

UNIVERSITI TEKNOLOGI MARA

CLASSIFICATION OF ZINC HYDROLASES

USING

A STRUCTURE-BASED PHYLOGENETIC

METHOD

NURUL ATIKAH BT MOHD SANY

Written Thesis Submitted to
Universiti Teknologi MARA

Dissertation submitted in full fulfillment of the requirement for the degree
of
Bachelor of Pharmacy

Faculty of Pharmacy

JULY 2014

ACKNOWLEDGEMENT

Firstly, I would like to express my gratitude to the Almighty for giving me the opportunity and strength to finish up my thesis as a requirement for my Bachelor of Degree in Pharmacy. Other than that I would also like to express my gratitude to my supervisor, Yuslina bt Zakaria for helping me and giving me knowledge and guidance and help me with all her hearts, so that I can completed my thesis.

I would also like to thanks my friends and family that had given me support and had helped me to overcome my own weaknesses in order for me to complete my thesis. Special thanks to my sister, Nurul Fajariah bt Mohd Sany for providing me with needed tools for completion of my thesis.

Thank you.

TABLE OF CONTENT

	Page
ACKNOWLEDGEMENT	ii
TABLE OF CONTENTS	iii
LIST OF ABBREVIATIONS	V
LIST OF TABLES	vi
LIST OF FIGURES	vii
ABSTRACT	viii
CHAPTER 1 INTRODUCTION	
1.1 Introduction	1
1.2 Problem Statement	3
1.3 Objectives and aim	4
CHAPTER 2 LITERATURE REVIEW	
2.1 Protein	6
2.2 Zinc Hydrolases	9
2.3 Protein Classification	12
2.4 Phylogenetic Method	15
CHAPTER 3 METHODOLOGY	18
CHAPTER 4.0 RESULTS AND DISCUSSION	
4.1 Results	30
4.2 Comparisons with SCOP	33
4.3 A Comparisons with Previous Work by Wouters et al. (2001)	36
4.4 Comparisons with Sequence	38

ABSTRACT

Computational method of structural alignment is a method that had managed to grasp the attention of many people in the research world. Compared to the sequence alignment method, structural alignment method can bring people closer to the functions of proteins as the proteins structure are closely related to its function. SCOP databases had been used as a standard databases in order to obtain information about structural classification. However, problem rises as the SCOP database is developed manually and the computational method that is on par with this method has not been developed yet. The problem in this research is to determine whether the computational method is comparable to the SCOP classification. The objectives of this research is to assess the similarities of zinc hydrolases based on phylogenetic method, to construct phylogenetic tree by using similarity values from protein alignment and validate phylogenetic tree against standard protein classification. The result obtained show that structural alignment managed to produce tree with higher cluster order compared to the sequence alignment tree. When compared to Wouters et al. (2001), result obtained is almost similar and with increment of sample size, problems found in Wouters et al. managed to be solved in this research. Then it can be concluded that the structural alignment is a great method to be use in producing an accurate phylogenetic tree, conducting multiple structural alignment and identifying function of unknown proteins. However this method is not really accurate in classifying proteins with high sequence identity.

CHAPTER 1.0 INTRODUCTION

Chapter 1.0 introduces the aim and the purpose of the research. A slight exposure to the core and important part of the research is also be given in this chapter. This includes the importance of this research the contributions to the pharmaceutical research society.

1.1 Introduction

Protein is one of the most important elements in human body. It is derived from the human ribonucleic acid (RNA), which is synthesized in the human nucleus. RNA is synthesized based on the nucleotide arrangement of deoxyribonucleic acid (DNA) that holds important genetic and heritance data of all living organism on the earth (Gu & Bourne, 2011). So, indirectly, we can see that protein contains data and information, which are inherited from one generation to another generation of every organism. Though some data maybe lost during the transmission from one generation to another, some of it are still conserved which can be observed through the protein expression in the organism (Orengo & Thornton, 2005).

Protein can be classified into four groups according to the degree of folding of the protein. The groups are primary structure, secondary structure, tertiary structure and quaternary structure (Alberts et al., 2002; Gu & Bourne, 2011). For this research, the focus was on the tertiary structure of protein.