A rapid identification system for metallothionein proteins using expert system

Bhoopathi Praveen[1*], Savariar Vincent[3], Upadhyayula Suryanarayana Murty[1], Amirapu Radha Krishna[2] and Kaiser Jamil[1]

[1]Biology division, Indian Institute of Chemical Technology, Hyderabad-500 007, India;
[2]Computer division, Indian Institute of Chemical Technology, Hyderabad-500 007, India;
[3]Department of Zoology, Loyola College, Chennai - 600 034, India;
Bhoopathi Praveen* - Email: bhoopathip@gmail.com; * Corresponding Author

Abstract:

Metallothioneins (MT) are low molecular weight proteins mostly rich in cysteine residues with high metal content. Generally, MT proteins are responsible for regulating the intracellular supply of biologically essential metal ions and they protect cells from the deleterious effects of non-essential polarizable transition and post-transition metal ions. Due to their biological importance, proper characterization of MT is necessary. Here we describe a computer program (ID3 algorithm, a part of Artificial Intelligence) developed using available data for the rapid identification of MT. Tissue samples contains several low molecular weight proteins with different physical, chemical and biological characteristics. The described software solution proposes to categorize MT proteins without aromatic amino acids and high metal content. The proposed solution can be expanded to other types of proteins with specific known characteristics.

Keywords: metallothionein; rules; artificial intelligence; expert system; isolation; detection; tissue samples; purification

Back ground:

Metallothioneins (MT) constitute a super family of low molecular weight, cysteine-rich metalloproteins and metallopeptides responsible for regulating the intracellular supply of biologically essential zinc and copper ions. They have role in protecting cells from the deleterious effects of high concentration metal ions. [1] MT is now known to occur in all animal phyla examined so far as well as in certain fungi, plants and cyanobacteria. The function of MT is to detoxify non-essential metals such as mercury, cadmium and essential metals such as zinc and copper. MT is generally induced during stress conditions. [2] Hence, identification of MT from tissues/animal models is very important during analysis. Here, we describe a method for the identification of metallothionein proteins using expert system.

An expert system is a set of computer programs that mimic a human expert. [3] Expert system has been used for the identification of proteins using mass spectrometry in peptide mapping by ProFound. [4] It has also been used for detecting the protein domain structure in sequence [5] and for the prediction of protein localization sites in Gram-negative bacteria. [6] Classification and characterization of proteins are generally laborious dealing with large datasets and experimentations. [4] Here, we describe a framework using expert system developed from a set of physical and chemical properties of known proteins, especially MT. This procedure utilizes knowledge organized as "IF-THEN" rules for metallothionein protein. This methodology finds application in protein recovery, isolation and detection of proteins.

Methodology:

The physical and chemical characteristics of MT proteins were used in the development of a rapid identification system *in silico*. These characteristics were derived based on the experiments conducted using animal models (wistar rats). Different groups of animals were exposed to different concentrations of heavy metals like Cr(VI), Cd, Zn and Hg. Tissue samples containing MT were isolated from various organs after a series of centrifugation steps with alcohol extraction and purified by Ion exchange followed by Gel filtration chromatography. Purified MT was further subjected for the estimation of metal content, molecular weight and non-aromatic amino acid content using atomic absorption spectro-photometry, SDS-PAGE and Bradford assay, respectively. The derived characteristics from experimental results were divided into a set of rules. These rules include, (1) protein with low molecular weight / high molecular weight, (2) proteins with metal content / without metal content, (3) aromatic amino acids / without aromatic amino acids and (4) sulphur content / without sulphur content. The derived rules at every step were trained in an expert system (ID3 algorithm) as described elsewhere. [7-9] A minimum number of rules were selected (each rule is a series of conditions consisting of attributes and value pairs, followed by a single conclusion that contains the class and the corresponding class value) using human expertise to maximize true positive identification. The rules were then formulated into an IF – THEN – ELSE algorithm and translated into VISUAL BASIC language statements. Thus a software solution is proposed for the identification of MT during protein recovery. The methodology for the identification of MT is illustrated in Figure 1.

14
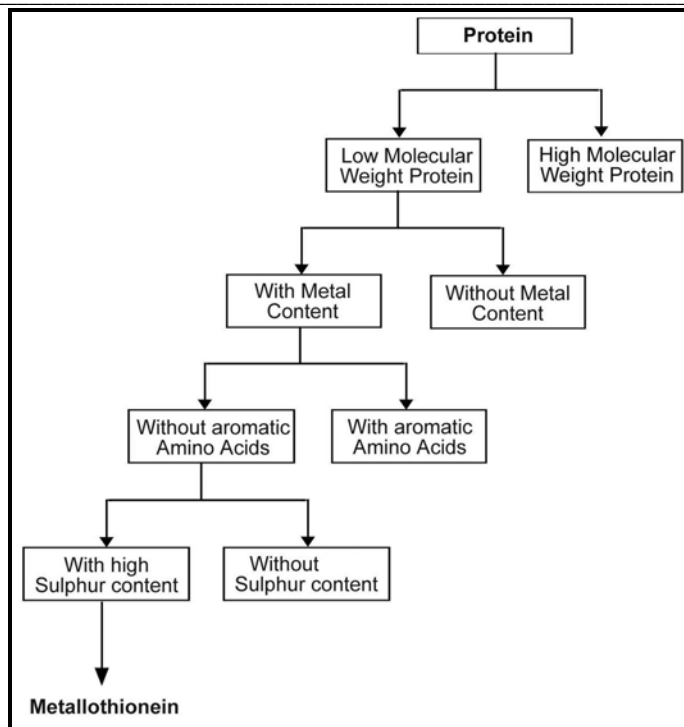
Figure 1: Identification of MT proteins. The rules are given in boxes

Limitation:

Expert systems are useful tools for the identification of MT during protein characterization. Here, we propose the utilization of expert system for the identification of MT from a mixture of unknown proteins during a series of purification steps. A pilot development shows that the procedure is successful in identifying proteins with high specificity towards low or very low molecular weight components. The major limitation in the use of expert system is the need to train specific systems using specific sets of rules. This requires domain expert regarding the protein of interest. In other words, this procedure cannot help in the identification of unknown proteins other than metallothionein, unless the methodology is modified for other proteins.

References:

[1]   F. Silvestre, *et al*., *Comp Biochem Physiol C Toxicol Pharmacol*., 140:39 (2005) [PMID: 15792621]
[2]   P. Dziegiel, *et al., Rocz Akad Med Bialymst.,* 49:43 (2004) [PMID: 15638370]
[3]   D. L. Rotin, *et al., Arkh Patol*., 66:47 (2004) [PMID: 15154385]
[4]   W. Zhang & B.T. Chait, *Anal Chem.,* 72:2482 (2000) [PMID: 10857624]
[5]   E. Benatan, *Proc Int Conf Intell Syst Mol Biol.,* 2:37 (1994) [PMID: 7584415]
[6]   K. Nakai & M. Kanehisa*. Proteins,* 11:95 (1991) [PMID: 1946347]
[7]   U. S. N. Murty, *et al., Comput Appl Biosci.,* 12:491 (1996) [PMID: 9021267]
[8]   M. Edwards, *et al., Comput Appl Biosci.,* 3:1 (1987) [PMID: 3330955]
[9]   R. Hofestadt, *Medinfo.,* 2:964 (1995) [PMID: 8591601]