1. Introduction

Carbohydrates have rarely been a matter of research by medicinal chemistry-oriented scientists in the past, but has recently gained substantially more publicity [1]. Although equally or even more abundant in nature than peptides/proteins or nucleic acids, they were often simply neglected as potential drug targets and/or drug leads or even drug candidates. There are a number of reasons for this. Glycans – complex carbohydrates linked to proteins or lipids are essential components of every cell surface as all cells are coated with a so-called glycocalyx layer. These cell surface carbohydrates allow for intercellular communication by binding to the carbohydrate-binding proteins (CBPs). Many physiological and even pathophysiological processes like pathogen-cell contact rely upon these interactions. Currently, more than 80 human CBPs have been identified [2,3], so one might immediately recognise CBPs as promising targets for anti-infective therapy. However, only a few of them have been thoroughly studied and as a result, few CBPs have been recognised and validated as drug targets. Furthermore, CBPs’ affinity is generally weak per monosaccharide unit, so CBPs usually form strong interactions by binding massive glycans that bear several hundred terminal carbohydrate units, many of which form contact with a single CBP or even a cluster of CBPs on the cellular surface. To inhibit glycan-protein interaction efficiently, one would have to consider designing and synthesizing multivalent carbohydrate ligands. This is by no means an easy task, and makes CBPs relatively unattractive targets in terms of druggability. Even if one would consider producing multivalent carbohydrates with optimal pharmacodynamic properties, they possess unattractive pharmacokinetic properties. Carbohydrates, especially large multivalent ones, share some intrinsic pharmacokinetic shortcomings as drugs; they are rapidly digested in the gut in most cases, and even if they survive the gut metabolism obstacles, they are practically unable to passively diffuse through the enterocyte layer in the small intestine. This inevitably means that no oral application can be guaranteed for a carbohydrate drug, which is the preferred application
route. Once in the body, carbohydrates are mostly a source of energy, but are also excreted rapidly with glomerular filtration. Due to these drawbacks, carbohydrates or carbohydrate-derived drugs were often considered unappealing even before their design would take place.

To tackle the first drawback, advances in CBP biochemistry has led to a progressive gain of knowledge on CBPs. A vast international research network named Consortium for Functional Glycomics (CFG, http://www.functionalglycomics.org/static/consortium/consortium.shtml) was founded in 2001 with the ultimate goal of enlightening the role of glycans and glycan-binding proteins. To date, hundreds of members of this network as well as outside researchers have revealed the secrets of CBPs and many are now considered potential drug targets: their specific ligands were discovered with screening on large glycoarrays and their binding sites were mostly elucidated [3]. Understanding the 3D structures of CBPs and their binding modes is indispensable for drug design. In order to offer a central repository of knowledge available about CBPs, PROCARB has been constructed [4]. PROCARB is a database of known and modelled carbohydrate-binding protein structures with sequence-based prediction tools, and is a single resource where all the relevant information about a pair of interacting proteins and carbohydrates is available. A similar, although a bit less appealing and informative database on lectin 3D structure is the database of lectins of Centre de Recherches sur les Macromolécules Végétales (CERMAV), a French fundamental research centre devoted to glycosciences (http://www.cermav.cnrs.fr/lectines/). Glycoscience.de (http://www.glycosciences.de/) is a German internet portal to support glycomics and glycobiology research. These web sites offer proof that the scientific community has recognised the potential of CBPs. Accordingly, many national and international associations have been involved in deciphering CBPs fundamental roles in human pathophysiological processes, which inevitably leads to on-going cognizance of how CBPs might be exploited to help develop novel therapeutic approaches for treatment of human diseases.

CBPs may be classified primarily into two broad categories: lectins (which are further subdivided into specific types, like C-, H-, I-, L-, P-, R-, and S-type lectins) and sulphated glycosaminoglycan-binding proteins [3]. Extracellular lectins are the most promising drug targets among CBPs and CFG divides them into 4 specific groups: C-type lectins, galectins (or S-type lectins), siglecs (a subclass of I-type lectins that bind sialic acid) and other. Out of 174 total CBPs records on the CFG webpage, 120 of them are classified as C-type lectins, which in itself highlights the importance of this class of CBPs. Like all extracellular lectins, C-type lectins bind terminal carbohydrate epitopes of glycans that originate from either pathogens or other cells. DC-SIGN, (Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin) is a type II C-type lectin that functions as an adhesion molecule located exclusively on dendritic cells (DCs). Originally discovered in 1992 [5], it was defined as a C-type lectin capable of binding the HIV-1 envelope glycoprotein gp120, but it took time until Steinman [6] and van Kooyk [7] unravelled its specific role in immunology and pathology and revealed the information to the broader scientific community. Since then, it has been shown to have a major role in primary immune response [8], but it also enhances several pathogens (like HIV, Ebola) infection of T cells and other cells of the immune system.
This makes DC-SIGN a very interesting CBP and a target of interest towards novel immunomodulatory and anti-infective therapy. DC-SIGN function can be modulated by small molecules termed DC-SIGN antagonists that bind to DC-SIGN and prevent binding of native DC-SIGN ligands. These molecules could act as novel anti-infectives, and are currently in the early phase of drug development. Latest in vitro studies demonstrate that DC-SIGN antagonists effectively block the transmission of pathogens like HIV-1 and Ebola to CD4+ T cells [10]. Although DC-SIGN has not been validated in vivo as a drug target yet, DC-SIGN antagonists are a fruitful example of how inhibition of a C-type lectin function might be achieved by small synthetic molecules. As exposed before, CBPs’ affinity is generally weak per carbohydrate unit, so the design of molecules that bind to DC-SIGN with high affinity is a demanding assignment. In general, the affinity issue and pharmacokinetic drawbacks of carbohydrates may be overcome by two predominant strategies implied in the design of DC-SIGN antagonists: screening of non-carbohydrate compounds to obtain ligands that are completely devoid of carbohydrate nature [11] and the design of glycomimetics, the compounds designed based on carbohydrate leads which usually still retain some or even a significant degree of carbohydrate nature [3,12]. DC-SIGN binds mannose- and fucose-based oligo- and polysaccharides, so their glycomimetics have been designed and proved to inhibit pathogen-DC-SIGN interaction potently. The author foresees that the approach used to design glycomimetic DC-SIGN antagonists is of general applicability when designing lectin antagonists and will be the thoroughly presented in the present chapter.

2.1. C-type lectins

Probably the largest type or family of lectins is the C-type found on animal (and human) cells [13]. This family includes several endocytic receptors and proteoglycans, and all collectins and selectins identified to date. They are of paramount importance for normal function of the immune system, as they mediate innate immunity, inflammation and immunity to tumour and virally infected cells. Although these CBPs vary greatly in structure among themselves, they have in common a domain, named C-type lectin-like domain (CTLD) or carbohydrate recognition domain (CRD). It is a compact globular structure responsible for selective binding of terminal units found in large carbohydrates [14]. The unique structural hallmark of such a domain is that it binds a carbohydrate by Ca²⁺ ion, which acts as a bridge between the protein and the “core monosaccharide” unit through complex interactions with sugar hydroxyl groups [15]. Namely, the binding with Ca²⁺ ion involves just one terminal saccharide unit – the “core monosaccharide” (as illustrated in Figure 1 for two distinct C-type lectins, the mannose-binding protein and DC-SIGN), while other ligand carbohydrate units (if present) form structural and bonding complementarity with the CRD. Several amino acid residues of the CRD offer 6 coordinate bonds for a Ca²⁺ ion and the carbohydrate donates 2 coordinate bonds with its hydroxyls, so that the Ca²⁺ ion is octacoordinated [16]. Distinct CRD aminoacid residues form hydrogen bonds with other hydroxyl groups on the carbohydrate directly or through water bridges. Changes in the amino acid residues that interact with the “core monosaccharide” modify the carbohydrate-binding specificity of the lectin, so that specific carbohydrate is recognised.
The free energy of such interactions is relatively weak per carbohydrate unit, as we have to take into account high desolvatation penalties of numerous hydroxyls upon carbohydrate binding. The C-type lectin family however has means to obtain high binding affinity, or better, avidity; some C-type lectins oligomerize in order to promote high avidity for specific glycans [19]. The clustering of the CRDs influences not only avidity, but also the lectin selectivity, since each individual CRD can act independently to bind end mono- or oligosaccharide moiety [20]. Although the majority of lectins contain a single C-type CRD, the macrophage mannose receptor has multiple independent CRDs in a single polypeptide. The adjacent CRDs in the mannose receptor may help to direct its binding to specific multivalent, mannose-containing glycans [21].

2.2. DC-SIGN function and structure

DC-SIGN (Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin) is a C-type lectin that functions as an adhesion molecule expressed specifically by dendritic cells (DCs), a class of professional antigen presenting cells (APCs). The intrinsic role of DCs is guidance of adaptive immune responses, since they are the major APCs that capture, process and present antigens [22]. DC-SIGN is a specific pattern recognition receptor (PRR) that recognizes distinct molecular patterns (PAMPs – Pathogen-Associated Molecular Patterns) of a number of pathogens [23]. It induces intracellular signalling pathways and triggers DC maturation upon pathogen binding [8,9]. To promote efficient transport of DCs towards effector cells – T cells, DC-SIGN binds human adhesion molecules ICAM-2 on vascular and lymphoid endothelium and enables cell interactions during DC migration [24]. Furthermore, DC-SiGN binding to ICAM-3 allows early antigen-nonspecific contact
between DC and T cells in the lymph nodes, enabling T cell receptor engagement by stabilizing the DC-T-cell interaction [25]. Thus, we may conclude that DC-SIGN enables some of the normal DC functions by binding endogenous ligands, but also modulates immune responses to diverse pathogens via its ability to induce antigen-specific intracellular signalling.

DC-SIGN is a transmembrane C-type lectin that consists of one CRD, which defines the ligand specificity of the receptor, a neck region composed of seven complete and one incomplete tandem repeats, and a transmembrane region followed by a cytoplasmic tail containing recycling, internalization and intracellular signalling motifs (Figure 2) [9].

![DC-SIGN schematic structure](image)

Figure 2. DC-SIGN schematic structure: Ca²⁺ CRD responsible for ligand binding/selectivity, a neck region composed of seven complete and one incomplete repeats, and a transmembrane region followed by a cytoplasmic tail (modified from Švajger et al.) [9].
Like many other C-type lectins, DC-SIGN tetramerizes to increase binding affinity and oligomerization occurs through association of the DC-SIGN neck domain (Figure 3) [26]. The oligomerization status of the DC-SIGN and related C-type lectins depends on the number of helical repeats of the neck region; at least 6 repeats are needed for tetramerization [27]. The stacked CRDs of tetramerized DC-SIGN provides the means of increasing the specificity for multiple repeating (oligo)saccharide units on host molecules, thereby defining the set of pathogens or endogenous molecules that are recognized by DC-SIGN. DC-SIGN neck domain, while allowing stacking and tetramerization, plays a central role as a pH-sensor that balances the equilibrium between the monomeric and tetrameric states of DC-SIGN [28]. In this sense, affinity for carbohydrates may be changed markedly by changing the pH, which helps DC-SIGN to realize its native function. Namely, upon binding, pathogen particles are internalized and further degraded into smaller particles and conjugated to MHC class-II proteins. As degradation proceeds in an acidic endosomal environment, the acid-triggered pathogen release from DC-SIGN is needed for successful degradation. Apart from tetramerization, DC-SIGN forms clusters that organize in membrane microdomains [29]. This organization on the plasma membrane is important for the binding and internalization, suggesting that clustered assemblies act as functional docking sites for pathogens (Figure 3).

Apart from allowing oligomerization, DC-SIGN neck repeat domains separate CRDs from the cell surface to enable multivalent interaction with glycans. DC-SIGN binds particularly well to viral and bacterial glycans with closely spaced terminal oligo- or monosaccharides with roughly 5 nm between the units [18,30,31]. DC-SIGN is not a totally rigid macromolecule and shows a degree of flexibility upon ligand binding [31]. Namely, DC-
SIGN adapts to an arrangement of terminal oligo- and monosaccharides, so all CDRs are allowed to interact with their ligands. Taken together, the tetrameric form of DC-SIGN and its conformational flexibility enable effective and selective binding of various glycans. Furthermore, due to the nature of the receptor-mediated cellular signalling, tetramerization of DC-SIGN could contribute to signal transduction after ligand binding.

2.3. DC-SIGN, a target for anti-infective therapy

Probably the most important feature of DC-SIGN is its ability to bind a great number of highly virulent pathogens [32]. Accordingly, it has been recognised as a potential new target for anti-infective therapy that perpetuates basic research of considerable importance. Apart from the already mentioned CFG, the importance of DC-SIGN as a new drug target is reflected in several other projects aimed to thoroughly clarify its function in pathogen infection. For example, several prominent European research teams have joined to form CARMUSYS, a collaborative training project aimed at designing and synthesizing carbohydrate multivalent systems to be used as inhibitors for pathogen attachment and penetration into target cells (CARMUSYS - Carbohydrate Multivalent System as Tools to Study Pathogen Interactions with DC-SIGN).

DC-SIGN modulates the outcome of the immune response of DCs by binding and recognizing a variety of microorganisms, including viruses (HIV-1, HCV, CMV, Dengue, Ebola, SARS-CoV, HSV, coronaviruses, H5N1, West Nile virus, measles virus), bacteria (H. pylori, M. tuberculosis, L. interrogans), fungi (C. albicans, A. fumigatus) and parasites (Leishmania, S. mansoni) [9]. On mucosal surfaces of the body where immature dendritic cells sample pathogens, DC-SIGN serves as one of the very first pathogen attachment points and the usual result of this interaction is pathogen internalization, degradation and exposure of the pathogen PAMPs to recruit CD4+ T-cells and to start the humoral immune response [6,22]. HIV-1 exploits native DC-SIGN functions to enslave DCs as carriers to boost T cell infection with or without becoming infected themselves [33]. The very first interaction between HIV-1 and DCs occurs via HIV-1 envelope glycoprotein gp120 with DC-SIGN on immature DCs (Figure 3). The HIV-1-DC-SIGN complex is rapidly endocytosed into early endosomes, where the acidic medium causes ligands to dissociate from DC-SIGN [34]. Upon dissociation, most DC-SIGN-bound ligands are lysed and processed as a normal degradation pathway. Some HIV-1 particles however remain bound to DC-SIGN and thus protected from the host immune system, so a fraction of HIV-1 retains its infectiveness [35]. HIV-1 may rest in DCs in an infectious state for days, hidden in undefined bodies that differ from both endosomes and lysosomes [36]. N-glycan composition of surface proteins governs the viral faith upon interaction of viral envelope with DC-SIGN [37]. By altering the N-linked glycan composition from mixed to oligomannose-enriched, one increases the affinity of HIV-1 for DC-SIGN, which enhances the viral degradation and reduces virus transfer to target cells. On the contrary, HIV-1 with oligomannose-enriched N-glycans is presented to viral envelope-specific CD4+ T cells more efficiently. In the alternative scenario, HIV-1 binds to DC-SIGN and facilitates lateral binding of HIV-1 to CD4 and CCR5 receptors expressed on the same immature DCs [33]. The direct consequence of this interaction is HIV-1 infection
of DCs [38]. To conclude, HIV-1 uses DC-SIGN as an entry mode to DCs, and DCs may be regarded as a Trojan horse that takes HIV-1 to CD4+ T-cells while protecting it from the host immune system [39]. A similar entry mode has been observed for some other pathogens like viruses (Ebola) and bacteria (M. tuberculosis, H. pylori) [32]. Apart from this straightforward infection pathway, HIV-1-infected DCs are able to mediate transmission of HIV-1 to CD4+ cells by means of immune response modulation and infectious synapse formation. Namely, HIV-1 causes both inhibition of DC maturation while inducing formation of viral synapse, a process previously attributed to mature DCs only [40]. As only several pathogens are able to modulate DCs maturation process, it may be concluded that the DC maturation depends upon selective ligand recognition, possibly also by DC-SIGN.

The fact that DC-SIGN acts as an entry point and a mediator of pathogen infections points out the possible therapeutic usefulness of DC-SIGN. DC-SIGN antagonists work by inhibiting pathogen interaction with DC-SIGN, so the very first phase of pathogen infection can be inhibited, as proven in in vitro experiments [10,41,42]. DC-SIGN antagonists may be designed as monovalent glycomimetics based on the DC-SIGN-binding oligosaccharides and their multimeric presentation [43]. Alternatively, screening of compound libraries has been successful in obtaining non-carbohydrate DC-SIGN antagonists [44]. As mentioned, potential carbohydrate-derived drugs have poor pharmacokinetic properties in general and this might raise some concern as to whether there is potential for therapeutically useful glycomimetic DC-SIGN antagonists. In the case of HIV-1, the major initial contact site between HIV-1 and DC-SIGN is in the vaginal mucosa, so a DC-SIGN antagonist could be administered topically to prevent HIV-1 transmission without systemic application [45]. It has to be stressed that no clinically proven therapy based on inhibition of DC-SIGN-mediated pathogen infection has been presented [43].

2.4. How does DC-SIGN choose among terminal monosaccharides?

DC-SIGN has a highly regulated recognition of its ligands as it selectively binds glycans with terminal D-mannose- and L-fucose expressed on a number of bacteria, parasites, fungi and viruses [46]. However, the mere presence of D-mannose and L-fucose does not assure binding selectivity itself. Namely, monosaccharide binding to DC-SIGN CRD is generally very weak, with $K_i($D-mannose$) = 13.1$ mM and $K_i($L-fucose$) = 6.7$ mM being the strongest binders among monosaccharides [26]. DC-SIGN CRD forms a 1-to-1 complex with terminal mono- or oligosaccharides, which relies upon already mentioned octacoordination of Ca$^{2+}$ ion in the binding site by the “core monosaccharide”; the equatorial 3- and 4-hydroxyls each form coordination bonds with the Ca$^{2+}$ in the binding site common to all C-type lectins, but also offer hydrogen bonds with amino acids that also serve as Ca$^{2+}$ ligands [47]. A crucial structural feature of a mannose residue is an axial position of the 2-hydroxyl group; this allows tight surface complementarity of the core mannose with the binding site. Equatorial position of the 2-hydroxyl group would probably prevent this tight binding due to steric clash, so hexopyranoses with equatorial 2-hydroxyl groups do not form strong interactions. To increase binding affinity, other saccharide units form additional interactions with the binding site, while binding specificity is based on spatial constraints. An excellent
description of selectivity/specificity mechanism may be found in the work of Guo et al., who demonstrated that DC-SIGN selectively binds high-mannose and fucosylated oligosaccharides [18]. The difference in the affinity of each oligosaccharide results from a different spatial arrangement of the mannose- and fucose-based ligands demonstrated in two crystal structures of fucose-based tetrascarharide LNFP III and Man₄ tetrascarharide bound to DC-SIGN CRD; the Man₃ mannose moiety of Man₄ (the “core monosaccharide”) inclines the rest of the molecule towards Phe313 with high surface complementarity, while the fucose moiety (the “core monosaccharide” of LNFP III) makes hydrophobic contact with Val351 and positions the second saccharide in a vertical manner, away from the protein surface (Figure 4). The binding mode presented for Man₄ tetrascarharide (PDB code: 1SL4) in which Phe313 forms a steric hindrance is a prevalent one for a mannose-containing oligosaccharides. However several other PDB structures of DC-SIGN CRD (PDB code: 1K9I, 1K9J, 2IT5 and 2IT6) show that Phe313 residue is rather flexible and allows 2 distinct binding modes, both including coordination of the Ca²⁺ by one mannose residue. One mode is shown in Figure 4, while in the second, the binding orientation of the mannose at the principal Ca²⁺ site is reversed, and creates different interactions between the terminal mannose and the region around Phe313 [47,48]. Thus we conclude that the binding mode of a specific oligosaccharide does not depend solely on the “core monosaccharide” involved, but is highly sensitive to the substitution pattern and 3D structure of adjacent monosaccharides. Furthermore, the mannose- and fucose-based ligands have overlapping, but different binding modes, which might offer a rational explanation for the different biological effects of mannose- and fucose-based ligands [49]. Consequently, the two observed binding modes offer a solid basis upon which DC-SIGN antagonists with either fucose or mannose anchors can be designed [50].

On the supramolecular level, DC-SIGN tetramerization has a major impact on binding affinity. By forming tetramers, binding affinity for glycans with repetitive sugar motifs with high-mannose or fucose N-linked oligosaccharides increases markedly [46]. This simple observation drives us to the conclusion that only polyvalent ligands could efficiently bind to DC-SIGN and offer a rationale for design of multivalent carbohydrate systems as DC-SIGN antagonists [51].

DC-SIGN was first regarded as the major binding lectin for various mannose-glycosylated PAMPS including HIV-1 gp120. However, other C-type lectins bind both D-mannose and L-fucose-containing glycans, which brings under question DC-SIGN’s exclusive role in pathogen binding, but above all, raises the concern of binding selectivity when designing DC-SIGN antagonists. In particular, two C-type lectins - Langerin on epithelial Langerhans cells and a mannose receptor on dermal DCs - bind high-mannose oligosaccharides and HIV-1 gp120 with high affinity [52]. Their function in HIV-1 (and possibly other pathogen) infection differs: interaction of HIV-1 with DC-SIGN enables HIV-1 to survive a host immune system, while Langerin helps Langerhans cells to eradicate HIV-1 [53,54]. On the other hand, mannose receptor facilitates HIV-1 infection of DCs by the CD4/CCR5 entry pathway [53]. Therefore, blockade of both DC-SIGN and mannose receptors seems the right strategy for prevention of HIV-1 entry into DCs while Langerin inhibition suppresses viral
Figure 4. Selected amino acid residues (thin sticks) of DC-SIGN CRD (transparent surface) in complex with: a) Man$_3$ tetrasaccharide (solid sticks, PDB code: 1SL4), and b) fucose-based tetrasaccharide LNFP III (solid sticks, PDB code: 1SL5) [18]. The Man$_3$ mannose moiety of Man$_4$ or the “core monosaccharide” inclines the rest of the molecule towards Phe313 with high surface complementarity, while the fucose moiety or the “core monosaccharide” of LNFP III makes hydrophobic contact with Val351 and positions the second saccharide in a vertical manner, away from the protein surface, so only one galactose moiety makes additional contacts with DC-SIGN CRD.
clearance thus allowing a boost of HIV-1 infection. When designing an efficient HIV-1 entry inhibitor based on C-type lectin antagonism, an agent of choice should bind to both DC-SIGN and mannose receptors, while having no or marginal affinity to Langerin. Although all three receptors bind to virtually the same ligands, selectivity against Langerin is an achievable objective since Langerin binding sites differ from that of DC-SIGN [55]. As expected, selectivity versus Langerin does not necessarily rely on the “core monosaccharide” involved in the binding process, but on spatial constraints formed by adjacent glycan monosaccharide units [56].

2.5. Binding of monovalent DC-SIGN antagonists

As mentioned earlier, binding of DC-SIGN natural ligands depends upon the presence of an L-fucose or D-mannose hexopyranose unit. Their inherent disadvantages in terms of low activity and/or insufficient drug-like properties can be modified by the design of glycomimetics – compounds that mimic the bioactive function of carbohydrates but have far better drug-like properties [3]. This concept has been successfully used in the design of monovalent DC-SIGN antagonists. The term “monovalent” is used here to describe low-molecular weight molecules that can occupy only one DC-SIGN CRD at a time, so it incorporates mono- and oligosaccharide structures and their mimetics. The design of these monovalent glycomimetics can be structurally divided into the following sections, as depicted in figure 5:

- the choice of a monosaccharide unit,
- the choice of glycosidic bond surrogate,
- the choice of adjacent saccharides or structures that contribute to overall binding affinity.

![Figure 5](image.png)

**Figure 5.** Schematic presentation of the design of monovalent glycomimetics that act as DC-SIGN antagonists: systematic replacements in structure that lead to efficient DC-SIGN ligands.

2.5.1. The choice of a monosaccharide unit

The most extensive work regarding the choice of the monosaccharide unit comes from the work of Bernardi and Rojo’s groups from CARMUSYS [41,50,57]. In particular, they have
shown that L-fucose can be incorporated in glycomimetic surrogates of Lewis-X trisaccharide to obtain even better affinity than the native trisaccharide. The full Lewis-X mimetic 1 was shown to inhibit DC-SIGN binding of mannosylated BSA in the upper micromolar range (IC₅₀=350 μM), but the second generation of analogous compounds failed to give any significant improvement over 1 (Figure 6) [50,58]. STD-NMR studies of 1 with DC-SIGN ECD have shown that only fucose residue makes strong contact with the DC-SIGN CRD. A reasonable explanation for this observation might be the before mentioned binding mode of the L-fucose moiety: it positions the second saccharide in a vertical manner away from the protein surface, and thus the rest of the molecule fails to form tight interactions with protein. Just a glance at figure 4 reveals that DC-SIGN CRD is quite flat and high structural complementarity is one of the requirements for strong binding. In the absence of functional groups that would allow ionic interactions which are not highly dependent on the distance, the rest of the Lewis-X mimetic 1 and its analogues probably form the majority of interactions with the solvent. Lewis-X mimetics however share one important figure; their affinity for Langerin is insignificant, so they are selective DC-SIGN antagonists [58].

The L-fucose monosaccharide has the highest affinity for DC-SIGN among monosaccharides, and L-fucose should be the logical choice as the “core monosaccharide” when designing DC-SIGN antagonists. On the contrary, D-mannose has received most of the attention as the majority of mono- and polyvalent DC-SIGN antagonists incorporate D-mannose as the “core monosaccharide”.

Figure 6. The choice of “core monosaccharide” of DC-SIGN antagonists; L-fucose (L-Fuc), D-mannose (D-Man), 2-substituted D-mannose and reduced shikimic acid are useful “core monosaccharides”.
Pseudo-1,2-mannobioside 2 and its analogues of Reina et al. (Figure 6) contain a D-mannose unit substituted at the anomeric position with conformationally constrained cyclohexanediol [41]. The STD-NMR of DC-SIGN ECD with an azido derivative of 2 shows that the compound makes close contact with the protein, which is in agreement with the binding mode of Man₄ tetrasaccharide. The inhibitory activity of 2 on Ebola virus entry into DC-SIGN expressing Jurkat cells was quite high (IC₅₀=0.62 mM) and this was the first functional assay showing that DC-SIGN antagonism with small molecules might be used to inhibit viral transfection mediated by DC-SIGN. Another example of substituted D-mannose as the “core monosaccharide” might be found in the work of Mitchell et al.: they have synthesized a small library of 2-C-substituted branched D-mannose analogues of which compound 3 exhibited a 48-fold stronger binding to DC-SIGN (Ki=0.35 mM, Ki(mannose)=17.1 mM) [59].

![Figure 7. Design of reduced shikimic acid “core monosaccharide” as D-mannose mimetic [60].](image)

An innovative approach was used by Garber et al.: taking D-mannose as the lead structure, they have concentrated on the spatial relationship of hydroxyls at positions 2,3 and 4 and concluded that reduced shikimic acid should enable the same spatial relationship of hydroxyls (Figure 7) [60]. They have synthesized 192 derivatives of reduced shikimic acid and compound 4 (Figure 6) was found to be the most potent hit of this focused library.

2.5.2. The choice of glycosidic bond surrogate

The metabolic instability of glycosidic bonds renders it inappropriate for the design of drugs, and hence requires an appropriate surrogate when designing mimetics of oligosaccharides. Numerous changes in carbohydrate structures have been successful in the design of glycosidase inhibitors, and exactly these structures could be used to modify glycosidic bond instability [61]. However, unwanted inhibition of glycosidases may as well provoke side effects of potential drugs, so careful choice of glycosidic bond surrogates has to be made if only its stability is the ultimate goal. The α-glycosidic bond found in both L-fucose and D-mannose containing oligosaccharides that bind to DC-SIGN was successfully replaced by Timpano et al. by a stable α-fucosylamide structure (compound 1, Figure 6) and shown not to affect the binding affinity in a detrimental sense [50]. The derivatives of reduced shikimic acid (compound 4, Figure 6) have 2 features that influence glycosidic bond stability: first, they have a thioglycosidic bond, which is proven to be metabolically more stable towards glycosidases [62], and second, they are a constitute of reduced shikimic acid which is a carbasugar. Carbasugars lack anomeric reactivity, which implies their metabolic stability towards glycosidases and glycosyltransferases. Although glycosidic bond
surrogates were often not used in the design of DC-SIGN antagonists, the stability of glycosidic bonds has to be challenged when designing stable glycomimetics. According to the latest literature, a large number of alternatives already exist [63].

2.5.3. The choice of adjacent saccharides or structures that contribute to overall binding affinity

Monosaccharide moieties other than “core monosaccharide” in the structure of DC-SIGN oligosaccharide ligands form not only additional interactions with the binding site, but also influence the binding specificity with spatial constraints and point other monosaccharide units towards or away from the protein surface, as seen in the Figure 4. Notable quality of the adjacent monosaccharide units is that they do not form the same interactions like the “core monosaccharide”, but instead form a network of H-bonds, possibly through water molecules. For example, Man₄ tetrasaccharide makes contact with DC-SIGN through at least two water molecules, while one stabilizes its binding conformation [18,64]. Alternatively, both “core” and adjacent monosaccharides make surface complementarity and hydrophobic interactions. In particular, Van Liempt et al. demonstrated that Val351 in DC-SIGN creates a hydrophobic pocket that strongly interacts with the Fucα1,3/4-GlcNAc moiety of the Lewis antigens [65]. So, although highly polar in nature, the adjacent monosaccharide units contribute significantly to overall binding free energy also by hydrophobic interactions apart from being mere “linker” to other structures. This implies that altering hydrophilic character of adjacent monosaccharide units to more hydrophobic surrogates should increase free binding energy by increasing hydrophobic interactions. With this in mind, Timpano et al. have used the (1S,2R)-2-aminocyclohexanecarboxylic acid as a scaffold/linker to attach D-galactose mimetic into the structure of compound 1 and its derivatives (Figure 8) [50]. The molecule was carefully chosen to mimic Lewis-X structure while minimizing the hydrophilic ballast of the Lewis-X, and galactose mimetic was incorporated based on the observation, that galactose residue makes contact with the DC-SIGN CRD surface and is thus important in binding [18].

Figure 8. (1S,2R)-2-aminocyclohexanecarboxylic acid as a scaffold/linker to attach D-galactose mimetic into the structure of compound 1 and its derivatives, all mimic Lewis-X [50].
In the series of mannose-based DC-SIGN antagonists, Reina et al. and Sattin et al. (groups of Rojo and Bernardi) have incorporated (1S,2S,4S,5S)-dimethyl 4,5-dihydroxycyclohexane-1,2-dicarboxylate as adjacent monosaccharide mimicking “trans” conformation of 1,2-hydroxyls in D-mannose (Figure 9, compounds 2 and 5) [10,41]. Again, the rational for this change was to imitate the 3D relationship of key hydroxyls in the D-mannose moiety, while lowering the overall hydrophilicity. Furthermore, the cyclohexane saccharide mimic lacks glycosidic bond and is thus metabolically stable.

Starting from compound 2 or its azido derivative, the group of Bernardi (Obermajer et al.) continued to pursue the idea of increasing the binding affinity by identifying two binding areas around Phe313 in the DC-SIGN binding site that were only partially occupied by co-crystallized tetramannoside Man4 [64]. These hydrophobic areas were targeted by attaching different hydrophobic moieties to deprotected carboxylates of pseudo-1,2-mannobioside 2 (Figure 10). A number of mannose-based DC-SIGN antagonists were synthesized (an illustrative example is compound 6), and the majority of them inhibited DC adhesion at low micromolar concentrations improving the potency of the starting compound 2 by two orders of magnitude. Probably the same hydrophobic areas have contributed to the affinity of 4.

Figure 9. (1S,2S,4S,5S)-dimethyl 4,5-dihydroxycyclohexane-1,2-dicarboxylate as central monosaccharide surrogate that mimics “trans” conformation of 1,2-hydroxyls in D-mannose.

Figure 10. Increasing potency of DC-SIGN antagonists by attaching hydrophobic moieties to deprotected carboxylates of 2 to afford 6 and its derivatives [64].
2.6. Increasing affinity and/or avidity?

The evident disadvantage of monovalent DC-SIGN antagonists – their low affinity for DC-SIGN and thus low potency – can be overcome by conjugating monovalent units to various scaffolds for polyvalent presentation (Figure 11) [51]. Namely, even the most potent monovalent ligands have inhibitory constants in low micromolar concentrations, while therapeutically useful compounds have the same effect in nanomolar or even picomolar concentrations. The polyvalent antagonists are believed to interact with possibly all DC-SIGN CRDs on tetramers, or perhaps, they might interact and collate DC-SIGN clusters on the cell surface in the same manner as HIV-1 increases its avidity to DC surface (depicted in Figure 3). The avidity observed for polyvalent ligands in general originates from multiple binding: the polyvalent molecule with reversible mechanism of binding has higher possibility for being bound to at least one of receptor binding sites (i.e. one of CRDs), so that dissociation rate constant significantly decreases [66]. The other rationale may be derived from the observation of Andrews et al., who concluded that the average loss of overall rotational and translational entropy accompanying drug-receptor interaction is a constant for relatively small molecules and was estimated to be approx. 14 kcal/mol at 310 K [67]. In other words, more favourable free binding energy is obtained by combining binding epitopes/structures into one larger molecule, and this is also true for uniting same monovalent structures into a large polyvalent one because all rotational and translational entropic losses that occur during binding of individual monovalent molecules are reduced to one entropic loss for a larger molecule. According to both theories, when larger polyvalent molecules are employed, binding avidity is observed instead of just a linear increase in binding affinity. The main concern when choosing appropriate polyvalent support is the spatial relationship between individual monovalent ligands: monovalent ligands should be appropriately spaced to allow binding to at least two binding sites (or CRDs in case of C-type lectins), otherwise avidity cannot be achieved.

**Figure 11.** Strategy for increasing binding affinity/avidity of DC-SIGN antagonists: monosaccharides or monovalent glycomimetics are attached to polyvalent dendrimer or dendron core (modified from Anderluh et al.) [43].
The first polyvalent DC-SIGN antagonists were designed, synthesized and assayed in the laboratories of Rojo and Delgado [68]. A simple glycodendritic core bearing 32 mannose residues conjugated to BoltornH30 dendrimer through a succinic acid spacer (7, Figure 12) hindered DC-SIGN mediated Ebola virus infection at nanomolar concentrations (IC₅₀=337 nM). In comparison, the same inhibition was achieved with only millimolar concentrations of α-methyl-D-mannopyranoside (IC₅₀=1.27 mM); the core bears only 32 mannose residues and the difference in binding affinity is more than 3000-fold, which clearly indicates high binding avidity.

Figure 12. Glycodendritic structure bearing 32 mannose residues conjugated to BoltornH30 dendrimer as DC-SIGN antagonist [68].

Relatively simple glycopolymers with a different load of combined α-mannose and β-galactose in ratios were synthesized by Becer et al. and evaluated in inhibition of DC-SIGN-gp120 binding (8, Figure 13) [69]. Glycopolymer with 25% mannose failed to inhibit gp120 binding potently (IC₅₀ of 1.45 μM), while glycopolymer with 100% mannose was far more efficient (IC₅₀ of 37 nM) even when calculated per mannose unit, as it had approx. 40 times higher affinity compared with 4-times higher load of the mannose unit.

Polyvalent DC-SIGN antagonists described before relied on the use of “core monosaccharide” only. The design and choice of a potent monovalent DC-SIGN ligand might reduce the requirement for huge dendrimeric presentation while retaining desired effect. Designed monovalent glycomimetic 4 (Figure 6) of Garber et al. was loaded onto a carefully selected multivalent core designed to link at least 2 CRDs of DC-SIGN tetramer as monovalent units were roughly 40Å apart, which is exactly the width of DC-SIGN CRD (8, Figure 13) [60]. Glycopolymer 8 with 29 units of 4 was found to have 1000-fold higher affinity for DC-SIGN (IC₅₀=2.9 μM) than 4, showing that high avidity binding was obtained.
It is hard though to assess the relevance of designed monovalent ligand versus D-mannose, as the analogous polyvalent DC-SIGN antagonist bearing D-mannose or reduced shikimic acid was not synthesized. The direct evidence however came from the work of Sattin et al [10]. They have attached four copies of monovalent trimannoside mimetic 5 (Figure 9) to a tetravalent dendron 10 (Figure 14), which potently inhibited HIV-1 transfection to CD4+ T lymphocytes (>94 % inhibition at 100 μM). For comparison only, a tetrameric dendron bearing 4 copies of simple D-mannose (11, Figure 14) failed to inhibit HIV-1 transfection with even comparable potency (65 % inhibition at 100 μM). Furthermore, Dendron 10 inhibited Ebola cis infection of Jurkat T cells one order of magnitude more potently than dendron 11 [70]. This data clearly demonstrates that potent monovalent DC-SIGN antagonists reduce the need for high polyvalency number and influences binding affinity markedly.

Figure 13. Glycopolymers of Becer et al. and Garber et al. with high binding avidity [60,69].

Figure 14. Tetravalent dendrons of Sattin et al. bearing four copies of monovalent trimannoside mimetic 5 (10) or D-mannose residue (11) [10].
3. Conclusion; could it work for all C-type lectins?

As mentioned earlier, all C-type lectins share a structural feature, namely C-type lectin-like domain (CTLD) or carbohydrate recognition domain (CRD) responsible for Ca\(^{2+}\)-dependent selective binding of terminal mono- or oligosaccharide units of large carbohydrates. This implies that only the “core monosaccharide” makes contact with Ca\(^{2+}\) ion while other ligand carbohydrate units (if present) form structural and bonding complementarity with the CRD. From this point of view, the systematic approach presented herein could be of general applicability when designing glycomimetic C-type lectin antagonists. It consists of designing the monovalent ligand based on three distinct steps: the choice/design of a “core” monosaccharide unit, the choice/design of glycosidic bond surrogate and the choice/design of adjacent saccharides or structures that contribute to overall binding affinity. Still, the evident drawback of monovalent glycomimetics is their low affinity not only to DC-SIGN, but to lectins in general [3]. The other characteristic of C-type lectins is their ability to oligomerize, and further make clusters of functional oligomers. The tactics of polyvalent presentation targets exactly the oligomerized or even clustered structures and is a prerequisite of polyvalent structures as C-type lectin antagonists. Accordingly, high avidity for DC-SIGN and other C-type lectins can be achieved with high loading of monovalent ligands to various polyvalent systems. Taken together, the results of several groups presented in this chapter clearly demonstrate that this general procedure for designing glycomimetic DC-SIGN antagonists gives notable results, and according to the structural resemblance of diverse C-type lectin CRDs, there is a high probability of its applicability to distinct C-type lectins.

Author details

Marko Anderluh

Department of Medicinal Chemistry, University of Ljubljana,
Faculty of Pharmacy, Ljubljana, Slovenia

Acknowledgement

This work is dedicated to Prof. Dr. Slavko Pečar, a great mentor and teacher, who taught me not only pure science, but above all how to nourish my enthusiasm for science.

4. References


