**Thesis Outline**

**Chapter 1: Introduction**

1.1 Introduction of proteins and domains

1.2 Need for domain identification

1.3 Literature survey

1.3.1 Sequence based methods

1.3.2 Structure based methods

1.3.3 Manual & Semi-automatic methods

1.4 Organization of thesis

**Chapter 2: Materials & Methods**

2.1 Introduction

2.2 Clustering

2.2.1 Hierarchical clustering - Agglomerative and Divisive approaches

discuss here Newman's and Clauset's approach

2.2.2 non-Hierarchical clustering or partition methods - discuss here k-means

2.4 Domain identification by graph theoretic approach

2.4.1 Vishveshwara’s graph based approach

2.4.2 Hari Krishna’s method

2.5 Implementation details

2.7.1 Construction of data set

2.7.2 Libraries & tools

2.7.3. Distinguishing between single and multi-domain proteins

**Chapter 3: Results & Discussion**

3.1 Results of K-means clustering algorithm

3.2 Results of graph-based algorithm

3.3 Discussion of results and comparative analysis

**Chapter 4: Conclusion**