Electronic Supplementary Material (ESI) for RSC Chemical Biology.
This journal is © The Royal Society of Chemistry 2021

Molecular recognition of sialoglycans by streptococcal Siglec-like adhesins: to-ward the shape of specific inhibitors

Cristina di Carluccio[a], Rosa Ester Forgione[a], Andrea Bosso[b], Shinji Yokoyama[c], Yoshiyuki Manabe[c], Elio Pizzo[b], Antonio Molinaro[a,c], Koichi Fukase[c], Marco Fragai[d], Barbara A. Bensing[e], Roberta Marchetti[a*], Alba Silipo[a*].

E-mail: marchetti@unina.it; silipo@unina.it

Table of Contents

Experimental Procedures

Preparation of sialoglycans. The 3'-sialylactosamine (3'-SLn) was purchased from Tokyo Chemical Industry Co., Ltd. The sialyl-T-antigen linked to the threonine was chemically synthesized.[1] The sugar portion of the ganglioside GM1b was purchased from Elicictyl-oligotech.

Protein expression and purification. The Siglec and Unique domains of Siglec-like adhesins SLBR-H and SLBR-B were expressed as GST fusion proteins. Chimeras GST-SLBR-B and GST-SLBR-H were produced as previously reported for SLBR-H[11] with minor modifications. The expression plasmids pGEX encoding GST-SLBR-B or GST-SLBR-H were used to transform competent *E. coli* strain BL21(DE3) (Invitrogen). The bacterial culture was grown in Luria Broth Medium using 50 µg/ml ampicillin at 37°C and then induced with 0.1mM isopropyl-1-thio-d-galactopyranoside (IPTG) overnight at 28°C, when A_{600} reached 0.6 nm. Cells were harvested by centrifugation and resuspended in phosphate saline buffer (PBS) pH 7.4. To obtain the lysate, cells were disrupted by sonication and again harvested by centrifugation to separate the soluble proteins from the pellet. SLBR-H and SLBR-B were purified using a Glutathione Sepharose 4B column (GE Healthcare) and eluted with 10 mM GSH and 10mM DTT in 50 mM Tris-HCl, pH 8.0. For NMR purposes, the eluted proteins were dialyzed against PBS pH 7.4.

Fluorescence analysis. The quenching fluorescence titration curves have been achieved by using a Fluoromax-4 spectrofluorometer (Horiba, Edison,NJ, USA). All the measurements were acquired at 10°C, upon excitation at 285 nm and recording the emission spectra in the range of 295–600 nm. The slit widths were chosen at 5 nm for the excitation and 10 nm for the emission wavelength. A quartz cuvette with a path length of 1 cm was used.

A fixed concentration of both SLBR-B and SLBR-H was chosen at 0.25 µM in 1,2 mL PBS buffer (pH 7.4) and titrated by adding small amounts of each ligand (sTa-Thr and 3'-SLn), ranging from 0 to 500 µM. No emission of the ligands was noted in the same region under the abovementioned conditions.

For the evaluation of the binding interactions, the change in fluorescence at 329 nm was followed for SBLRN-H and at 341 nm for SBLRN-B. The data were analyzed by non-linear regression with O*ne Site- Specific Binding* model for the determination of the dissociation constant (K_d) as implemented in in OriginPro 2016, according to the following equation:[2,3]

$$Y = \frac{B_{max} * X}{K_D + X}$$

where X represents the ligand concentration, Y is the fluorescence intensity change at the maximum wavelength, B_{max} is the maximum specific binding and K_D is the equilibrium dissociation constant.

NMR analysis. The NMR experiments were recorded on a Bruker AVANCE NEO 600-MHz equipped with a cryo probe and data acquisition and processing were performed with TOPSPIN 4.1.1 software. Samples were prepared in H_2O:D_2O (90:10) buffered with 50 mM phosphate at pH 7.4, using the 3mm NMR tubes. All the NMR spectra were acquired in the same conditions

for SLBR-H and SLBR-B, using protein concentrations of 20-30μM, depending on the NMR experiments, ligand concentrations of 1mM and temperatures of 288K and 298K.

STD NMR analysis. STD NMR experiments were acquired with 32 k data points and zero-filled up to 64 k data points prior to processing. The protein resonances were selectively irradiated using 40 Gauss pulses with a length of 50 ms,setting the off-resonance pulse frequency at 40 ppm and the on-resonance pulse at 6.5 and 7.5 ppm. An excitation sculpting with gradient pulses (esgp) was applied for the suppression of water signals. The %STD displayed in the ligands' epitope maps were obtained by the ratio of the STD signals in the STD spectra ($I_0 - I_{sat}$) and each relative peak intensity of the unsaturated reference spectrum (off-resonance, I_0), at saturation time of 2s. The highest STD signal was set to 100% and all the other STD were normalized to this value.

Tr-NOESY analysis. Homonuclear 2D ^{1}H-^{1}H NOESY experiments were carried out by using data sets of 2048x512 points and mixing times of 300 ms. The *inter*-proton cross relaxation rates (σij) were measured by the integration of the NOE cross peaks of interest and normalizing against the corresponding cross peak on the diagonal in F1. The experimental distances (rij) were then obtained by employing the isolated spin pair approximation.

WaterLOGSY analysis. The water-LOGSY experiments were performed on the ligand 3'-SLn in absence and in presence of SLBR-H, using a conventional 1D NOE-ePHOGSY pulse sequence[4] provided in the Bruker library (ephogsygpno.2). The bulk water was selectively irradiated by 1D NOESY and the solvent suppression was modified into a double pulsed field gradient (DPFG) perfect-echo.[5] In order to compare the WaterLOGSY experiment with the STD NMR experiment, a mixing time of 2s was set.

Docking and MD simulations. The ligands alone and bound to the Siglec and Unique domains of SLBR-H and SLBR-B were considered for running 100ns MD simulations by means of AMBER 18 suite[6]. First, the ligands were manually docked into the proteins binding pocket according to the published co-structures (PDB: 6EFD and 5IUC for SLBR-H and SLBR-B, respectively). The complexes of the proteins with each ligand were refined by pre-processing the structures using Maestro Protein Preparation Wizard.[7] Before the MD simulation, the complexes were minimized using Sander in Amber tools. Atom types and charges of the proteins were assigned according to AMBER ff14SB force field, while the ligands were represented with GLYCAM-06j-1 force field. The molecular structures were neutralized by adding Na+ ions and then hydrated considering octahedral boxes containing explicit TIP3P water molecules buffered at 10 Å by using the Leap module. The MD simulations were performed by using the pmemd.cuda implementation within Amber 18 package. The smooth particle mesh Ewald method was used for computing long-range electrostatic interactions while each simulation was under periodic boundary conditions, and the grid spacing was set to 1 Å. In the equilibration procedure, the system was minimized by applying a restriction to the protein which was gradually released in the following steps. Then slow system thermalization from 0KK to 300 K was carried out applying a solute restraint. Temperature was increased from 0K to 100K at constant volume. Then, from 100K to 300K in an isobaric ensemble and kept constant at 300 K during 50 ps with progressive energy minimizations and solute restraint. Once completed the restraints were removed and the systems then advanced in an isothermal-isobaric ensemble along the production. Concerning the complexes of SLBR-H and SLBR-B with sTa-THR, a restraint to the peptide dihedral angle (O-CB-CA-N) was applied to keep its value around 60 degrees. Coordinates were archived to acquire 10000 structures of the progression of the dynamics. Trajectories were analyzed using the ptraj module within AMBER18 and VMD program was used to visualize the MD results. Each trajectory was submitted to cluster analysis with respect to the ligand RMSD using K-mean algorithm implemented in ptraj module. The representative structure of the most populated cluster was considered to depict the complexes interactions. For the MD simulations including the sugar portion of GM1b, three of the most representative poses with each protein were depicted.

The determination of hydrogen bonds was calculated using the CPPTAJ module in AMBER 18. The h-bond is defined as occurring between an acceptor heavy atom A, a donor hydrogen atom H and a donor heavy atom D. The distance cut-off was set to 3 Å and the A-H-D angle cut-off was 135°.

The frequency of the protein-ligand bonds established during the dynamics is reported with a cut-off of 5Å.

Supporting figure caption

Figure S1 Quenching of intrinsic fluorescence. a) sTa-Thr bound to SLBR-B **b)** sTa-Thr bound to SLBR-H **c)** 3'-SLn boud to SLBR-H. Insets: Determination of dissociation constant (K_d) using one site binding analysis. ΔF represents the decrease in fluorescence intensity relative to the fluorescence intensity of the free protein on ligand binding. For each data point, 10% Y error bars were shown. The insets show the non-linear fits for the titrations with R^2 values of 0,98, 0.97, and 0,96 respectively.

Figure S2. MD simulation of sTa-Thr in the free state and bound to SLBR-H. a) φ/ψ dihedral angles of Neu5Ac-Gal linkage along the MD trajectory of sTa in the free state; **b)** φ/ψ dihedral angles of Gal-GalNAc linkage along the MD trajectory of sTa in the free state; **c)** φ/ψ dihedral angles of Neu5Ac-Gal linkage along the MD trajectory of sTa bound to SLBR-H; **d)** φ/ψ dihedral angles of Gal-GalNAc linkage along the MD trajectory of sTa bound to SLBR-H; **e)** SLBR-H and sTa-Thr ligand RMSD variation along the MD. **f)** Analysis of the most representative polar interactions between SLBR-H and sTa-Thr obtained by MD simulations (above the 10% of presence). Abbreviations from MD: 0SA corresponds to Neu5Ac, 3LB is Gal and 3VA is GalNAc.

Figure S3. Conformational analysis of sTa-Thr in the free and bound states. a) ROESY NMR spectrum of sTa-Thr in the free state. b) Tr-NOESY spectrum of sTa-Thr bound to SLBR-H. c) Tr-NOESY spectrum of sTa-Thr bound to SLBR-B.

Figure S4. NOESY NMR spectrum of 3'SLn in the free state. The key NOE B3-$K3_{ax}$ is circled (see the table in figure 3).

Figure S5. MD simulation of sTa-Thr in the free state and bound to SLBR-B. a) φ/ψ dihedral angles of Neu5Ac-Gal linkage along the MD trajectory of sTa bound to SLBR-B; **b)** φ/ψ dihedral angles of Gal-GalNAc linkage along the MD trajectory of sTa bound to SLBR-B; **c)** SLBR-B and sTa-Thr ligand RMSD variation along the MD. **d)** Analysis of the most representative polar interactions obtained by MD simulations (above the 10% of presence). Abbreviations from MD: 0SA corresponds to Neu5Ac, 3LB is Gal and 3VA is GalNAc.

Figure S6. MD simulation of 3'-SLn bound to SLBR-H. a) φ/ψ dihedral angles of Neu5Ac-Gal linkage along the MD trajectory of 3'-SLn bound to SLBR-H; **b)** φ/ψ dihedral angles of Gal-GalNAc linkage along the MD trajectory of 3'-SLn bound to SLBR-H; **c)** SLBR-H and 3'-SLn ligand RMSD variation along the MD, **d)** $H3_{ax}$ Neu5Ac/$H3_{eq}$ Neu5Ac/H8 Neu5Ac – H3 Gal inter-ligand distances. **e)** Analysis of the most representative interactions between SLBR-H and 3'-SLn obtained by MD simulations (above the 10% of presence). Abbreviations from MD: 0SA corresponds to Neu5Ac, 3LB is Gal and 4YB is GlcNAc.

Figure S7. **Analysis of the interactions between SLBR-H and 3'-SLn mediated by water molecules. a)** WaterLOGSY experiment of SLBR-H – 3'-SLn. **b)** Variation of Ser295 with O9 and H9O atoms of Neu5Ac, displaying the stability of the occurrence of distances along the MD simulation.

Figure S8. STD analysis of SLBR-H (a) and SLBR-B (b) in interactions with GM1b. In the upper side, epitope maps of the ligand recognized by SLBR-H (left) and SLBR-B (right) calculated by $(I_0-I_{sat})/I_0$, where (I_0-I_{sat}) is the intensity of the signal in the STD-NMR spectrum and I_0 is the peak intensity of the unsaturated reference spectrum (off-resonance). The highest STD signal intensity of the acetyl group of Neu5Ac was set to 100% and the other protons were normalized to this. In the bottom, the reference ^{1}H NMR spectra (black) and STD 1D NMR (red) of GM1b with SLBR-H (left) and SLBR-B (right).

Figure S9. 3D view of the representative poses of the most populated clusters derived from MD of **a)** SLBR-H – GM1b and **b)** SLBR-B – GM1b. The Neu5Ac-α-(2,3)-Gal-β-(1,3)-GalNAc portion is mainly involved in the interaction with both proteins.

Figure S10. Complexes of SLBR-H **(a)** and SLBR-B **(b)** with sTa-Thr (colored structures) superimposed to a hypothetical O-glycan expressed on mucins (in grey, structure PDB: 5t78): the orientation assumed by the threonine to the solvent is coherent with the presence of a potential amino acid chain linked to the sTa usually found in O-glycans.

Figure S11. **Comparison between Siglec-like adhesin and Siglec binding sites**. Upper panel: 3D view of SLBR-H - sTa-Thr (blue), SLBR-B - sTa-Thr (pink) and SLBR-H - 3'SLn (wheat) complexes, here studied, showing that the ligands are entirely accommodated into the binding site of the Siglec-like adhesins. Lower panel: examples of Siglecs that recognize α2,3 sialoglycans (3'-Sialyllactose, 3'SL and 3'-Sialyllactosamine, 3'SLn). The 3D complexes of Sialoadhesin/3'SL (green, pdb: 1qfo), Siglec-10/3'SLn (light blue) and Siglec-5/3'SL (yellow, pdb: 2zg3) show that only Neu5Ac and Gal residues are recognized by the proteins, exposing the reducing end of the glycan far away from the binding pocket of the Siglec.

References

[1] J. B. Schwarz, S. D. Kuduk, X.-T. Chen, D. Sames, P. W. Glunz, S. J. Danishefsky, A Broadly Applicable Method for the Efficient Synthesis of α-O-Linked Glycopeptides and Clustered Sialic Acid Residues. *J. Am. Chem. Soc.* 1999, 121, 12, 2662–2673.

[2] Sindrewicz P, Li X, Yates EA, Turnbull JE, Lian LY, Yu LG. Intrinsic tryptophan fluorescence spectroscopy reliably determines galectin-ligand interactions. Sci Rep. 2019 Aug 14;9(1):11851.

[3] Cook GW, Benton MG, Akerley W, Mayhew GF, Moehlenkamp C, Raterman D, Burgess DL, Rowell WJ, Lambert C, Eng K, Gu J, Baybayan P, Fussell JT, Herbold HD, O'Shea JM, Varghese TK, Emerson LL. Structural variation and its potential impact on genome instability: Novel discoveries in the EGFR landscape by long-read sequencing. PLoS One. 2020 Jan 15;15(1):e0226340..

[4] C. Dalvit, Homonuclear 1D and 2D NMR Experiments for the Observationof Solvent – Solute Interactions, J. Mag. Reason., B112,282 – 288 (1996).

[5] F. De Biasi, D. Rosa-Gastaldo, X. Sun, F. Mancin, F. Rastrelli, Nanoparticle-Assisted NMR Spectroscopy: Enhanced Detection of Analytes by Water-Mediated Saturation Transfer, J. Am. Chem. Soc. 2019, 141, 4870−4877.

[6] D.A. Case, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J. Mermelstein, K.M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. Salomon- Ferrer, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman (2018), AMBER 2018, University of California, San Francisco.

[7] Maestro v9.2. Portland, OR: Schrodinger, Inc.; 2012.