 185-223.

5 Merritt, E. A.; Sarfaty, S.; Van den Akker, F.; L'hoir, C.; Martial, J. A.; Hol, W. G. J. *Protein Sci.*, **1994**, *3*, 166-175.

6 Verlinde, C. L. M. J.; Hol, W. G. J. *Structure*, **1994**, *2*, 577-587.

7 Kitov, P. I.; Sadowska, J. M.; Mulvey, G.; Armstrong, G. D.; Ling, H.; Pannu, N. S.; Read, R. J.; Bundle, D. R. *Nature*, **2000**, *403*, 669-672.

8 Thomas R. Branson, T. R.; Turnbull, W. B. *Chem. Soc. Rev*., **2013**, *42*, 4613-4622.

9 Branson, T. R.; McAllister, T. E.; Garcia-Hartjes, J.; Fascione, M. A.; Ross, J. F.; Warriner, S. L.; Wennekes, T.; Zuilhof, H.; Turnbull, W. B. *Angew. Chem. Int. Ed.* **2014**, *53*, 8323-8327.

10 Fu, O.; Pukin, A. V.; Ufford, H. C. Q. V.; Branson, T. R.; Thies-Weesie, D. M.E.; Turnbull, W. B.; Visser, G. M.; Pieters R. J. *ChemistryOpen*, **2015**, *4*, 471-477.

11 Zhang, Z.; Liu, J.; Verlinde, C. L. M. J.; Hol, W. G. J.; Fan, E. *J. Org. Chem.*, **2004**, *69*, 7737-7740.

12 Fan, E.; Zhang, Z.; Minke, W. E.; Hou, Z.; Verlinde, C. L. M. J.; Hol, W. G. J. *J. Am. Chem. Soc.*, **2000**, *122*, 2663-2664.

13 Kumar, V.; Venugopalan, P.; Kartha, K. P. R. *ChemPlusChem*, **2014**, *79*, 1605-1613.

14 Patil, P. R.; Kartha, K. P. R. *Green Chem.*, **2009**, *11*, 953-957.

15 Han, S. Y.; Kim, Y. A. *Tetrahedron*, **2004**, *60*, 2447-2467; Williams, A.; Ibrahim, I. T. *Chem. Rev.*, **1981**, *81*, 589-636; Carpino, L. A.; El-Faham, A. *Tetrahedron*, **1999**, *55*, 6813-6830; M. Bodansky, Principles of Peptide Synthesis; Springer-Verlag: New York, **1984**.

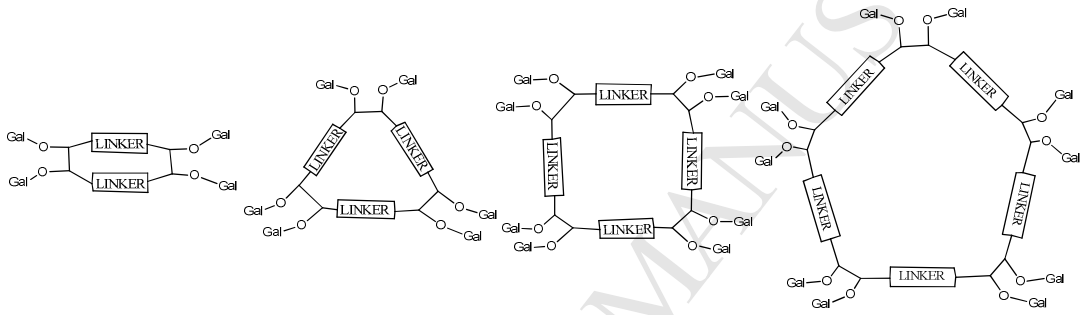
16 Please see the ESI for details.

17 Kumar, V.; Taxak, N.; Jangir, R.; Bharatam, P. V.; Kartha, K. P. R. *J. Org. Chem.*, **2014**, *79*, 3427-3439.

Graphical abstract

Synthetic multivalent ligands for cholera & cholera-like toxins: Protected cyclic neoglycopeptides

Vajinder Kumar, Narender Yadav and K. P. Ravindranathan Kartha



Captions to figures:

Figure 1. A cartoonic representation of the proposed glycoconjugates as potential polyvalent protein inhibitors

Figure 2. Structure of a previously prepared cross-linked galactopyranoside

Figure 3. Various divalent compounds synthesized in the protected form and the building blocks used for their synthesis

Figure 4. Synthetic tri- and tetravalent neoglycopeptide ligands

Figure 5. Unexpected by-products **28** and **29** in the preparation of the neoglycopeptide **27** using building blocks **23**/**26**

Figure 6. Unexpected by-products **32** and **33** in the preparation of the neoglycopeptide **31** using building blocks **30** and **26**

Figure 7. Synthetic multivalent neoglycopeptide ligands as potential CT inhibitors prepared in their protected form and the building blocks used for their preparation

Figures:

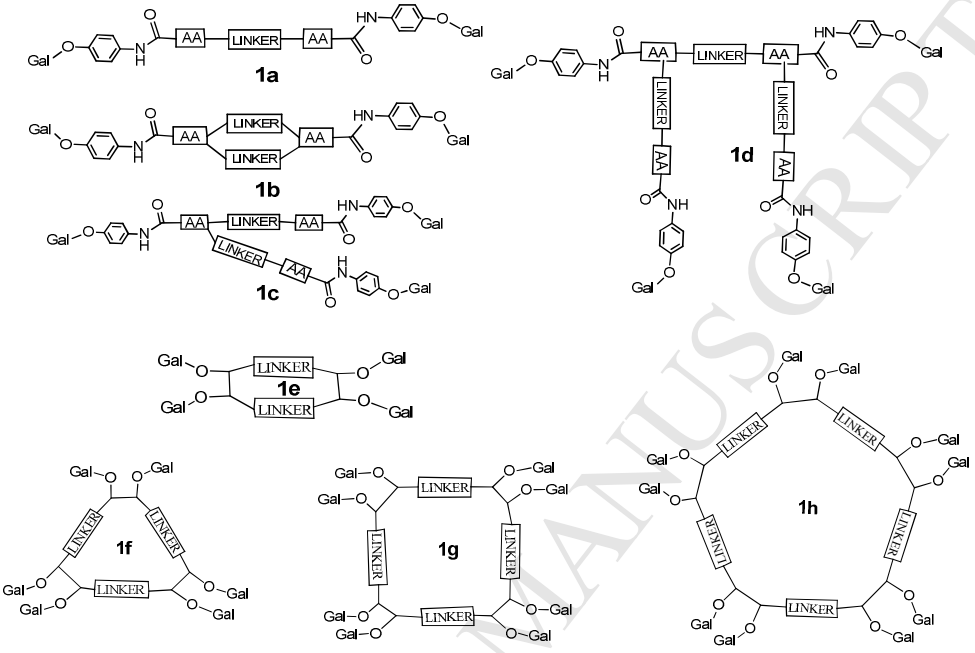


Figure 1. A cartoonic representation of the proposed glycoconjugates as potential polyvalent protein inhibitors

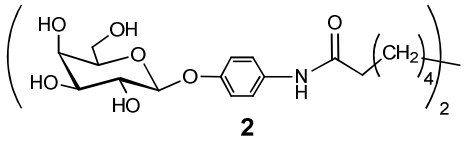


Figure 2. Structure of a previously prepared cross-linked galactopyranoside

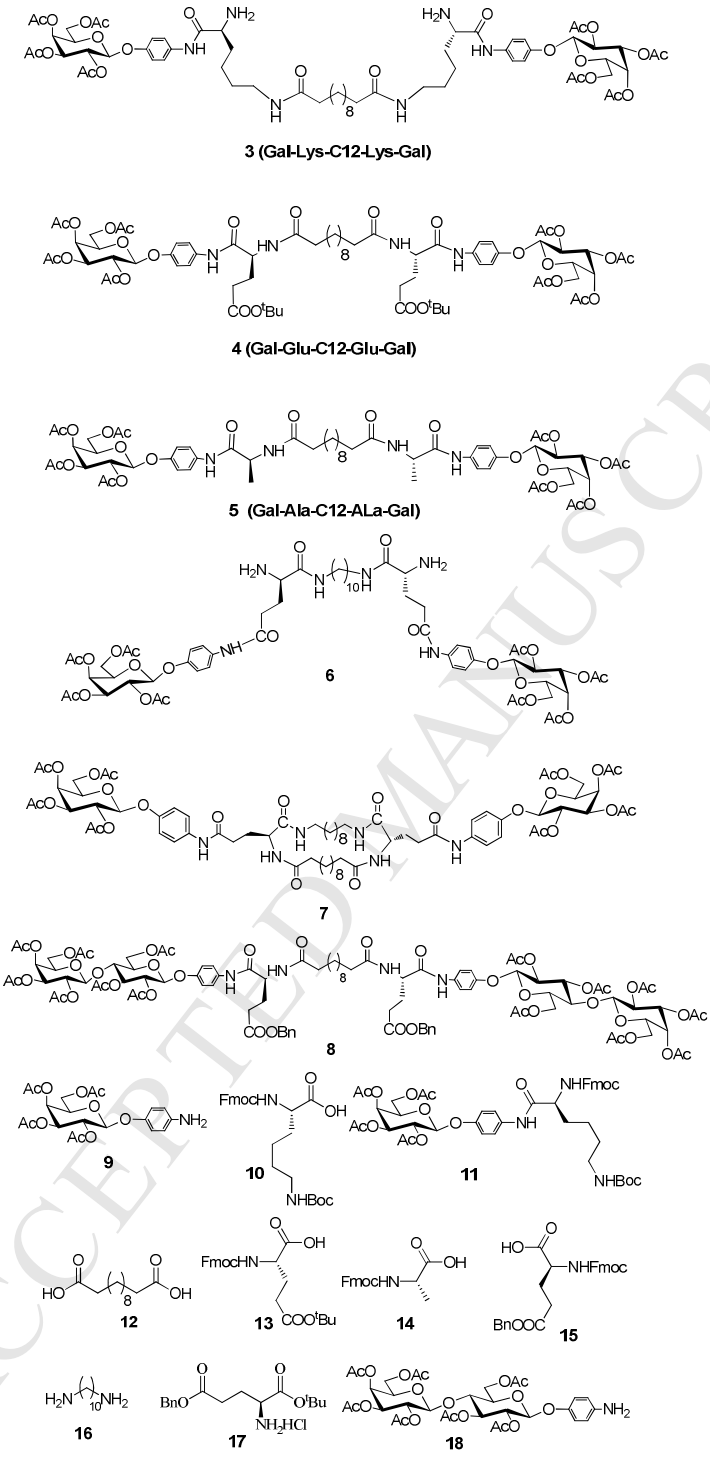


Figure 3. Various divalent compounds synthesized in the protected form and the building blocks used for their synthesis

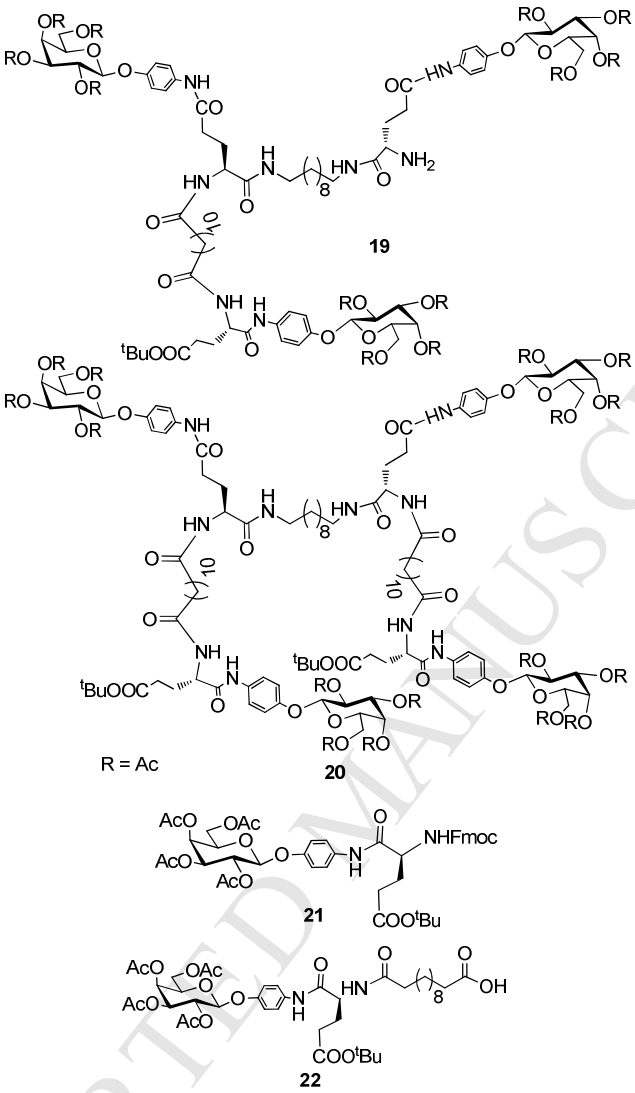


Figure 4. Synthetic tri- and tetravalent neoglycopeptide ligands

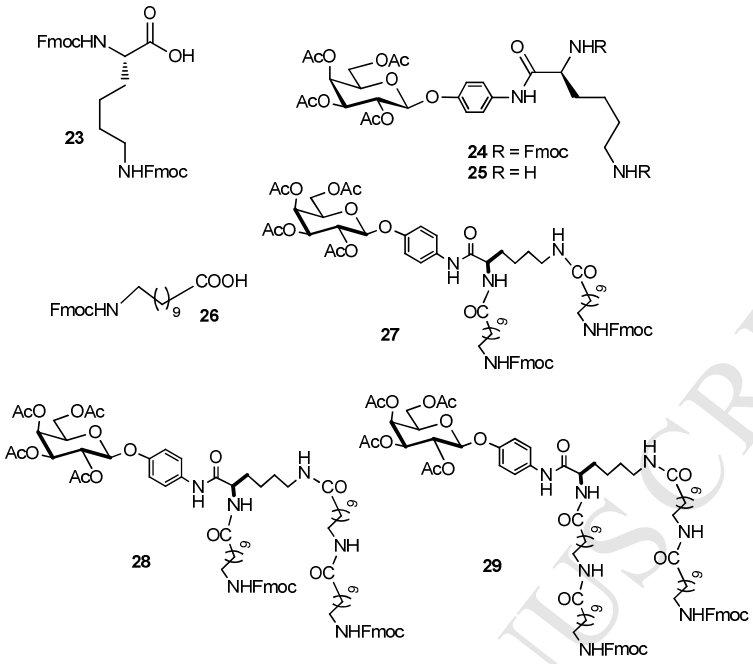


Figure 5. Unexpected by-products **28** and **29** in the preparation of the neoglycopeptide **27** using building blocks **23/26**

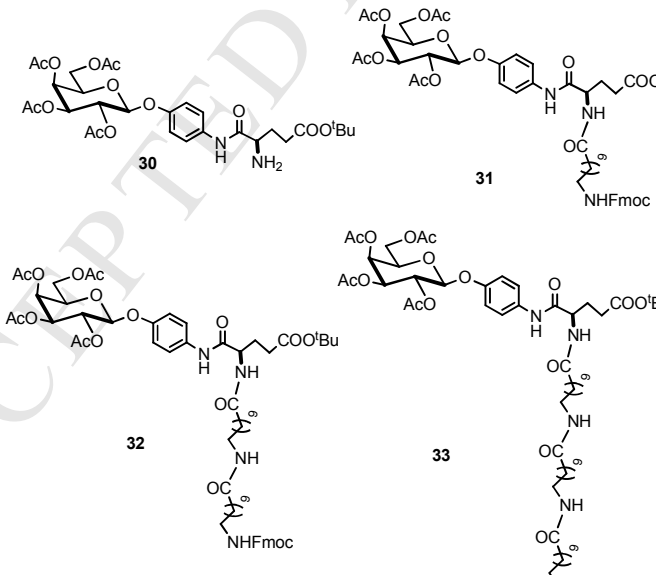


Figure 6. Unexpected by-products **32** and **33** in the preparation of the neoglycopeptide **31** using building blocks **30** and **26**

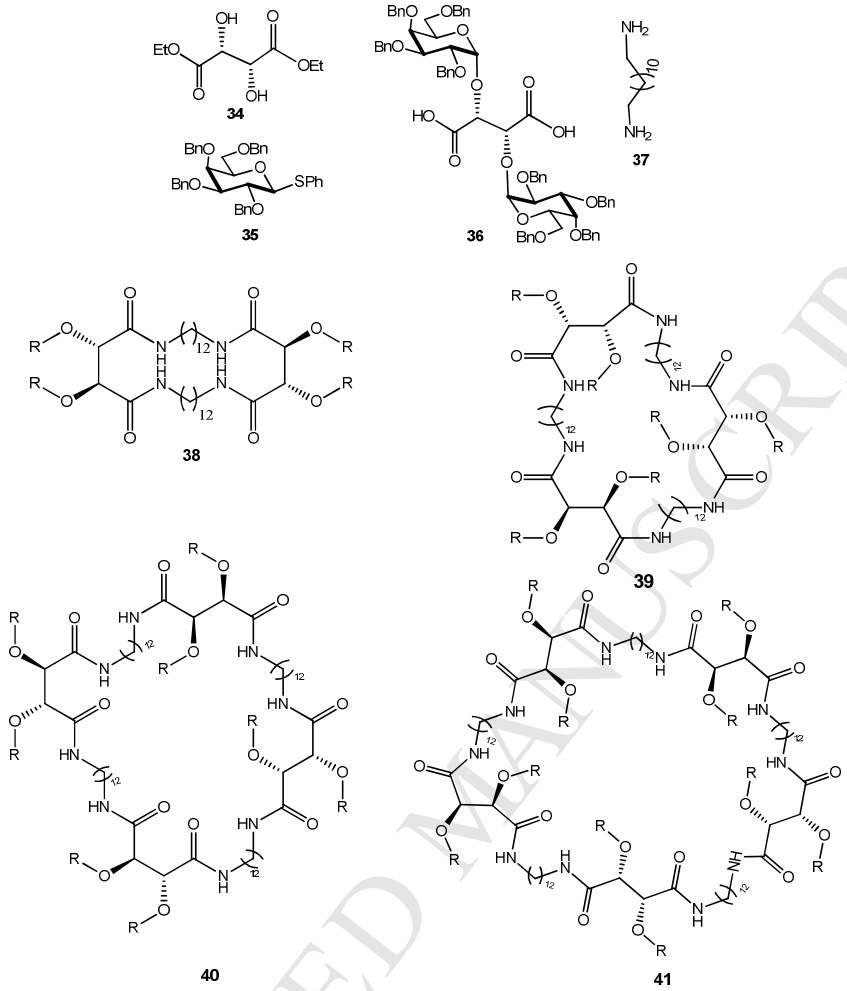


Figure 7. Synthetic multivalent neoglycopeptide ligands as potential CT inhibitors prepared in their protected form and the building blocks used for their preparation.