Bioorganic & Medicinal Chemistry xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Synthetic 1,2,3-triazole-linked glycoconjugates bind with high affinity to human galectin-3

Marcelo Fiori Marchiori [a], Dênio Emanuel Pires Souto [b,†], Leandro Oliveira Bortot [a,†], João Francisco Pereira [a], Lauro Tatsuo Kubota [b], Richard D. Cummings [c], Marcelo Dias-Baruffi [a], Ivone Carvalho [a], Vanessa Leiria Campo [a,*]

[a] Faculdade de Ciências Farmacêuticas de Ribeirão Preto, USP, Av. Café S/N, CEP: 14040-903, Ribeirão Preto, SP, Brazil
[b] Instituto de Química-Unicamp, R. Josué de Castro s/n°, Cidade Universitária, CEP: 13083-861, Campinas, SP, Brazil
[c] Department of Biochemistry, Emory University School of Medicine, Atlanta, GA 30322, USA

ARTICLE INFO

Article history:
Received 31 January 2015
Revised 2 April 2015
Accepted 13 April 2015
Available online xxxx

Keywords:
Galectin-3
Triazole
Glycoconjugates
'Click chemistry'
Surface plasmon resonance
Molecular modeling

ABSTRACT

This work describes the synthesis of the 1,2,3-triazole amino acid-derived-3-*O*-galactosides **1**–**6** and the 1,2,3-triazole di-lactose-derived glycoconjugate **7** as potential galectin-3 inhibitors. The target compounds were synthesized by Cu(I)-catalyzed azide–alkyne cycloaddition reaction ('click chemistry') between the azido-derived amino acids N_3-ThrOBn, N_3-PheOBn, N_3-*N*-Boc-TrpOBn, N_3-*N*-Boc-LysOBn, N_3-*O*-*t*Bu-AspOBn and N_3-L-TyrOH, and the corresponding alkyne-based sugar 3-*O*-propynyl-GalOMe, as well as by click chemistry reaction between the azido-lactose and 2-propynyl lactose. Surface plasmon resonance (SPR) assays showed that all synthetic glycoconjugates **1**–**7** bound to galectin-3 with high affinity, but the highest binders were the amino acids-derived glycoconjugates **2** (K_D 7.96 μM) and **4** (K_D 4.56 μM), and the divalent lactoside **7** (K_{D1} 0.15 μM/K_{D2} 19 μM). Molecular modeling results were in agreement with SPR assays, since more stable interactions with galectin-3 were identified for glycoconjugates **2**, **4** and **7**. Regarding compounds **2** and **4**, they established specific cation–π (Arg144) and ionic (Asp148) interactions, whereas glycoconjugate **7** was capable to bridge two independent galectin-3 CRDs, creating a non-covalent cross-link between two monomers and, thus, reaching a submicromolar affinity towards galectin-3.

1. Introduction

Galectins are an ancient family of glycan-binding proteins, being found in all animals, and are defined based on their affinity for β-galactosides-containing saccharides and amino acid sequence similarity in their CRDs (Carbohydrate Recognition Domain).[1–3] There are currently 15 described members of the galectin family in mammals and these proteins are expressed by a variety of cell types, and can participate in cell–cell and cell–extracellular matrix interactions, besides modulating several intracellular and extracellular pathways.[2–6] Thus, galectins may play an essential role in the development of inflammatory response, autoimmune diseases, atherosclerosis, infectious processes and cancer.[3–13] Galectin-3, for instance, is overexpressed in several types of tumor cells and is involved in important functions related to cancer, such as tumorigenesis, neoplastic transformation, tumor cell survival, angiogenesis, tumor metastasis and regulation of apoptosis.[7,10,11] Moreover, it can help tumors to escape from immune surveillance through modulation of immune and inflammatory responses.[12,13] Galectin-3 is therefore associated with the development and malignancy of colorectal, gastric, lung and breast cancers, pancreatic and hepatocellular carcinomas, melanoma and glioblastomas, among others, where it acts through various molecular mechanisms depending on the cancer cell type.[7–14] For example, galectin-3 can mediate neoplastic transformation by interacting with oncogenic K-RAS gene and promoting RAS-mediated signal transduction. It can also direct tumor progression by controlling the levels of regulators of cell cycle progression and cell proliferation, such as cyclin D1, c-MYC, and β-catenin, and exerting antiapoptotic functions by repressing the tumor suppressor p53 gene that induces apoptosis.[5,14] Therefore, considering all these tumor-promoting effects of galectin-3, synthetic galectin-3 inhibitors are of utmost relevance for development of new anti-tumor therapeutic strategies.[15]

* Corresponding author. Tel.: +55 16 33154285; fax: +55 16 33154879.
E-mail address: vlcampo@fcfrp.usp.br (V.L. Campo).
† These authors contributed equally to this manuscript.

http://dx.doi.org/10.1016/j.bmc.2015.04.044
0968-0896/

CRDs of galectins contain approximately 135 amino acid (aa) residues which directly lead to the specificity of galectins for saccharides and are composed by subsites (A–E) which, along their 135 aa, form a groove able to bind up to a tetrasaccharide.[4,16,17] Galactose binds to the most conserved subsite C, whereas the second more important subsite D is occupied by another pyranoside represented by (1→4)Glc/GlcNAc or (1→3) GlcNAc/GalNAc linked to galactose (subsite C).[4,17,18] Interestingly, the preferential interaction of these subsites of each galectin CRD with different carbohydrates illustrate the diversity in their binding specificity and biological activities.[4–19] Depending on their structural organization, mammalian galectins can be subdivided into three groups known as prototypical, tandem-repeat and chimeric. Prototypical galectins have one CRD and exist either as monomers or as non-covalent dimmers (galectins-1, -2, -7, -13 and -14); tandem-repeat galectins have two non-identical CRDs (galectins-4, -6, -8, -9 and -12); and the only chimeric galectin is galectin-3, which contains one C-terminal CRD and N-terminal extension constituted by repeated collagen-like sequences of proline and glycine. Galectin-3 is unique since it can associate, through its amino-terminal extension, to form oligomers in the presence of multivalent carbohydrate ligands.[1,3,4]

The CRD of galectin-3 is comprised of eight conservative amino acids (Arg144, His158, Asn160, Arg162, Asn174, Trp181, Glu184 and Arg186) responsible for the lectin binding to carbohydrates. In general, the main interactions of galectin-3 to the natural disaccharide ligands Lac/LacNAc are represented by hydrogen bonds between the OH groups of Gal (C-4 and C-6) and Glc/ GlcNAc (C-3) with His158, Asn160, Arg162, Glu184 and Asn174, and by Van der Waals contacts of Gal and Glc/GlcNAc residues with Trp181 and Arg186.[20] Therefore, chemical modifications of natural ligands, such as at C-3 of Gal and at C-1 of Glc/GlcNAc can maintain the cited interactions and increase additional contacts, such as with Arg144 (subsite B).[20–23] In this regard, the synthesis of anomeric and O-3 modified analogs of natural monosaccharide (Gal) and disaccharides (Lac/LacNAc), especially those containing aromatic groups on the galactose C3 position, has been explored by Nilsson et al. and others, since they are able to establish favorable cation–π interactions with arginine residues, leading to galectin-3 inhibition.[24–29] The synthesis of carbohydrate based 1,2,3-triazole analogs as galectin-3 inhibitors by Cu(I)-assisted 1,3-dipolar azide–alkyne cycloaddition (CuAAC) has also been greatly developed, considering the advantageous physicochemical properties of the triazole linkage between the carbohydrate and the aglycone moiety, if compared to an ether, ester or amide linkage.[15,17,29–31] Thus, the stability of triazole toward oxidation, reduction and hydrolysis, its property to acts as a rigid link, besides the fact that CuAAC reactions from suitable azide and alkyne precursors are, generally, easily executed, fast and highly selective, have contributed to the increasing interest in the CuAAC synthesis of 1,2,3-triazole-linked carbohydrates as galectin-3 inhibitors.[28,32]

We have previously described the synthesis of the amino acid derivative 1,2,3-triazole fused threonine-3-*O*-galactose **1** (Fig. 1) by CuAAC reaction, and its activity as an inhibitor of *Trypanosoma cruzi trans*-sialidase (TcTS) enzyme.[32] However, considering the potential of *O*-3 triazole-galactose analogs to interact with galectin-3 CRD, we asked whether or not compound **1** should be tested as potential galectin-3 inhibitor, as well as other 1,2,3-triazole amino acids-derived-3-*O*-galactosides, such as compounds **2**–**6**, derived from phenylalanine, tryptophan, lysine, aspartic acid and tyrosine amino acids, respectively (Fig. 1), which could allow us to investigate the influence of the galactose scaffold bearing different side chains on galectin-3 inhibition. Interestingly, galectin-3 can form lattice structures upon binding of multivalent carbohydrates,[5] in connection to the concept of 'glycoside cluster effect',[15] and thus, the lectin binding affinity observed for multimeric ligands is often greater than for monomeric species, depending on their steric bulk, relative distance and three-dimensional arrangement.[33,34] These considerations led us to also design the divalent lactoside represented by 1,2,3-triazole di-lactose-derived glycoconjugate **7** as a potential galectin-3 inhibitor (Fig. 1). Therefore, here we describe the synthesis of the 1,2,3-triazole-linked glycoconjugates **1**–**7** by CuAAC reactions and their corresponding binding affinity to galectin-3 performed by surface plasmon resonance (SPR) assays. The molecular interactions between glycoconjugates **1**–**7** and galectin-3 by molecular modeling studies will also be reported herein.

2. Results and discussion

2.1. Synthesis

The target glycoconjugates are represented by the amino acids derivatives 1,2,3-triazole linked threonine- (**1**),[32] phenylalanine- (**2**), tryptophan- (**3**), lysine- (**4**), aspartic acid- (**5**) and tyrosine- (**6**) 3-*O*-galactose, and by the 1,2,3-triazole di-lactose-derived glycoconjugate **7** (Fig. 1). The synthesis of compounds **1**–**7** by CuAAC reaction started from the previously prepared azido-derived amino acids N_3-L-ThrOBn **8**,[32] N_3-L-PheOBn **9**,[35] N_3-L-*N*-Boc-TrpOBn **10**, N_3-L-*N*-Boc-LysOBn **11**, N_3-L-*O*-*t*Bu-AspOBn **12** and N_3-L-TyrOH **13**,[36] and by the sugar azido-lactose **14**,[37] as well as the use of the corresponding alkyne-based sugars 3-*O*-propynyl-GalOMe **15**[32,38] and 2-propynyl-lactose **16**.[38]

2.1.1. Synthesis of azido- and alkyne-functionalized precursors

With the aim to get a similar triazole pattern for compounds **1**–**6**, the α-NH_2 groups of the precursor amino acids were selected for conversion to azido group, allowing us to compare the distinct interactions of their natural side chains with galectin-3 CRD. Thus, the azido-functionalized amino acids **8**–**12** were prepared by treatment of their corresponding *N*-Fmoc *O*-benzyl esters **17**–**21** with 50% morpholine in DMF for removal of the *N*-Fmoc group, followed by diazo transfer reaction, utilizing in situ generated triflyl azide.[32,39,40] Compounds **8**–**12** were obtained in yields varying from 50% to 78%, and were characterized by ^{1}H NMR and IR [2100 cm^{-1} (N_3)] spectra. Regarding compound **13**, since the preliminary carboxyl protection reaction using benzyl bromide would also lead to the undesired benzylation of *p*OH-phenyl side chain, it was prepared directly by diazo transfer reaction of the commercial NH_2-L-TyrOH **22**, being obtained in quantitative yield (IR [2100 cm^{-1} (N_3)]. The alkyne-based sugar 3-*O*-propynyl-GalOMe **15** was synthesized from commercial α-GalOMe, previously described, being obtained in 70% yield.[32] The synthesis of the sugar azido-lactose **14** and 2-propynyl-lactose **16** were carried out by glycosylation reactions of 2-[2-(2-azidoethoxy)ethoxy]ethanol **23** or propargyl alcohol **24** with peracetyllactose **25**,[41] in the presence of $BF_3{\cdot}Et_2O$ as catalyst and DCM as solvent, affording compounds **14** (36%) or **16** (54%), respectively, with exclusive β configuration.[42]

2.1.2. Synthesis of 1,2,3-triazole-linked glycoconjugates

In general, the synthesis of compounds **1**–**7** by Cu(I)-assisted 1,3-dipolar azide-alkyne cycloaddition reactions were performed in a microwave reactor utilizing the catalytic system $CuSO_4$/sodium ascorbate and DMF as solvent.[32,43,44] The progress of the reactions was followed by TLC analysis, which revealed consumption of starting materials after 15 min irradiation bursts (1–3) at 100 °C. Thus, the condensation of azido-functionalized amino acids **8**–**13** with 3-*O*-propynyl-GalOMe **15** afforded the peracetylated products **26** (37%), **27** (35%), **28** (45.5%), **29** (47%), **30** (34%) and **31** (40%) after purification by column chromatography (Scheme 1), with regioselective formation of only 1,4-disubstituted