International Journal of Advanced Research in Biological Sciences
ISSN: 2348-8069
www.ijarbs.com
Volume 3, Issue 5 - 2016

Research Article

SOI: http://s-o-i.org/1.15/ijarbs-2016-3-5-35

Structural studies on binding ability of Siglec-3 to ligand using molecular modeling techniques

Madhumita Dandopath Patra*
Department of Chemistry, Acharya Prafulla Chandra College, Kolkata-700131, West Bengal, India
E-mail: *madhumita_ucst@yahoo.co.in*

Abstract

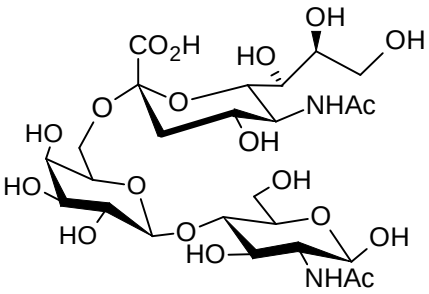
Siglecs, the major homologous subfamily of I-type lectins contains elements of sialic acid, immunoglobulin and lectin. The primary role of Siglecs is to recognize bacterial pathogens that express sialic acids. Siglec-3, a member of the Siglec family expressed on myeloid progenitor cells in the bone marrow and on peripheral blood monocytes. In this work, 3-D structure of human Siglec-3 was predicted using molecular modeling techniques. The structure of the complex in solution of Siglec-3 with ligand, 6′- SialylLacNAc: NeuAcα2,6Galβ1,4GlcNAc) containing α(2,6)-linked sialic acid was predicted using a novel docking technique. The structural analysis of the complex as well as theoretical dissociation constant value will help to ascertain functional roles of such sugar binding protein.

Keywords: Modeling, Sialic acid, Siglec-3, Structure

Introduction

The Siglecs, subgroup of the Ig superfamily, can recognize sialylated glycoconjugates (Crocker *et al.*, 1998). Sialic acid (Neu5Ac), a nine-carbon monosaccharide on the cell surface is responsible for non-specific electrostatic repulsion between cell types and to mediate cell adhesion, protein–protein interactions, and protein trafficking. (Blixt *et al.*, 2003; Kelm *et al.*, 1994; Brinkman-Van der Linden and Varki, 2000; Crocker *et al.*, 1998; Cornish *et al.*, 1998; Powell and Varki, 1994; Nicoll *et al.*, 1999; Varki, 1997; Karlsson, 1998; Crocker and Varki, 2001). Siglecs are type 1 membrane proteins, with an N-terminal sialic acid-binding V-set Ig domain, transmembrane domain, a cytoplasmic tail and variable number of C2-set Ig-like domains Angata *et al.*, (2001). Siglecs can be divided into two subgroups: Sialoadhesin (Siglec-1), Siglec-2 (CD22), MAG (Siglec-4) and Siglec-15 constitute one subgroup, share ~25–30% sequence identity in the extracellular region, and the second subgroup consists of the CD33-related Siglecs. They share 50–80% sequence similarity and have in their cytoplasmic tails two highly conserved tyrosine-based motifs. Expression of each human Siglec in a cell type-specific fashion, suggesting involvement in discrete functions ranging from maintenance of myelination in the nervous system (MAG) (Li *et al.*, 1994; Montag *et al.*, 1994) and control of myeloid cell interactions (sialoadhesin [Crocker *et al.*, 1997]) and CD33 (Freeman *et al.*, 1995) and activation of B cells (CD22 [reviewed in Cyster and Goodnow, 1997]) .

In the present study, I have predicted the 3-D structure of human Siglec-3 (hSiglec-3) along with the specific ligand, 6′-SialylLacNAc (Fig. 1). The structural analysis of the predicted complex was done. The theoretical dissociation constant value was also calculated for the complex which helped me to compare the relative binding affinity.



6′- SialylLacNAc: NeuAcα2,6Galβ1,4GlcNAc)

Fig. 1: Glyco-chain structure of the ligand used in this study.

Materials and Methods

The starting scaffold for modeling was the x-ray crystallographically determined structure of hSiglec-7 (PDB ID: 107V) in its unliganded from. The initial structure of hSiglec-3 was obtained using our in-house software package of ANALYN and MODELYN (Mandal, C., 1998). At first, the ligand molecule was constructed using 'BUILDER' module of insightII and superposed taking the sialic acid part of the ligand with the equivalent part of the x-ray structure containing GT1b (2HRL). Then the modeled protein was superposed with the GT1b bound protein (2HRL) with respect to the structurally conserved regions followed by transfer of the superposed 6′-SialylLacNAc molecule to the binding site. The structural refinement was using DISCOVER module of InsightII 2005 of Accelrys (San Diego, CA). Structural optimization was done using cff91 force-field and energy minimization (100 steps each of steepest descent and conjugate gradient methods) followed by dynamics simulations. At the end of the dynamics simulation, lowest potential energy conformation with was picked for the next cycle of refinement using the module ANALYSIS of InsightII. This combination of dynamics and minimization were repeated until satisfactory conformational parameters were obtained.

Water molecules were added using the Assembly/Soak option of Insight II, as a sphere of radius 18 Å having its center at an atom roughly at the center of the ligand molecule so as to surround it completely to investigate the influence of water on the ligand binding. In the aqueous environment, structure optimization of the ligand was done using energy minimization and molecular dynamics simulation in presence and absence of the protein molecule. From the values of the free energies of complex formation of ligand in water and water-protein environments the absolute binding energy was calculated using the $\Delta G_{bind} = \alpha \Delta \langle V^{el}_{l\text{-}s} \rangle + \beta \Delta \langle V^{vdw}_{l\text{-}s} \rangle$ where ΔG_{bind} is the absolute binding energy, Δ stands for differences in the electrical ($V^{el}_{l\text{-}s}$) and van der Waals ($V^{vdw}_{l\text{-}s}$) components of the free energies of the ligand solvent (l-s) systems i.e. in pure water and protein containing water environments following the linear interaction energy approximation method of (Aqvist *et al*., 1994) The weight factors of the electrical and van der Waals contributions were as as 0.5 (α) and 0.16 (β) respectively as proposed by Åqvist *et al*. and used by earlier workers (Aqvist and Mowbray, 1995; Hulte´n *et al*., 1997). K_a was calculated using the thermodynamic relation $\Delta G_{bind} = -RT\ln K_a$ where R is the ideal gas constant and T is the absolute temperature. Dissociation constant K_d was calculated by taking the inverse of K_a (association constant).

MODELYN was run in the windows environment and in the IRIX environment. Altrix 350 server of Silicon Graphics, Inc. in the IRIX environment and FUEL workstation were used to run Insight II. The electrostatic potential surface of the protein was determined by MOLMOL (Koradi *et al*., 1996) and structural parameters were checked using the PROCHECK (Laskowski *et al*., 1993). The binding affinity of the Siglec-ligand complex was obtained using the DOCKING module of InsightII.

Results

General structural characteristics of the predicted model was determined by measuring all the bond distances and bond angles and the deviation of these parameters from the standard values were calculated for appropriate types of bonds and angles. By calculating the phi and psi dihedral angles and drawing Ramachandran's plots, the quality of backbone conformations were determined for the structure. Table 1 presents the RMSD (root mean square deviation) of bond lengths and bond angles of the predicted structure along with the percentages of

backbone Phi-Psi angles in different areas of Ramachandran's plots obtained after the prediction of 3D structures.

RMSD from the respective standard values of the bond lengths around 0.02 Å and those of bond angles around 3 degrees indicate good general structural parameters of the modeled structure. The good quality of the backbone conformations of the modeled structure indicated by the values of above 95% Phi-Psi pairs in the core and allowed areas of Ramachandran's plot.

Table 1: General and backbone structural parameters of the modeled structure of the target sequence as well as the x-ray structure of the Siglec.

Siglecs	Accession No	% of AA Identity (positive score)	RMS deviation		% of Phi-Psi pairs in the area			
			Bond (Å)	Angle (°)	Core	Allowed	Generously allowed	Dis-allowed
hSiglec-7	107V	100	0.015	2.35	87.7	10.5	0.9	0.9
hSiglec-3	NP_001763	57(70)	0.012	2.75	89.9	10.1	0.0	0.0

PROCHECK was used for side chain planarity of the planar groups in phenylalanine, tyrosine, tryptophan, histidine, arginine, glutamine, asparagines, glutamic acid, and aspartic acid and deviations from planarity were identified by measuring RMS (root mean square) distances of planar atoms from the best-fitted plane; residues having RMS distances >0.03 Å for rings and 0.02 Å for other groups were marked as outliers (Laskowski *et al*., 1993) (Table 2). We checked protein geometry of the modeled structure by calculating clashscores and rotamer outliers using MOLPROBITY (Davis *et al*., 2004) (Table 2).

Table 2: General and backbone structural parameters of the modeled structure of the target sequence in comparison with the x-ray structure of the Siglec.

Siglecs	Accession No	All atom clashcore (per 1000 atom)	Rotamer outliers (%)	Planarity outliers (%)
hSiglec-7	107V	6.50	1.77	0.0
hSiglec-3	NP_001763	5.52	2.65	3.70

CD33, an immunoglobulin (Ig) superfamily protein expressed on peripheral blood monocytes and myeloid progenitor cells in the bone marrow (Brinkman-Van der Linden *et al*, 2003). It has also been reported on dendritic cells, cord blood–derived natural killer (NK) cells, in vitro–expanded T cells, and some bi-phenotypic leukemias (Grobe and Powell, 2002). CD33 is known to recognize α(2,3)- and α(2,6)-linked sialic acids which are expressed mostly at the non-reducing termini (outermost positions) of glycan chains (Brinkman-Van der Linden *et al*., 2003). CD33 (Siglec-3) is becoming increasingly important as a target of antibody mediated therapy in acute myeloid leukaemia (AML) (Hauswirth *et al*, 2007). We have modeled the 3-D structure of hSiglec-3 taking the crystal structure of hSiglec-7 (PDB ID: 107V) in its unliganded from as template. hSiglec-3 (Accession No: NP_001763) showed 57% sequence identity and about 70% sequence similarity with the query sequence. Modeled structure was refined as mentioned in the materials and method section. The ligand 6′-SialylLacNAc: NeuAcα2,6Galβ1,4GlcNAc)containing α(2,6)-linked sialic acid and specific for hSiglec-3 was docked into the binding site of the protein (Varki and Angata, 2006). The ligand occupies a large area of the binding face of hSiglec-3 and makes a number of contacts with the protein. The 6′-SialylLacNAc (three times) molecule makes six direct hydrogen bonds with the protein, five water-mediated hydrogen bonds and a number of hydrophobic interactions (Fig. 2). These interactions are summarized in Table. 3.

Table 3: Hydrogen-bond network within the binding site of hSiglec-3 in complex with 6′-SialylLacNAc. Glycosidic linkages are shown in parenthesis. Distances are measured between hydrogen and acceptor or donor atom.

Ligand–protein hydrogen-bonds		
Atoms of 6′-SialylLacNAc	Atoms of hSiglec-3	Distance(Å)
Neu5Ac (α2,6)		
O1B	Ser133 N/OG	2.44/2.10
N5	Lys131 CO	1.94
Nag		
O6	Arg124 NE	2.04
O5	Arg 124 NH2	2.10
N2	Asp 75 OD1	1.92
Ligand–protein hydrogen-bonds mediated by water		
Atoms of 6′-SialylLacNAc	Atoms of hSiglec-3	*Distance(Å)
Nag O1	Arg 67 O/Gly 74 O	5.62/5.19
Nag O1	Asp 75 O/ Arg 124 NH2	4.04/3.80
Neu5Ac (α2,6) O4	Lys131 N/Ser 129 CO	4.74/5.44
Neu5Ac (α2,6) O10	Thr130 OG1	4.95/5.89
Neu5Ac (α2,6) O9	Asp23 OD1/OD2	4.96/4.01
Intramolecular hydrogen-bonds		
Atoms of 6′-SialylLacNAc	Atoms of 6′-SialylLacNAc	Distance(Å)
Neu5Ac (α2,6) O4	Neu5Ac O1A	1.81
Neu5Ac (α2,6) O8	Neu5Ac O1B	1.82
Neu5Ac (α2,6) O7	Neu5Ac O9	1.83
Gal O4	Nag O6	1.90

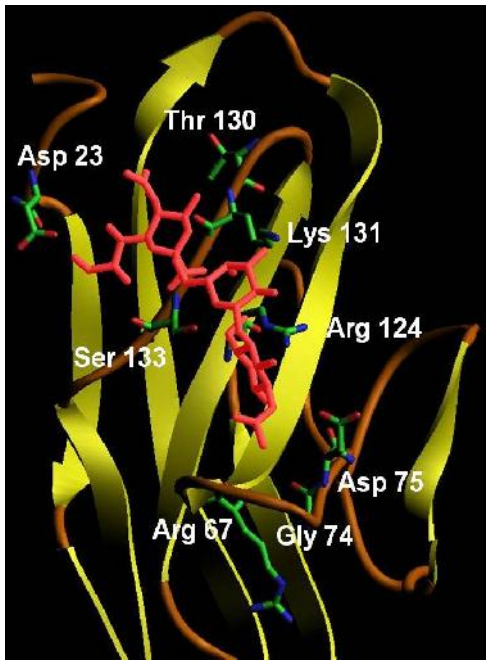


Fig. 2: Mode of ligand binding in hSiglec-3: Ligand binding environment is shown in the secondary structure environment of the modeled lectin. Beta sheets are shown in yellow with an arrow indicating the C-terminus and random coils as thin cylinder coloured in maroon. The residues of the protein involved in hydrogen bonding with the ligand are shown in stick representation, coloured as atoms (C=Green, O=Red and N=Indigo) and the ligand 6′-SialylLacNAc in red colour.

Table 4: Empirical free energies, their difference in water and water-protein environments and corresponding ΔG and K_d values for the complex formation between hSiglec-3 and the specific ligand in the aqueous solution. Abbreviations used in this table: VdW, van der Waals; Elect, electrical.

complex	Free energy in kcals/mol			Difference		ΔG_{bind} in kcals/ mol	K_d
	Vdw	Elect	Total	Vdw	Elect		
Siglec-3-6′-SialylLacNAc in solution	-94.51	-204.59	-299.10	-14.53	-7.31	-5.98	46.77 μM
6′- SialylLacNAc *	-79.98	-197.28	-277.26				

* Value corresponding to the interaction energy in presence of water molecules only.

This geometry is supported by interactions with the protein and by four intramolecular hydrogen bonds. The majority of the observed intermolecular interactions involve the sialic acid moiety of the ligand are the same as reported in hSiglec-7 complexes (Attrill *et al*., 2006a; Attrill *et al*., 2006b; Alphey *et al.*, 2003). One direct hydrogen-bonding between C-5 nitrogen of sialic acid with Lys-131 backbone oxygen is conserved. But one of the major salt bridge formation interactions between the sialic acid carboxylate and the guanidinium group of Arg-124 is absent which is reflected in the low dissociation constants (K_d) value of the hSiglec-3 and 6′-SialylLacNAc complex. Arg-124 is involved in direct hydrogen-bonding with N-acetyl glucose. Sialic acid also makes one direct hydrogen bonds with the protein via carboxylate group to Ser-133 and three water - mediated hydrogen bonds with Lys-131, Thr-130 and Asp-23. The C-C' loop (residues 68-74) participates in ligand binding via two water-mediated hydrogen-bonding through residues Arg-67, Gly-74, Asp-75 and one direct hydrogen-bonding through the residue Asp-75 with N-acetyl glucose of the ligand as mentioned in case of hSiglec-7 complexes (Attrill *et al*., 2006a; Attrill *et al*., 2006b; Alphey *et al*., 2003). The calculated ΔG_{bind} value using linear interaction energy approximation method as described in materials and methods section for the complex of hSiglec-3 and 6′-SialylLacNAc is negative indicating that the complex formation in the aqueous medium is thermodynamically favorable (Table.4). The ΔG_{bind} value leads to dissociation constants (K_d) of 46.77 μM.

Discussion

I have modeled the 3-D structure of human Siglec-3. The predicted structure was refined to obtain best backbone and sidechain conformations. The structure encompassed the important segments known to participate in their biological activities.

I have also predicted the structure of the complex of the modeled human Siglec-3 with the specific ligand, 6′-SialylLacNAc known so far from experimental studies. The nature of interactions of the ligand with the Siglec-3 was examined in details in order to understand the origin of their specificity at the atomic levels. The involvement of the crucial amino acids, identified by experimental techniques, was confirmed from the modeled structure by exploring the involvement of evolutionary conserved amino acids. Binding constant was predicted for the modeled complex and compared with the experimental values. Thus, my structural studies using predicted model of human Siglec-3 and the complex with specific ligand, 6′-SialylLacNAc have contributed significantly in understanding the interactions involving sialic acid containing bioactive molecules which are implicated in many important biochemical phenomena.

Acknowledgments

The author thank Dr. Chhabinath Mandal, Ex-Scientist, Structural Biology and Bioinformatics Division, Indian Institute of Chemical Biology, Jadavpur, India for his support.

Abbreviations

Glc, Glucose; Gal, Galactose; GlcNAc, N-acetyl glucose; GalNAc, N-acetyl galactose; LacNAc, N-acetyl lactose; NeuAc/ Neu5Ac, N-acetyl neuraminic acid (sialic acid); hSiglec, human Siglec; sia, sialic acid.

References

Alphey, M.S., Attrill, H., Crocker, P.R., van Aalten, D.M.F. 2003. High resolution crystal structures of Siglec-7. Insights into ligand specificity in the Siglec family. *J. Biol. Chem*., 278: 3372–3377.

Angata, T., and Varki, A. 2000. Siglec-7: a sialic acid-binding lectin of the immunoglobulin superfamily. *Glycobiology*, 10: 431-438.

Angata, T., Hingorani, R., Varki, N. M., and Varki, A. 2001. Cloning and characterization of a novel mouse Siglec, mSiglec-F: differential evolution of the mouse and human (CD33) Siglec-3-related gene clusters. *J. Biol. Chem*., 276: 45128–45136.

Aqvist, J.; Medina, C.; Samuelsson, J.-E. 1994. A New Method for Predicting Binding Affinity in Computer-Aided Drug Design. *Protein Eng*., 7: 385-391.

Aqvist, J.; Mowbray, S.L. 1995. Sugar recognition by a Glucose/Galactose receptor: Evaluation of binding energies from molecular dynamics simulations. *J. Biol. Chem.*, 270(17): 9978-9981.

Attrill, H., Takazawa, H., Witt, S., Kelm, S., Isecke, R., Brossmer, R., Ando, T., Ishida, H., Kiso, M., Crocker, P.R., van Aalten, D.M.F. 2006a.The structure of Siglec-7 in complex with sialosides: leads for rational structure-based inhibitor design. *Biochem. J.*, 397: 271-8.

Attrill, H., Imamura, A., Sharma, R.S., Kiso, M., Crocker, P.R., van Aalten, D.M 2006b. Siglec-7 undergoes a major conformational change when complexed with the alpha (2,8)-disialylganglioside GT1b. *J. Biol. Chem*., 281: 32774-83.

Blixt, O., Collins, B.E., Van den Nieuwenhof, I.M., Crocker, P.R., Paulson, J.C. 2003. Sialoside specificity of the Siglec family assessed using novel multivalent probes: identification of potent inhibitors of myelin associated glycoproteins, *J. Biol. Chem.*, 278: 31007–31019.

Brinkman-Van der Linden, E.C., Varki, A. 2000. New aspects of siglec binding specificities, including the significance of fucosylation and of the sialyl-Tn epitope. *J. Biol. Chem*., 275: 8625– 8632.

Brinkman-Van der Linden, Els C. M., Angata, T., Reynolds, S.A., Powell, L.D., Hedrick, S.M., Varki, A. 2003. CD33/Siglec-3 Binding Specificity, Expression Pattern, and Consequences of Gene Deletion in Mice. *Molecular and Cellular Biology*, 23: 4199–4206.

Cornish, A.L., Freeman, S., Forbes, G., Zhang, J. Ni, M., Cepeda, M., Gentz, R., Augustus, M., Carter, K.C., Crocker, P.R. 1998. Characterization of siglec-5, a novel glycoprotein expressed on myeloid cells related to CD33, *Blood*, 92: 2123– 2132.

Crocker, P. R., and Varki, A. 2001. Siglecs, sialic acids and innate immunity. *Trends Immunol*., 22: 337–342.

Crocker, P.R., Hartnell, A., Munday, J., Nath, D. 1997. The potential role of sialoadhesin as a macrophage recognition molecule in health and disease. *Glycoconj. J.*, 14: 601–609.

Crocker, P. R., Clark, E. A., Filbin, M., Gordon, S., Jones, Y. Kehrl, J. H., Kelm, S., Le Douarin. N., Powell. L., Roder. J., Schnaar. R. L., Sgroi, D. C., Stamenkovic, K., Schauer, R., Schachner, M., van den Berg, T. K., van der Merwe, P. A., Watt, S. M., and Varki, A. 1998. *Glycobiology*, 8: v

Cyster, J.G., and Goodnow, C.C. 1997. Tuning antigen receptor signaling by CD22: integrating cues from antigens and the microenvironment. *Immunity*, 6: 509–517.

Davis, I.W., Murray, L.W., Richardson, J.S., Richardson, D.C. 2004. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Research*, 32: W615–W619.

Freeman, S.D., Kelm, S., Barber, E.K., Crocker, P.R. 1995. Characterization of CD33 as a new member of the sialoadhesin family of cellular interaction molecules. *Blood*, 85: 2005–2012.

Grobe, K. and Powell, L.D. (2002). Role of protein kinase C in the phosphorylation of CD33 (Siglec-3) and its effect on lectin activity. *Blood*, 99, 3188-96.

Hauswirth, A.W, Florian, S., Printz, D., Sotlar, K., Krauth, M.T., Fritsch, G., Schernthaner, G.H., Wacheck, V., Selzer, E., Sperr, W.R., Valent, P. 2007. Expression of the target receptor CD33 in CD34+/CD38-/CD123+ AML stem cells. *Eur. J. Clin. Invest.*, 37(1): 73-82.

Hulte´n, J.; Bonham, N.M.; Nillroth, U.; Hansson, T.; Zuccarello, G.; Bouzide, A.; Aqvist, J.; Classon, B.; Danielson, U.H.; Karle´n, A.; Kvarnstro¨m, I.; Samuelsson, B.; Hallberg, A. 1997. Cyclic HIV-1 Protease Inhibitors Derived from Mannitol: Synthesis, Inhibitory

Potencies, and Computational Predictions of Binding Affinities. *J. Med. Chem.* 40: 885-897.

Karlsson, K. A. 1998. Meaning and therapeutic potential of microbial recognition of host glycoconjugates. *Mol. Microbiol*., 29: 1–11.

Kelm, S., Schauer, R., Manuguerra, J.C., Gross, H.J., Crocker, P.R. 1994. Modifications of cell surface sialic acids modulate cell adhesion mediated by sialoadhesin and CD22, *Glycoconj. J*., 11: 576– 585.

Koradi, R., Billeter, M., Wuthrich, K. 1996. Related Articles, MOLMOL: A program for display and analysis of macromolecular structures *J. Mol. Graph.* 14: 51–55.

Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thornton, J.M. 1993. PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Cryst.* 26: 283-91.

Li, C., Tropak, M. B., Gerlai, R., Clapoff, S., Abramow-Newerly, W., Trapp, B., Peterson, A., Roder, J. 1994. Myelination in the absence of myelin-associated glycoprotein. *Nature*, 369: 747-750.

Mandal C 1998. MODELYN – A molecular modelling program version PC-1.0 Indian Copyright No 9/98.

Montag, D., Giese, K.P., Bartsch, U., Martini, R., Lang, Y., Bluthmann, H., Karthigasan, J., Kirschener, D.A., Wintergerst, E.S., Nave, K.A., et al. 1994. Mice deficient for the myelin-associated glycoprotein show subtle abnormalities in myelin. *Neuron*, 13: 229–246.

Nicoll, G., Ni, J., Liu, D., Klenerman, P., Munday, J., Dubock, S., Mattei, M.G., Crocker, P.R. 1999. Identification and characterization of a novel siglec, siglec-7, expressed by human natural killer cells and monocytes, *J. Biol. Chem.*, 274: 34089–34095.

Powell, L.D., Varki, A. 1994. The oligosaccharide binding specificities of CD22 beta, a sialic acid-specific lectin of B cells, *J. Biol. Chem*., 269: 10628– 10636.

Varki, A. 1997. Sialic acids as ligands in recognition phenomena. *FASEB J*., 11: 248–255.

Varki, A. and Angata, T. 2006. Siglecs—the major subfamily of I-type lectins, *Glycobiology*, 16: 1R–27R.

Access this Article in Online	
	Website: www.ijarbs.com
Quick Response Code	**Subject:** Chemistry

How to cite this article:
Madhumita Dandopath Patra. (2016). Structural studies on binding ability of Siglec-3 to ligand using molecular modeling techniques. Int. J. Adv. Res. Biol. Sci. 3(5): 254-260.