

Abstract

Glycodelins (Gds) are dimeric, glycosylated proteins belonging to the lipocalin superfamily which are primarily transporters of small hydrophobic molecules. Gd is classified under the lipocalin superfamily as it shows 55 % sequence identity to β -Lactoglobulin (BLG), a prominent member of the lipocalin family. There are several isoforms of glycodelin identified namely GdA, GdS, GdF and GdC which are distributed in different reproductive tissues and perform various functions. Although all isoforms share a common protein backbone, they differ in the glycosylation profile which in turn dictates the activity of these proteins. One of the well-studied, multifunctional glycodelin isoforms is GdA, which is a novel pregnancy-related galectin like lectin. GdA is a 162 amino acid long, dimeric, secretory protein synthesized by decidualized endometrium and secreted into the uterine luminal cavity during the luteal phase of an ovulatory cycle. It is the major progesterone-regulated protein as its levels are correlated well with the levels of progesterone. GdA is known to perform a variety of functions like immunosuppression, contraception, neo-vascularization, cell differentiation and tissue remodeling which makes it an important molecule to study. The important contraceptive activity of GdA is shown to be contributed by the unique glycosylation pattern present in it, indicating glycosylation to be an important post translational modification regulating the function of glycodelin's.

GdA is also known to play a critical role in the establishment, maintenance, and progression of pregnancy as subnormal levels of it are associated with early pregnancy loss, recurrent miscarriages, and unexplained infertility. Fetus is a semi allograft, which contains paternal antigens that can be considered as foreign by the maternal immune system. There are many factors which work in concert to protect the fetus from getting rejected. GdA is one such factor known to modulate the maternal defense mechanisms by affecting both the innate and adaptive arms of the immune system i.e., GdA inhibits proliferation of B cells, inhibits activity of NK cells, induces apoptosis in monocytes and activated T cells. Earlier studies from our laboratory have shown that the protein backbone of GdA is sufficient to induce apoptosis in T cells and the glycans modulate this activity. Thus, to understand the mechanism of action of GdA and to determine the domain of GdA

involved in apoptogenic activity, this study was aimed at mapping the functional domain/s of the GdA. The overall results obtained in the study were categorized into three chapters, each describing various bioinformatic and experimental approaches employed to identify the functional domain of GdA involved in inhibition of proliferation/inducing apoptosis in T cells.

Chapter 1 discusses the deletion analysis carried out on glycodelin gene to obtain various truncated forms of the protein, which were subjected to activity assay (i.e., inhibition of proliferation and induction of apoptosis) to further narrow down on the region (functional domain) of GdA involved in inducing apoptosis in T cells. The truncated gene constructs were cloned and expressed in baculoviral system, as the recombinant protein expressed only in this system showed activity like that of the native GdA. The N-terminal deletion construct Gd Δ N2 in which 40 amino acids from the N terminus of GdA had been deleted showed activity similar to wildtype glycodelin indicating that the N terminal 40 amino acids are not required for the apoptogenic activity of the GdA. The second deletion construct Gd Δ C2, in which 57 amino acids from the C terminus had been deleted showed reduced activity, as its IC₅₀ value (estimated by inhibiting the proliferation of Jurkat T cells) were 5-10 fold higher than the wildtype glycodelin. The results further suggest that the functional domain of the GdA protein may probably lie towards the C terminus of the truncated protein, Gd Δ C2 thus resulting in reduced activity. The deletion of glycodelin gene further from either N or C terminus, lead to insolubility of the truncated proteins hindering us from using this approach to characterize the functional integrity of these proteins. Hence, we resorted to mutagenesis approach to map the functional domain of GdA.

The results of mutagenesis approach to further map the functional domain of GdA is presented in **chapter 2**. To identify the residues of GdA which must be mutated to determine its functional domain we employed bioinformatic approach. The primary aim of the tool (surface probability by Emini) used was to identify the amino acids or regions in GdA protein which have high probability of being on the surface of the protein and thus likely to bind to the receptor on T cells to trigger apoptosis in these cells. As the structure of GdA is not available in literature, we went ahead to compare the sequence of GdA with a well-studied non-immunosuppressive member of lipocalin family, the β -Lactoglobulin

(BLG). As BLG is a homolog of GdA, both the protein sequences were compared to identify regions in GdA which were different from BLG and were also surface exposed. The region between E77 to E128 in GdA was found to be interesting as this region had high probability of being surface exposed, had amino acid sequence dissimilar to BLG and more importantly formed a part of the probable apoptogenic region as identified in chapter 1. The high surface probable amino acids in the identified region of GdA were mutated to the corresponding amino acid residue of BLG and the mutants were characterized. The mutants generated in region E77 to E128 were functionally as active as wildtype glycodeclin indicating that this region may not be a part of the functional domain of the GdA protein. Mutagenesis was then carried out in the second most surface probable region which spanned between D52 to S65. The proteins generated post mutation in D52-S65 region showed reduced activity as the IC₅₀ values (estimated by inhibition of proliferation of Jurkat T cells) were 5-7-fold higher than wildtype glycodeclin indicating that the apoptogenic region lies between D52 to S65. Since, from literature we know that GdA is a galactose specific lectin, we were interested to know whether it has multiple sugar binding pockets like other lectins. So, we expressed in the baculoviral system a double truncated protein, Gd (M24...L105) in which 23 amino acids from N terminus and 57 amino acids from the C terminus had been deleted. This truncated protein was found to be active indicating region from M24...L105 to be sufficient to inhibit proliferation and induce apoptosis in T cells like wildtype glycodeclin.

Since GdA is a known dimeric lectin, we investigated the role of oligomerization on the apoptogenic activity of this protein in **chapter 3**. In BLG, the residues R40, H146 and R148 were known to be part of domain forming the dimer interface and for maintaining the monomer-dimer equilibrium. Since the structure of GdA is not known and BLG is a close homolog of GdA, we decided to investigate the role of these residues on the oligomerization of GdA. Hence these residues in GdA were mutated to identify their role in oligomerization of GdA. The result suggest not only the single mutants but even the triple mutant of these amino acids belonging to the dimer interface did not affect the oligomeric status and also the apoptogenic activity of GdA indicating that these residues may not be involved in the oligomerization of GdA. Thus, the results implied that the residues involved in dimerization of BLG are not the same for GdA and even though both

these proteins show high sequence similarity and identity, there are subtle differences in their structures which account for the uniqueness in their functions.

To summarize, our study suggests that the functional domain of the GdA protein resides in the region, M24...L105 as a truncated protein containing only this region was shown to be sufficient to induce apoptosis in T cells. Thus, this region from M24...L105 appears to be necessary for GdA to perform its function whereas the mutagenesis approach within this region showed that the functional domain most probably lies between D52...S65.