Identification of Multiple Druggable Secondary Sites by Fragment Screening against DC-SIGN

Jonas Aretz, Hannes Baukmann, Elena Shanina, Jonas Hanske, Robert Wawrzinek, Viktor A. Zapol’skii, Peter H. Seeberger, Dieter E. Kaufmann, Christoph Rademacher

Multiple binding sites: DC-SIGN, known for 15 years for its role in HIV trans-infection of T cells, is one of the most attractive targets among glycan-binding proteins. Still, drug-like effectors are sparse but its undruggable primary site might be bypassed by targeting druggable secondary sites.
Identification of Multiple Druggable Secondary Sites by Fragment Screening against DC-SIGN
Jonas Aretz[a,b], Hannes Baukmann[a,b], Elena Shanina[a,b], Jonas Hanske[a,b], Robert Wawrzinek[a], Viktor A. Zapol’skii[c], Peter H. Seeberger[a,b], Dieter E. Kaufmann[c], Christoph Rademacher[a,b]

Abstract: DC-SIGN is a cell surface receptor for several pathogenic threats such as HIV, Ebola virus or Mycobacterium tuberculosis. Multiple attempts to develop inhibitors of the underlying carbohydrate–protein interactions have been undertaken in the past fifteen years. Still, drug-like DC-SIGN ligands are sparse, which is most likely owed to its hydrophilic, solvent-exposed carbohydrate binding site. Here, we report on a parallel fragment screening against DC-SIGN applying SPR and a reporter displacement assay, which complements previous screenings using 19F NMR and chemical fragment microarrays. Hit validation including SPR and 1H-15N HSQC NMR revealed that although no fragment bound in the primary carbohydrate site, five secondary sites are available to harbour drug-like molecules. Building on key interactions of the reported fragment hits, these pockets will be targeted in future approaches to accelerate the development of DC-SIGN inhibitors.

Many members of the C-type lectin (CLR) family are expressed as transmembrane receptors on cells of the innate immune system and serve as pattern recognition receptors regulating the immune response. Recognition of surface carbohydrates by these receptors is mediated via a central Ca2+ co-factor and promotes immune cell signaling as well as pathogen uptake resulting in antigen processing and presentation on MHC molecules. DC-SIGN is a prominent member of the CLR family that recognizes mannose- and fucose-type ligands on viruses such as HIV, Ebola virus, Dengue virus, Phlebovirus as well as bacteria such as Mycobacterium tuberculosis[1]. In particular, DC-SIGN-expressing cells are amongst the first immune cells to encounter HIV and promote transinfection of T cells[2]. These findings have generated significant interest in the development of inhibitors of DC-SIGN.

The extracellular domain of DC-SIGN forms homotetramers which each consists of a neck and a carbohydrate recognition domain (CRD). This increases the avidity of the low-affinity monovalent carbohydrate recognition (Kd (mannose) = 3.5 mM[2]) into the nanomolar range for multivalent interactions (Kd (gp120) = 1-2 nM[3]). Utilizing the multivalent display of simple carbohydrates on various supports lowered the inhibitory constants of several reported inhibitors into the nano- to picomolar affinity range[4]. On the other hand, low-molecular weight carbohydrate-derived inhibitors show only modest affinity in the medium to high micromolar regime[5]. Moreover, non-carbohydrate, drug-like inhibitors for DC-SIGN are limited to quinoxalinones, which display high potency in cell-based assays but are electrophiles with partially limited stability[6].

In general, the development of inhibitors for mammalian glycan-binding proteins is challenging owing to their shallow and hydrophilic binding sites. Starting from the primary carbohydrate-binding site elaborating on the original carbohydrate as a scaffold has proven to be a promising route for some targets[7]. However, such approaches capitalize on the presence of extended binding sites that allow for affinity maturation of the carbohydrate scaffold by the attachment of hydrophobic substituents. Often such secondary sites are neither well described nor directly accessible from the analysis of crystallographic protein structures as they might stem from minor alterations of the protein geometry. Previous reports indicated the presence of binding sites for drug-like molecules to DC-SIGN[8,9]. As these drug-like inhibitors lack functional groups to directly interact with the primary Ca2+ ion, an allosteric mechanism was proposed[9], but the nature of their pockets remained elusive. Here, we applied fragment-based screening to identify low molecular weight ligands and their respective binding sites for DC-SIGN.

A library of 986 fragments was screened in parallel by surface plasmon resonance (SPR) and a reporter displacement assay (RDA). Along with the results from our previously reported screening of a sublibrary of 281 fluorinated fragments using 19F NMR and chemical fragment microarrays, a large set of orthogonal screening data was available for DC-SIGN[8,9]. A detailed analysis of the entire parallel screening dataset is available in the Supplementary Information (Figures 1, S2, S3). Most importantly, SPR and 19F NMR screening resulted in 49 (5.0% hit rate) and 46 hits (16.3%), respectively (Figure 1 a,b). These 95 compounds were counter-screened by SPR and RDA (Figure 1 c,d), resulting in 61 (64%) hits validated by SPR and 4 (4%) hits validated by SPR and RDA. The best affinity of 0.3 nM was estimated for compound 6, the ligand with the highest efficiency that has been reported to date for DC-SIGN (LE = 0.34 kcal HA−1 mol−1). Moreover, these results suggest a moderate druggability of DC-SIGN[11], which is in accordance to our previous results[8]. It is noteworthy that only four compounds showed activity in the RDA with IC50 values above 1 mM. This assay directly probes for inhibitors rather than binders and consequently implies a low druggability of the carbohydrate binding site[11]. Taken together, the large discrepancy between binders and competitors indicates potential druggable secondary sites.

[a] J. Aretz, H. Baukmann, E. Shanina, Dr. J. Hanske, Dr. R. Wawrzinek, Dr. P. H. Seeberger, Dr. C. Rademacher
Department of Biomolecular Systems
Max Planck Institute of Colloids and Interfaces
Am Mühlenberg 1, 14476 Potsdam (Germany)
E-mail: christoph.rademacher@mpikg.mpg.de
[b] J. Aretz, H. Baukmann, Dr. J. Hanske, Dr. P. H. Seeberger, Dr. C. Rademacher
Department of Biology, Chemistry, and Pharmacy
Freie Universität Berlin
Takustraße 3, 14195 Berlin (Germany)
[c] Dr. V. A. Zapol’skii, Dr. D. E. Kaufmann
Institute of Organic Chemistry
Clausthal University of Technology
Lebnizstraße 6, 38678 Clausthal-Zellerfeld (Germany)

Supporting information can be found under https://doi.org/10.1002/anie.201701943
Figure 1. Examples from fragment screening and counter-screening against DC-SIGN. a) Scatter plot of relative response versus compound ID from SPR screening. A response (RU) larger than 3.0 or a calculated LE > 0.5 were chosen for hit identification (black) whereas compounds exceeding the calculated RU<sub>max</sub>, showing negative response, or slow or irreversible kinetics were excluded as potential false positives. b) Example from the 19F NMR-based fragment screening described previously. The expansion from a T<sub>2</sub>-filtered 19F NMR spectrum (T = 1 s, ν<sub>CPMG</sub> = 50 Hz) shows three compounds that were increased in signal intensity after mannan addition. c) SPR sensorgram and one-site-binding model fit of 8 from the counter-screen. The apparent dissociation constant was estimated to be 500 µM (n=3). d) Second counter-screening using a FITC-dextran-based RDA. Compounds were measured in triplicates at 1 mM concentration and compared to a DMSO control.

To expand our insight into the presence of secondary sites, we applied 1H-15N HSQC NMR analysis using 15N isotope-labeled DC-SIGN CRD (Figure 2 a) and computational pocket predictions. For the NMR analysis, we first tested the DC-SIGN ligand mannanose, which is a millimolar binder (Figure S6, S7). Here, mannanose addition to DC-SIGN CRD induced chemical shift perturbations (CSPs) and reduced resonance intensities of residues close to the primary site as already observed earlier for the carbohydrate Lewis X (Figure 2 b,c). Then, we analyzed 28 fragments to identify their potential binding sites (Figure 2). These fragments were chosen based on the aforementioned RDA and SPR results. In total, 14 of 19 (74%) validated fragments from the SPR screening and all ten (100%) from the 19F NMR screening induced changes in the HSQC spectra. As expected from compounds passing a cascade of orthogonal techniques, the validation rates obtained by 1H-15N HSQC NMR increased compared to the first and second counter-screen. Most importantly, all fragments induced CSPs or altered peak intensities of subsites compared to mannanose (Figure 2 b,c). In addition, these changes were distinct between individual fragments (Figure 2 b, S8-15), indicating the presence of several secondary binding sites.

In parallel, three computational pocket prediction algorithms were applied to the structure of DC-SIGN CRD. In addition to our previously reported results employing DoGSiteScore, hot spots for binding of drug-like molecules were identified by computational solvent mapping using FTMap and potentially inducible sites were predicted by CryptoSite. We determined five pockets with FTMap which locate in front and in the back of the primary carbohydrate binding site (site I and II, respectively), two pockets close to the neck domain (site III and IV), and one cluster of hot spots between the α<sub>1</sub> and α<sub>2</sub> helix opposite the primary site (site V; Table S3, Figure 3, S16). Among these five binding sites, site II to V also contain potentially inducible residues according to CryptoSite. Site I, III, IV, and V were also identified by DoGSiteScore (Table S4, S5). Of note, the primary carbohydrate binding site was not identified by any of these methods, which is in agreement with our 1H-15N HSQC NMR fingerprinting. Furthermore, site I was recently targeted using diaryl-substituted carbohydrate-analogs. In summary, DC-SIGN supposedly harbors five potential secondary sites that can accommodate drug-like molecules. Four of these pockets might also be inducible cryptic sites which only open upon ligand binding and are invisible in the static apo crystal structure.
Finally, the predicted binding sites were compared to our experimental HSQC NMR results. From 22 actives, six (27%) were assigned to site I, four (18%) to site II, two (9%) to site III, three (14%) to site IV, and nine (41%) to site V (Table S6, Figures S8-S15). The highest affinities estimated by SPR for three (14%) to site IV, and nine (41%) to site V (Table S6). These pockets have the potential to guide the development of DC-SIGN effectors. We mapped six substructures (3 to 7 and 9) of drug-like DC-SIGN inhibitors in site I and II, which are adjacent to the primary carbohydrate site (Figure S17). Hence, our data suggest that these reported effectors, which have in vitro and in vivo activity[6, 19], might also interact with these pockets (Figure S17). Future expansion of fragments 1 to 10 capitalizing on their key interactions could improve the efficiency of drug-like inhibitors[6]. In analogy, ligands binding to site I could be extended towards the primary site to induce a steric clash with the natural ligand. Alternatively, fragments interacting with sites I and II could be linked to glycomimetics combining their key interactions with those of the primary site. Such strategy was successfully followed for inhibitors of MAG[2] and E-Selectin[7]. Hence, the series of substituted D-mannoside inhibitors of DC-SIGN targeting site I could be improved by substituting the diaryl with fragments 1 to 6[8].

Interestingly, mannose as well as most fragments also induced chemical shifts or changes in dynamics of residues that were distant from their assigned primary pocket. Exemplarily, G264, which is part of the β2 sheet close to the neck domain (Figure S16), changed in chemical shift upon binding of 2, 3, 7, 8, and 9 in sites I and II. These long distant perturbations might be caused by an allosteric network, which we recently identified to regulate Ca2+ binding and release in the related CLR Langerin[9]. Activation of such an allosteric network might propagate via the neck domain, which could explain signaling initiated by monovalent ligands[10]. Larger structural rearrangements of DC-SIGN following carbohydrate recognition have been recently inferred from surface force measurements supporting our hypothesis[11]. As the initiation of DC-SIGN signaling has potential role in the treatment of fibrosis[10] and induce transplantation tolerance[12], DC-SIGN agonists may have important clinical implications.

To generalize the effect of multiple secondary sites on the outcome of a fragment screen, we investigated by which screening method hits validated by 1H-15N HSQC NMR were identified (Table S6, Figure S18). Among the nine compounds originating in the 19F NMR screening, five (56%) were mapped to binding site I whereas only one hit originating in the SPR screening bound to this site (7%). On the other hand, eight hits

Table 1. Summary of validation results and binding site mapping of chosen hits[a].

<table>
<thead>
<tr>
<th>ID</th>
<th>Structure</th>
<th>Screen[b]</th>
<th>SPR EDTA[c]</th>
<th>SPR Ca2+[d]</th>
<th>RDA[e]</th>
<th>Site[h]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NMR, array</td>
<td>0.8 ± 0.1</td>
<td>2.6 ± 0.1</td>
<td>Yes</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>NMR, array</td>
<td>4.8 ± 0.9</td>
<td>3.6 ± 0.3</td>
<td>Yes</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>NMR</td>
<td>0.3 ± 0.2</td>
<td>0.5 ± 0.1</td>
<td>Yes</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>NMR</td>
<td>1.1 ± 0.1</td>
<td>2.2 ± 0.2</td>
<td>No</td>
<td>II</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>SPR</td>
<td>0.4 ± 0.1</td>
<td>No</td>
<td>No II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>NMR, SPR, array</td>
<td>1.0 ± 0.1</td>
<td>3.5 ± 0.1</td>
<td>No III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>SPR</td>
<td>0.7 ± 0.1</td>
<td>0.7 ± 0.2</td>
<td>No IV</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(57%) from SPR screening were mapped to binding site V which was the case for only two fragments (22%) from the 19F NMR screening. For microarray hits no preferred binding site was observed (29% hits in site I and V). As discussed in detail in the Supplementary Information, these differences are most likely caused by the immobilization of DC-SIGN to the SPR chip and by addition of CaCl2 and mannan to identify hits in 19F NMR screening. Thus, the choice of the screening method can impact the chemical matter of hits[10], the hit rate[13], and the binding site. Taken together, fragments from a parallel screening against DC-SIGN were validated by orthogonal techniques and analyzed by 1H-15N HSQC NMR and computational pocket prediction methods. Finally, five potential secondary binding sites were identified, which are able to accommodate drug-like molecules. As three of these sites are moderately druggable they represent suitable targets for future drug discovery campaigns on DC-SIGN.

Experimental Section

Experimental Details are given in the Supplementary Information.

Acknowledgements

The authors thank the Max Planck Society and the German Research Foundation (DFG, RA1944/2-1) and the Boehringer Ingelheim Stiftung for financial support. Olaf Niemeyer is acknowledged for NMR technical support.

Keywords: Drug discovery • NMR spectroscopy • Carbohydrates • Immunology


[14] see [15]


