Synopsis

Lectins are a diverse group of proteins which can recognize different types of carbohydrates specifically and reversibly. Though first characterized in plants, lectins have now been identified in all other forms of life. A large number of lectins have been studied from animal and plant sources, however, studies on lectins from protozoan sources have been far and few. There are some reports of functional characterization of these proteins, where they were found to play crucial roles in adherence, interaction, and pathogenicity of some protozoans. However, the structural characterization is not yet available. The work carried out in this thesis attempts to fill the gap in our understanding of protozoan lectins and will shed light on the phylogeny of lectins found in higher eukaryotes.

Chapter 1 provides an overview of different folds of lectins in protozoa. It describes their overall structure, carbohydrate binding features and biological functions.

Chapter 2 gives details of various experimental techniques and procedures employed during the work carried out for this thesis.

Chapter 3 reports characterization of the first β-trefoil lectin (EntTref) from a protozoan source. EntTref from *Entamoeba histolytica* was cloned, expressed and purified. The protein was active as a lectin and could bind to rhamnose, galactose and galactose-linked sugars. The binding affinities were quantified by Isothermal calorimetry (ITC). Out of the three potential carbohydrate binding sites, EntTref exhibited carbohydrate binding in only one site. The protein was observed to undergo a cysteine-mediated dimerization. Such mode of dimerization has also been observed in other class of lectins. Further, we crystallized and determined the structure of the rhamnose-bound protein. Analysis of EntTref structure revealed the conserved nature of the carbohydrate binding site. Docking studies on EntTref could explain the increased binding affinity of some sugar molecules over others as observed during ITC studies.

Chapter 4 deals with the structural characterization of EntLec from *E. histolytica*. EntLec is a homolog of ERGIC-53 and belongs to the family of cargo lectins. Members of this class of proteins exhibit the legume lectin fold. EntLec was found to exist as a monomer in solution. Sequence comparison of the carbohydrate binding site revealed that the residues important for carbohydrate binding were not conserved. The crystal structure of this lectin-like protein has shed light on organism-specific differences in the structure of this class of proteins. Loops play an important role in carbohydrate recognition by legume lectins. In case of EntLec, the length of loops C and E were found to be shorter than that observed in plant legume lectins and ERGIC-53 making the carbohydrate binding site much wider. This shortening of loops together with non-conserved binding site could be a reason for the observed lack of carbohydrate binding in EntLec.

Chapter 5 describes the identification and analysis of putative lectins and lectin-like domains in protozoans, using sequence-based search approaches. A large-scale sequence analysis for all types of lectin domains was carried out through PSI-BLAST against the proteome of representative protozoans. The carbohydrate binding ability was evaluated by assessing the sequence conservation in the carbohydrate binding site. This search led us to interesting results: the occurrence of many plant and animal specific lectin domains in protozoans and the occurrence of these domains along with protein domains of unrelated functions pointing at the involvement of lectins in multiple processes within the cell.