Intelectins, also known as X-lectins or eglectins, are secreted proteins with Ca^{2+}-dependent sugar binding activity. They are found in chordates, but not in invertebrate metazoans and have roles in physiological processes ranging from fertilization and embryogenesis to host defense. Intelectins have a characteristic, conserved sequence motif (GGWTLVASVHEN) (Lee et al., 2004). They have a non-conserved N-terminal signal sequence and a single putative carbohydrate recognition domain (CRD) (Lee et al., 2004), which has no known similarity to any other protein domain, except an N-terminal region that is related to the globular domain of fibrinogen and to ficolins.

The first X-type lectin to be discovered was the X-cortical granule lectin of *Xenopus laevis*, which plays a significant role in fertilization and embryogenesis through the recognition of endogenous glycan ligands (Lee et al., 2004). Homologs have since been identified and studied in fishes, amphibians, reptiles and mammals (Abeet al., 1999) (French et al., 2009) (Nagata, 2005) (Tsuji et al., 2007) (Tsutsui et al., 2011) (Yokosawaet al., 1982) (Darville et al., 2012))(Yan et al., 2013) (Yan et al., 2012). The main function of these X-type lectins appears to be host defense through innate immunity.

Human intelectins 1 (hITLN1) and 2 (hITLN2) are present on the chromosome 1 (1q21.3 and 1q22-23.5, respectively) in the endothelial cell adhesion locus that also houses the selectin family of lectins (Tsuji et al., 2001). hITLN1, also known as omentin, intestinal lactoferrin receptor and endothelial lectin HL-1, is a disulfide-bonded, N-glycosylated, trimeric, secreted protein that is expressed in the endothelial cells of the intestine, heart, testis, colon, salivary gland, skeletal muscle, pancreas and thyroid, and to a lesser extent in the

uterus, spleen, prostrate, lymph node and thymus (Tsuji et al., 2001) (Suzuki et al., 2001) (Lee et al., 2004)(Lee et al., 2001).

Despite being known for more than a decade, few biochemical studies of intelectins were reported in the literature, and there was no report detailing the ligand binding parameters of these proteins, likely due to technical challenges in purifying and working with these disulphide-rich, Ca^{2+}-binding, secreted proteins, at the time of initiation of this thesis work. Recently, in year 2016 during the course of this thesis, Wesener *et al* published the 1.6-Å-resolution crystal structure of β-Galf-bound hITLN1 (Wesener et al., 2015). The studies indicated that hITLN1 discriminates between microbial and human glycan epitopes through Ca^{2+} ion-dependent coordination of an exocyclic, terminal 1,2-diol, which is a component of multiple microbial saccharides (Wesener et al., 2015) and not by binding to furanose monosaccharides such as galactofuranose and ribofuranose which are present exclusively in microbes), as Tsuji *et al* had previously suggested (Tsuji et al., 2001).

Biochemical studies reported in the literature suggest that despite the overall high sequence similarity, intelectins from various chordates have different monosaccharide specificities. hITLN1 binds to D-galactofuranose (Tsuji et al., 2001), *Halocynthia roretzi* intelectin binds to D-galactose (Abe et al., 1999); *Oncorhynchus mykiss* intelectin binds to N-acetylglucosamine (Russell et al., 2008); *Branchiostoma floridae* intelectins bind to bacterial lipopolysaccharide and peptidoglycan (Yan et al., 2012); alligator intelectin binds to mannose (Darville et al., 2012); and *Xenopus laevis* XL35 lectin binds to melibiose (Lee et al., 2004). Although most studied, questions still remain about the saccharide binding specificity of hITLN1 and its homologs.

The expression of hITLN1 is associated with various diseases and in inflammation, and therefore ligand binding by hITLN1 has a biological significance. hITLN1 is known to be upregulated in the airway epithelium of asthmatics and is required for IL-13 induced expression of chemokines, MCP-1 and MCP-3. MCP-1 and MCP-3 expression was downregulated in mice by attenuating the hITLN1 expression by shRNA and galactose, a known ligand of hITLN1, thus inhibiting airway inflammation (Gu et al., 2010).

Since ligand binding by hITLN1 can potentially modulate the outcome of disease, the focus of the work outlined in the thesis was to screen easily available saccharides to obtain some potent ligands of hITLN1 which in future may be explored to determine effect on hITLN1 levels and disease pathology in various disease conditions. Moreover, the ligand binding by hITLN1 was reassessed in the light of the findings published by Wesener *et al* in 2015. Apart from this ligand binding studies of intelectin homologs from *Halocynthia roretzii* (HaloITLN), *Xenopus laevis* (XL35ITLN) and an uncharacterized bacterial lectin with FReD domain from *Pseudomonas syringae* were also conducted with an objective to compare and understand ligand binding specificity of members of the X-type family of lectins.

One of the main objectives was to obtain purified recombinant hITLN1 in sufficient yield so as to enable biochemical studies. To accomplish this goal, various strategies were adopted to clone, express and purify this difficult-to express protein, which are described in **Chapter 2**. Full length hITLN1 and hITLN2, retaining their N-terminal signal peptide sequences to enable protein secretion into medium supernatants were successfully cloned in pcDNA5/FRT/V5-His-TOPO

vector having the strong CMV promoter and co-transfected with pOG44 vector (expressing FLP recombinase) to generate Flp-In-CHO stable cell lines. The stable cell lines were generated successfully, as assayed by the expression of tagless hITLN1 and hITLN2 in medium supernatants of the Flp-In-CHO cells by immunoprecipitation and flow cytometry. However, hITLN1 and hITLN2 expression levels were very low. Transiently transfection of Flp-In-CHO with hITLN1 and hITLN2 genes (cloned in the pcDNA5/FRT/V5-His-TOPO vector) was also attempted and secreted hITLN2 and hITLN2 were detected, but recombinant protein expression was insufficient to enable purification. Flp-In CHO cell lines were also transiently transfected with hITLN1 cloned in pEFBOS vector (a kind gift from Dr. Shoutaro Tsuji). On optimizing various conditions, tagless hITLN1 could be successfully expressed and purified in active form by Sepharose CL-6B affinity chromatography from hITLN1-pEFBOS transfected Flp-In-CHO cells. The protein yield was approximately 1 mg from 1 litre of cell culture media. The protein binding was Ca^{2+} dependent and EDTA did not have an irreversible effect on lectin activity. The purified hITLN1, in contrast to Tsuji's report, did not precipitate upon addition of Ca^{2+} and retained lectin activity. Thus, hITLN1 could be successfully expressed and purified from Flp-In-CHO cells transiently transfected with hITLN1-pEFBOS. Under the same conditions, optimal expression of tagless and 6XHis tagged hITLN2 and 6XHis tagged hITLN1 was not achieved. hITLN1 and hITLN2 have putative C-terminal GPI anchor sequences and there is a possibility that adding 6XHis tag at the C-terminus causes problems in protein folding and/or secretion as assayed by the western analysis of cell lysates and debris expressing hITLN1 and hITLN2. The 6XHis tagged proteins were mostly detected in detergent resistant

membrane debris. In contrast to hITLN1, insignificant expression of tagless hITLN2 was observed in lysates and debris.

To achieve cost effective high expression of hITLN1, bacterial expression systems were also attempted simultaneously. hITLN1 gene and hITLN2 gene, codon-optimized for expression in *E. coli*, were cloned in different vectors; pET-28a(+), pET-26b(+), pet-40b(+) and pET-32a(+) and expressed in various bacterial strains; BL21(DE3) cells, Rosetta(DE3) cells, Lemo21(DE3) cells, BL21(DE3) pLysS cells and Origami2(DE3) cells. The 6XHis tagged protein was detected only in insoluble fractions. The protein was tried to be solubilized under various conditions, however protein was solubilized only under two of these conditions – 1) Origami2(DE3) cells at low temperature and low IPTG conditions, and 2) BL21(DE3) cells in the presence of chaperones, pGKJE8 and pG-Tf2. The hITLN1-6XHis protein so expressed was not found to be functionally active as it did not bind to the Sepharose CL-6B beads. 6XHis tagged hITLN2 expression was only observed in Lemo21(DE3) cells, and the protein so expressed was not found to be functionally active as it did not bind to Sepharose CL-6B beads. Tagless hITLN2 with a periplasmic signal (pET-26b) was expressed in Lemo21(DE3) cells. Tagless hITLN2 was also expressed in Origami2(DE3) cells and in BL21(DE3) cells with chaperones, pGKJE8 and pGTf2, but these recombinant hITLN2 proteins did not show significant binding to Sepharose CL-6B beads.

Considering that hITLN1 is cysteine rich and presence of disulphide bonds could interfere in its proper folding, urea denaturation and refolding of hITLN1 was also done. A fraction of purified, refolded hITLN1 was found to retain lectin activity

as it bound to Sepharose CL-6B beads in a Ca^{2+} dependent manner. Under the same set of conditions hITLN2-6XHis got aggregated and could not be purified.

Since hITLN1 is a glycosylated, cysteine protein and bacterial expression system lacks post translational machinery for mammalian proteins, hITLN1 expression was also attempted in *P. pastoris* which is capable of post translational modifications like protein folding, disulphide bond formation and glycosylation. However, no protein expression was detected.

In summary, recombinant tagless hITLN1 could be successfully purified only from Flp-In-CHO cells transfected with hitln1-pEFBOS (without any tag).

In **Chapter 3**, purified tagless hITLN1 was used to study ligand binding. Out of the various ligands used in a Sepharose CL-6B bead binding inhibition assay, the best binding ligands obtained were 2-deoxy-D-galactose, 2-deoxy-D-ribose, 2-C-hydroxymethyl D-ribose, D-talose, D-idose, D-altrose and sorbitol. Of these, D-talose, D-idose, D-altrose 2-C-hydroxymethyl ribose and sorbitol are novel ligands which have not been reported in previous studies. We also found that hITLN1 bound to free N-acetylneuraminic acid, contrary to an earlier report by Tsuji *et al.* In addition, hITLN1 bound to β-D-Galactofuranose and 2-deoxy-D-ribose which have been earlier reported as ligands by Tsuji *et al*, too (Tsuji et al).

The D-ribose and D-galactose derivatives, 2-deoxy-D-ribose and 2-deoxy-D-galactose were found to bind much better than D-ribose and D-galactose. hITLN1 did not show significant binding to dissacharide sugars, similar to Tsuji's study (Tsuji et al., 2001).

Although Wesener *et al* findings had suggested that hITLN1 binds specifically to saccharides such as glycerol 1-phosphate only through the exocyclic 1,2-diol group, the findings of the study described in this thesis are that glycerol was only just as good as the other ligands, 2-deoxyribose, N-acetylneuraminic acid, sorbitol, ATP and CTP, in eluting hITLN1 from Sepharose CL-6B beads; all ligands including glycerol effectively eluted hITLN1 only at 0.1 M concentration.

By analyzing and comparing our results with the previous studies, it is concluded that there are several factors which may contribute to ligand binding by hITLN1. The presence of terminal exocyclic 1, 2-diol is not a sufficient criterion for ligand binding as also reported by Wesener *et al.* as some some sugars with terminal exocyclic 1,2-diol, D-allose and D-gulose, did not show efficient binding with hITLN1. Moreover, the presence of terminal exocyclic 1,2-diol does not seem to be strictly necessary for ligand binding by hITLN1, as some of the top glycan binders in our assay did not have this group - 2-deoxy-D-galactose, 2-C-hydroxymethyl ribose, D-ribose and 2-deoxy D-ribose. Rational site-directed mutagenesis of hITLN1 was done to mutate the residues involved in ligand binding as revealed by X-ray crystal structure; H268, W288 and Y297. The mutants did not bind to Sepharose CL-6B beads, confirming that these residues are critical for binding to ligands without the exocyclic 1,2-diol, too. The results of our binding assays suggest that ligands without exocyclic 1,2-diol are bound by hITLN1 through the same binding site but by a mode different from that observed in the solved crystal structure of the published hITLN1-ligand complex (Wesener et al., 2015).

An ELISA based assay demonstrated that hITLN1 binds to biotinylated human lactoferrin with an apparent binding affinity (K_d) of 11 nM in a Ca^{2+} dependent manner and binding affinity was found to be much higher as compared to what was shown in earlier studies 360 nM and 500 nM (Wesener et al., 2015) (Suzuki et al., 2001). hITLN1 bound to lactoferrin was partially eluted not by D-galactose but by its saccharide ligand, D-ribose and *vice versa*. The results suggest that there may be an overlap between the saccharide binding site and lactoferrin binding interface of hITLN1 and binding of a saccharide may modulate binding to lactoferrin and vice versa.

Chapter 4 describes the cloning, expression and purification of hITLN1 homologs – HaloITLN, XL35ITLN, OncoITLN and an uncharacterized bacterial protein with a FReD domain (psFReD). The lectins cloned in pET-28(+) were expressed in BL21(DE3) *E.coli* cells as tagless as well as 6XHis-tagged proteins but the proteins were detected mostly in the insoluble pellet, and could not be brought into the soluble fraction by various strategies. The proteins were finally solubilized by urea denaturation and refolded. OncoITLN-6XHis got precipitated upon addition of Ca^{2+}, therefore it was not used for further studies. A minor fraction of refolded psFreD-6XHis bound to Sepharose CL-6B beads indicating that the protein may not be properly folded in the refolding conditions used or that the binding to the Sepharose CL-6B matrix is weak. HaloITLN-6XHis and XL35ITLN-6XHis retained their lectin activity as assessed by binding to Sepharose CL-6B beads. Saccharide binding studies demonstrated that both HaloITLN-6XHis and XL35ITLN-6XHis bound well to melibiose, in contrast with hITLN1 which did not show any binding to disaccharides. HaloITLN-6XHis bound well to N-

acetylneuraminic acid and glycerol, like hITLN1. In contrast, XL35ITLN-6XHis bound very weakly to glycerol. XL35ITLN-6XHis bound well to N-acetylneuraminic acid and D-altrose and weakly to D-fucose, too. The saccharide binding profiles of XL35ITLN, HaloITLN and hITLN1 were thus very distinct. The reason for different saccharide specificities among these X-type lectins could be attributed to differences in the residues present in the binding pocket critical for ligand binding. In a multiple sequence alignment of hITLN1, XL35ITLN-6XHis and HaloITLN-6XHis sequences it was found that in HaloITLN-6XHis, the **Trp (W)** and **His (H)** residues of the binding pocket are conserved, but **Tyr (Y)** is replaced with **Val (V)**. Interestingly the **His (H)** and **Trp (W)** residues of a homology model of HaloITLN-6XHis and the hITLN1 three-dimensional structure were found to be completely overlapping. In XL35ITLN-6XHis, the binding pocket **His (H)** was found to be conserved while **Trp (W)** was found to be replaced with **Phe (F)** and **Tyr (Y)** was found to be replaced with **Asn (N).** These differences in the binding pocket may account for some of the different saccharide binding specificities among hITLN1, haloITLN-6XHis and XL35ITLN-6XHis.

HaloITLN-6XHis and XL35ITLN also showed binding to the lactoferrin with apparent binding affinities (K_d) of 35.2 nM and 45.9 nM, respectively. In the absence of lactoferrin in lower chordates, the binding displayed by XL35ITLN and HaloITLN is not expected to have a physiological role, but this confirms that binding is just facilitated by charge-charge interactions as previously proposed (Wesener et al., 2015).

Chapter 5 of the thesis is focused on a novel F-type lectin, Ranaspumin-4 (Rsn-4), from the Tungra frog, *Engystomops pustulosus*, which binds to D-

galactose, and not to L-fucose, which is the usual ligand of F-type lectins. This chapter describes the cloning, expression, purification and ligand binding by Rsn-4-6XHis and rationally designed mutants.

The F-type lectin domain (FLD) has a characteristic fucose-binding motif with a His/Arg/Arg triad that mediates hydrogen bonding with alpha-L-fucose. There are 437 FLD sequence clusters (clustered at 80% sequence identity), that are selectively distributed in eukaryotic, eubacterial and viral proteins. It has been revealed that approximately one-third of these sequences have a typical fucose-binding motif [HX(26)RXDX(4)R/K], but there are many sequences with variations in key amino acid residues and/or motif length, and about one-third of the non-redundant FLD sequences have substitutions of the critical His and/or Arg residues (Bishnoi et al., 2015).

Site directed mutagenesis was done to generate Rsn-4 mutants substituted at His67, Arg94 and Arg100 in order to know whether His and Ars residues are critical for ligand binding or not. While His67Asn seemed to retain weak binding, none of the other mutated proteins, even those with polar amino acids substituting His, displayed binding to D-galactose. Moreover the mutation of the first and second Arg to naturally substituting residues, Ile, Ser or Cys, also resulted in loss of lectin activity. The possible explanations for these results are 1) these mutants might have altered carbohydrate-binding specificities 2) in naturally occurring FLDs with atypical FLD motifs, a compensatory mutation counterbalances the deleterious effect of substitution of conserved residues, 3) only a subset of naturally occurring FLDs with atypical FLD motifs actually retain lectin activity; others might bind to non-carbohydrate ligands, or might have other, as yet unknown functions.

Also, by rational site directed mutagenesis, an Rsn-4 mutant (F97CM31Q) was successfully generated which is capable of binding to L-fucose and D-glucose in addition to its natural ligand D-galactose. This study provides a promising framework for engineering Rsn-4 or other F-type lectins to obtain designer lectins with altered binding specificity and binding affinity to desired ligands.