Identification of Immunogenic Peptides of Metadherin Protein for Cancer Vaccine

Dissertation

Submitted in the partial fulfillment of the requirement for

the award of the degree of

MASTER OF TECHNOLOGY

IN

BIOTECHNOLOGY

By

GouravDhiman

601204007

Under the guidance of

Dr. ManojBaranwal

Assistant Professor

DEPARTMENT OF BIOTECHNOLOGY

THAPAR UNIVERSITY, PATIALA (PUNJAB)-147004

JULY 2014

CANDIDATE'S DECLARATION

I, hereby declare that the work presented in the thesis entitled **"Identification of Immunogenic Peptides of Metadherin Protein for Cancer Vaccine"** in the partial fulfillment of the requirement for the award of the degree of Master of Technology in Biotechnology, Department of Biotechnology, Thapar University, Patiala, is an authentic record of my work during the period of one year from July 2013 to June 2014, under the guidance of Dr. ManojBaranwal, Assistant Professor, Thapar University, Patiala. I have not submitted the matter embodied in this thesis for the award of any other degree or diploma.

Place: Patiala GouravDhiman

Date: 18/07/14 (601204007)

CERTIFICATE

This is to certify that the thesis entitled **"Identification of Immunogenic Peptides of Metadherin Protein for Cancer Vaccine"** submitted by Gourav Dhiman in partial fulfillment of the requirement for the award of Degree of Master of Technology in Biotechnolgy to Thapar University, Patiala, is a record of student's own work carried out by him. The report has not been submitted for the award of any other degree or certificate in this or any other University or Institute.

Dr. Manoj Baranwal
Assistant Professor, Supervisor
Department of Biotechnology, TU
Patiala

Dr. Dinesh Goyal
Professor & Head
Department of Biotechnology, TU
Patiala

Dr. S.K. Mohapatra
Dean
Academic Affairs,
Thapar University
Patiala

ACKNOWLEDGEMENTS

I bow before the almighty for blessing me with his guidance. Thanks to my parents and family for giving me freedom and opportunity to pursue my own interest, even when they appeared incomprehensible.In the completion of the present work, there are feelings of achievement and satisfaction. In moment of happiness I take opportunity to record my sincere gratitude to all those who mattered.

First and foremost, I would like to express my profound gratitude and great indebtedness to my supervisor, Assistant Professor **Dr. Manoj Baranwal**, for his benevolent supervision of my M. Tech Thesis. I am very grateful to him for giving me the opportunity to conduct this research project within his fascinating research group, for his continuous support and constructive criticism, and for encouraging and challenging me throughout the progress of the study; never accepting less than my best efforts.

I express my sincere thanks to **Dr. Dinesh Goyal**, Professor and Head, Department of Biotechnology (DBT), Thapar University, Patiala, or providing the best laboratory facilities.

Special thanks to**Miss. Neha Lohia** research scholars in DBT, for their immense help and valuable suggestions. I also wish to express my sincere thanks to my cousin **Mr. Suraj Kumar and** friend **Satvinder Singh** for their constant moral support at the time of trouble. I was fortunate enough to avail the ungrudging help and assistance from all our laboratory staff, **Mr**. **Ram Newal,** Mr. **Lallan**and Mrs. **Lalita**.

Last but not the least I am thankful to my wonderful friends for their encouragement. I thank my friends **Vivek Sharma, Sarbjot Singh and Jhilmil Tandon** and all my classmates for help whenever required. I am forever indebted to my **parents** for their love and support.

GouravDhiman

(601204007)

CONTENTS

A

ABSTRACT

Metadherin overexpression and its role in development of diverse cancer makes it an attractive target for cancer vaccine. In the current study, combination of epitope prediction algorithm and molecular docking were employed to find the peptide containing T and B cell epitopes in this protein. Three peptide containing overlapping CD4+ and CD8+ T cell epitope were obtained using six different immunoinformatics prediction programs. Population coverage analysis of these peptide fragments have indicated that they have the capacity to induce a potent immune response among the individuals belonging to different ethnicities around the world. Molecular docking study has shown that four nonamers common for both CD4+ and CD8+ T cell epitopes has comparative binding energy with naturally bound peptides. Interestingly, B cell epitope prediction study resulted into two peptide fragments containing both T and B cell epitope which shows that these peptides have the capacity to induce cellular and humoral immune response.One of the predicted peptide M3 found to stimulate the PBMC (Peripheral blood mononuclear cell) proliferation during initial screening in the *in vitro* system. It is the first report on metadherin protein for immunogenic peptide identification based on immunoinformatics approach. Hence these peptides may be considered as vital vaccine candidate for multiple cancer.

B

LIST OF ABBREVIATIONS

ABC	ATP-binding cassette
AEG-1	Astrocyte elevated gene-1
APC	Antigen Presenting Cell
FDA	Food and Drug Administration
HER-2	Human Epidermal Growth Factor Receptor
HMMs	Hidden Markov Models
hTERT	Human telomerase reverse transcriptase
LEF-1	Lymphoid Enhancer Binding Factor-1
LYRIC	Lysine-rich CEACAM-1–associated protein
MHC	Major Histocompatibility Complex
MRP	Multidrug resistance-associated protein
MTDH	Metadherin
NLS	Nuclear Localizing Signal
QSAR	Quantitative Structure Activity Relationship
SART	Squamous cell carcinoma antigen recognized by T cells
SVMs	Support Vector Machines
TAAs	Tumor associated antigens
mg	Milligram
KDa	Kilo Daltons
ml	Millilitre
µg	Microgram
^{o}C	Degree celcius
µl	Microlitre

C

LIST OF TABLES

D

LIST OF FIGURES

Chapter 1. INTRODUCTION

Cancer remains a serious threat to human life despite significant advancement done in its treatment. Different cancers accounted for 8.2 million deaths in the year 2012 and it is projected to surge in the count of cancer patients to 22 million from present 14 million in next two decades (http://www.who.int/mediacentre/factsheets/fs297/en/). Cancer researchers are trying to understand the relationship between cancer and immune system with intension of finding novel immunotherapeutic approaches to control the cancer. There are several existing cancer immunotherapy approaches which include anti-tumor monoclonal antibodies, adoptive T cell therapy, cytokines based therapy and cancer vaccines (Raval et. al., 2014). Monoclonal antibodies (mAb) directed against tumor associated antigens like CD20, HER-2, CD30 are used in different cancer treatment (Raval et. al., 2014).

Vaccine against cancer is a promising approach as it involves the patient's own immune system to control the cancer. Gardasil, a human papilloma virus (HPV) vaccine is first FDA approved vaccine which is used to protect from cervical cancer (Emens et.al. 2008). Hepatitis B virus vaccine is another vaccine approved for liver cancer prevention. These are preventive vaccine while therapeutic vaccine is used to protect cancer after diagnosis. Sipuleucel-T (Provenge^{TM}) a dendritic cell vaccine is the only FDA approved therapeutic vaccine which is used against the prostate cancer , where dendritic cells from patients are taken and injected back after labelling with the prostate acidic protein which induce the T-cell response in the patient against cancer (Guo *et al* 2013). A diverse categories of vaccines such as tumor cell vaccine, protein/peptide vaccines, and genetic (DNA, RNA and viral) vaccines are under the stages of development or in clinical trials (Guo *et al* 2013). Peptide based vaccine is one of the promising approach for cancer therapeutics. Several categories of tumor antigens have been described and a peptide database available at cancer immunity web site has reported 403 well defined tumor antigenic peptides till 2013 (Vigneron *et al*., 2013). These tumor antigens are of particular interest for cancer vaccine because they are normally less expressed or absent in normal tissues and have potential to induce immune response. Peptides derived from tumor associated antigens (TAAs) like squamous cell carcinoma antigen recognized by T cells (SART) 2, SART3, multidrug resistance-associated protein (MRP) 3, alpha-

fetoprotein (AFP) and human telomerase reverse transcriptase (hTERT) have shown peptide specific T cell immune response in hepatocellular carcinoma which confirms the immunogenic potential of these TAAs (Mizukoshi *et al*., 2011).

Metadherin (MTDH) protein also known as Lysine-rich CEACAM-1–associated protein (LYRIC)/astrocyte elevated gene-1 (AEG-1) has emerged as an important protein which is widely expressed in the different types of cancer. It has been reported as a type II trans membrane protein, which possesses an extracellular lung homing domain held responsible for breast to lung cancer metastasis (Brown *et al*, 2004) signifying its name (metastasis and adhesion). The increased expression of MTDH is observed in breast cancer, oesophageal squamous cell carcinoma, prostate cancer and hepatocellular carcinoma (Hu *et al*., 2009) while in the normal tissues like epithelial cells of mammary gland ducts, purkinjee neurons in early postnatal stages and adult cerebellum expresses its expression is at null or a very low level (Brown *et al*, 2004). These expression studies suggest that MTDH may acts as potential TAAs. Inhibition of MTDH in neuroblastoma cells was reported to increase the chemo sensitivity of anticancer drug cisplatin or doxorubicin (Liu *et al*., 2009) and radiosenstivity in cervical cancer cell line (Zhao *et al*., 2012). Antibodies against this protein were also reported in different cancer patients which confirms its immunogenic potential (Chen et.al. 2012). With these studies, MTDH appears to be a promising target to identify peptide enriched epitope as probable candidates for vaccine design. An immunoinformatics-driven T and B cell epitope mapping and prediction algorithms appear to be leading the charge in identifying novel epitopes which save a considerable time and efforts required to screen epitopes. Recently, it has reported conserved peptides containing T cell epitopes in the H1N1 influenza virus based on immunoinformatics approach (Lohia et.al. 2014).

Considering the critical role of MTDH protein in different cancers, present study is focussed on the identification of immunogenic peptide fragment by applying *in-silico* approach. A combination of different computer based prediction algorithms were used to find T and B cell epitopes for the identification probable peptide candidates as cancer vaccine target. Immunogenic potential of these predicted epitopes were further evaluated by performing molecular docking with different MHC class I and II molecules. It is the first of its kind of study on metadherin protein for immunogenic peptide identification based on immmunoinformatics and molecular docking approach.

Chapter 2. REVIEW OF LITERATURE

The body is made up of trillions of living cells. Normal body cells grow, divide into new cells, and die in an orderly fashion. During the early years of a person's life, normal cells divide faster to allow the person to grow. After the person becomes an adult, most cells divide only to replace worn-out or dying cells or to repair injuries.

Cancer begins when cells in a part of the body start to grow out of control. There are many kinds of cancer, but they all start because of out-of-control growth of abnormal cells.

Cells become cancer cells because of damage to DNA. DNA is in every cell and directs all its actions. In a normal cell, when DNA gets damaged the cell either repairs the damage or the cell dies. In cancer cells, the damaged DNA is not repaired, but the cell does not die like it should. Instead, this cell goes on making new cells that the body does not need. These new cells will all have the same damaged DNA as the first cell does.

Cancer cells often travel to other parts of the body, where they begin to grow and form new tumors that replace normal tissue. This process is called metastasis. It happens when the cancer cells get into the bloodstream or lymph vessels of our body. No matter where a cancer may spread, it is always named for the place where it started like breast cancer that has spread to the liver but still called breast cancer, not liver cancer. Likewise, prostate cancer that has spread to the bone is metastatic prostate cancer, not bone cancer.

2.1 Vaccines

Louis Pasteur developed the concept of vaccine through his innovative work in microbiology and coined the term vaccine to all the protective preparations to honour the Edward Jenner. Now, vaccination is the administration of antigenic agents applied to stimulate the immune system of an individual and to develop adaptive immunity to a disease. Vaccines can improve, or often even prevent, the effects of infection. Vaccination is generally considered to be the most effective method of preventing infectious diseases, and the efficacy of vaccination has been extensively studied and verified. The administration of some vaccines is conducted after the patient has already been infected by the pathogen. The basis behind all the vaccinations is the ability of the vaccine to initiate an immune response in a quicker fashion than the pathogen itself. The purpose of every vaccination is to present a particular antigen or set of antigens to the immune system in order to evoke a relevant immune response. The main active component of a vaccine may be inactive, but still intact (attenuated

bacteria or viruses), or purified components of the pathogen that are known to induce immune reaction. There are different types of vaccines such as live attenuated vaccines, inactivated vaccines, subunit vaccines, peptide vaccines, DNA vaccines, therapeutic vaccines etc. Therapeutic vaccines development against chronic diseases has picked pace now days. It differs from the conventional vaccine as administrated to the person suffering from the disease. Significant efforts are being made to develop therapeutic vaccines against cancer, AIDS, hepatitis and tuberculosis. Copolymer 1 is used as therapeutic vaccine which show similarity with myelin basic protein is a great success in the field of therapeutic vaccines (Sela *et al*., 2002)

2.2 Peptide vaccines

The improved knowledge of antigen recognition at molecular level has contributed to the development of rationally designed peptide vaccines. The general concept behind the peptide vaccines is based on the chemical approach to synthesize the identified B-cell and T-cell epitopes that are immunodominant and can induce specific immune responses. B-cell epitope of a target molecule can be conjugated with a T-cell epitope to make it immunogenic. The first epitope-based vaccine was created in 1985. They introduced recombinant DNA and expressed epitopes against cholera in Escherichia coli. Thus, epitope-based vaccines can be constructed for T and B lymphocytes. The T-cell epitopes are typically peptide fragments, whereas the B-cell epitopes can be proteins, lipids, nucleic acids or carbohydrates. Peptides have become desirable vaccine candidates due to their comparatively easy production and construction, chemical stability, and absence of infectious potential. The peptide vaccines against various cancers have been developed, and entered phase I and phase II of clinical trials, with satisfactory clinical outcome as listed in Table 1. The peptide vaccination is commonly being studied for application in both ameliorating and prophylactic immunotherapy (Naz and Dabir, 2007). Yet there is more to be improved in order to eliminate obstacles, such as the need for a better adjuvant and carrier or the low immunogenicity.

Table 1: Peptide based vaccination trials in cancer patients

Tumor	Peptide vaccine	Clinical Trial phase	Reference
Melanoma	gp100(g209)-2M	II	Rosenberg *et al*., 1998
	MART-127–35	I	Cormier *et al.*, 1997
	gp100	I/II	Rosenberg *et al*., 1999
	MART-127–35	I/II	Wang *et al*., 1999
	gp100 (210M) + tyrosinase	II	Lee *et al*., 2001
	MAGE-3.A1	I/II	Marchand *et al*., 1999
Melanoma and others	Tyrosinase	II	Scheibenbogen *et al*., 2000
	NY-ESO-1	I/II	Jager *et al*.,2000
Pancreatic cancer	K-Ras/12	I/II	Gjertsen *et al.*, 2001
CIN	HPV-16/E7 + KSS/PADRE	I/II	Muderspach *et al*., 2000

MART-1 = melanoma antigen recognized by T cells; MAGE-3.A1 = melanoma antigen-3 peptide restricted by HLA-A1; NY-ESO-1 = New York- esophagus antigen-1; K-Ras/12 = K-Ras mutated at position 12; CIN = cervical intraepithelial neoplasia; HPV = human papilloma virus; E7 = early protein 7; KSS = amino acid sequence of the linker peptide; PADRE= pan-DR Epitope

2.3 Immunoinformatics

The accelerating growth of bioinformatics techniques and applications along with the substantial amount of experimental data has made a significant impact on the immunology research. This has led to a rapid growth in the field of computational immunology, and a number of immunology-focused resources and software, which help in understanding the

properties of the whole immune system, have become available. This has given rise to a new field, called immunoinformatics. This field can be described as a branch of bioinformatics concerned with in silico analysis and modelling of immunological data and problems. Immunoinformatics research stresses mostly on the design and study of algorithms for mapping potential B- and T-cell epitopes, which speeds up the time and lowers the cost needed for laboratory analysis of pathogen gene products. Using such tools and information, an immunologist can analyse the sequence areas with potential binding sites, which in turn leads to the development of new vaccines. The methodology of analysing the pathogen proteome to identify the potential immunogenic protein is known as reverse vaccinology. Conventional methods need to dedicate time to pathogen cultivation and subsequent protein extraction. Although pathogens grow quickly, extraction of their proteins and then testing of those proteins on a large scale is expensive and time-consuming. Immunoinformatics is capable of reducing time and saving resources for the development of relevant vaccines by directing towards the development of more rational peptide based vaccine approach. Normally, the investigation of the binding affinity of antigenic peptides to the MHC molecules is the main goal when predicting epitopes. The experimental techniques are found to be difficult and time-consuming, and therefore several *in silico* methodologies are being created and constantly improved to identify epitopes. The list of approaches includes matrix-driven methods, QSAR analysis, identification of structural binding motifs, protein threading, homology modelling, docking techniques, and design of several machine-learning algorithms and tools (Tomar and RK, 2010). In the past, computational techniques could only identify sequence characteristics, but new improved algorithms and tools are being designed to increase the predictive performance. The methods used for development of prediction models can be divided into structure-based methods that derive information from the three-dimensional structure of the proteins, and sequence-based methods that analyse the amino acid sequence. Cancer vaccine development have using immunoinformatics has become important area, HER-2/neu (Gritzapis *et al*., 2009) and pokemon (Yuan *et al*., 2012) specific immunogenic peptides has been identified by this approach.

2.4 T-cell and B-cell epitopes and their prediction algorithms

The epitope is recognizable by the immune system part of the antigen, and in particular by antibodies, B cells or T cells. The epitopes may belong to both foreign and self-proteins, and they can be categorized as conformational or linear, depending on their structure and integration with the paratope. T-cell epitopes are presented on the surface of an antigen presenting cell (APC), where they are bound to major histocompatibility (MHC) molecules in

order to induce immune response. MHC class II proteins bind oligopeptide fragments derived through the proteolysis of pathogen antigens, and present them at the cell surface for recognition by CD4+ T cells. If sufficient quantities of the epitope are presented, the T cell may trigger an adaptive immune response specific for the pathogen. Class II MHCs are expressed on specialized cell types, including professional APCs such as B cells, macrophages and dendritic cells, whereas class I MHCs are found on every nucleated cell of the body. The MHC I molecule binds to a peptide of approximate 9 amino acids in length within a closed groove. In contrast, because the antigen-binding groove is open at both ends, the MHC II molecules can present much longer peptides, generally varying from 12 to 25 amino acids, nine of which occupy the binding groove. This difference between MHC I and MHC II is very important for the development of distinct prediction algorithms (Larson *et al*., 2006). The recognition of epitopes by T cells and the induction of immune response is the sole basis of cell mediated immune response.

One of the key issues in T-cell epitope prediction is the prediction of MHC binding, as it is considered a prerequisite for T cell recognition. All T-cell epitopes are good MHC binders, but not all good MHC binders are T-cell epitopes. Determining the peptide binding preferences exhibited by this extensive set of alleles is beyond the present capacity of experimental techniques, necessitating the development of bioinformatics prediction methodologies. The most successful prediction methods for T-cell epitopes developed to date have been data-driven. T-cell epitope prediction typically involves defining the peptide binding specificity of specific class I or class II MHC alleles and then predicting epitopes *in silico*. Using peptide sequence data, experimentally determined affinity data have been used in the construction of many T-cell epitope prediction algorithms. Such methods include motif-based systems, support vector machines (SVMs), and hidden Markov models (HMMs), quantitative structure–activity relationship (QSAR) analysis and structure-based approaches.

Compared to T cell epitopes prediction algorithms, the B cell epitope prediction is more complicated, especially for the conformational B cell epitopes because, in addition to the sequence composition, the 3D-structure of protein must also be considered. The development of B cell epitopes prediction algorithms has been less successful compared to T cell epitope prediction, especially in accuracy (Anderson *et al*., 2006). There are several reasons for this. For instance, the majority of B cell epitopes are discontinuous so that it is hard to determine the relevant amino acids and the distribution of the antigen surface. Moreover, much of the experimental data based on the prediction algorithms are still controversial because of the poorly understood recognition properties of cross reactive antibodies (Davies *et al*., 2007). 16

Nevertheless, in spite of these difficulties, there are several methods available for B cell epitope prediction for both linear and conformational epitopes. The prediction algorithms for linear B are similar to that of T-cell. The accuracy of primary sequence based algorithm is low and modified algorithms based on machine learning were subsequently developed, such as ABCpred (Saha et. al., 2006) and BepiPred (Larsen *et al*., 2006) with significant improvements in accuracy. Prediction algorithms for conformational B cell epitopes based on 3D structure are also available owing to the ever-increasing 3D structure of antigen-antibody complex data. Some prediction servers based on this algorithm are accessible, for example DiscoTope and CEP (http:// bioinfo.ernet.in/cep.htm) (Kulkarni *et al*., 2005). These methods make use of information carried in the structure of antibodies against proteins of interest to reveal the 3D folding of target proteins.

2.5 Metadherin and its Potential for vaccine development

Metadherin (MTDH) (Brown *et al*, 2004) also known as LYRIC (Lysine Rich CEACAM1) (Britt *et al*, 2004) and Astrocyte Elevated Gene-1(MTDH) protein (Kang *et al*, 2005) is emerging as an important protein with multiple role in cancer progression and metastasis. It was discovered by three different groups independently and named accordingly. In a study it was reported as a type two trans membrane protein, possesses the extracellular lung homing domain which is held responsible for breast to lung cancer metastasis (Brown *et al*, 2004) thus named as Metadherin which means Metastasis and Adhesion. It is highly expressed in the tumors but the normal tissues like epithelial cells of mammary gland ducts, Purkinjee neurons in early postnatal stages and adult cerebellum expresses it at null or a very low level (Brown *et al*, 2004). In another independent study Metadherin was co isolated with the CEACAM1 protein named as LYRIC and found to be highly conserved in different species. In polarized epithelial cells it coexist with tight junction protein ZO-1 and occludin and is not the native component of tight junction but otherwise believed to be incorporated in it during tight junction complex maturation (Britt *et al*, 2004). It has identified as HIV-1 and TNF-α inducible gene encodes single pass trans membrane protein of molecular mass of 64 α KDa and mainly located in Endoplasmic Reticulum(ER) and perinuclear Space (Kang *et al*, 2005).

2.6 Molecular Basis and Functions

MTDH is having three putative NLS (Nuclear Localizing Signal) from amino acid 79 to 91,432 to 451 and 561 to 580 (Sarkar *et al*, 2008). But the location of MTDH protein varies depending upon cell's physiological state. In benign prostate hyperplasia or in other benign tissues the MTDH is found to be located in nucleus, whereas in tumorous cell an abundance

of MTDH is seen in cytoplasm, which affects the patient's survival rate (Thirketle *et al*, 2009). It may be concluded that as the cells becomes tumorous the MTDH protein starts localizing in cytoplasm and must play some role in cancer progression and metastasis.

The N-terminal domains of MTDH are found to be involved in interacting with the NF-κB and lead to the activation of NF-κB mediated gene expression (Sarkar *et al*, 2008). With the increase in MTDH expression the level of p65 subunit of NF-κB increases in the cytoplasm and vice- versa, at the same time there is a significant reduction in the NF-κB inhibitor IκBα which binds to the NF-κB and keeps inactivated. Thus MTDH may be involved in the degradation of INF-κB and hence increasing the NF-κB mediated gene expression. On treatment with the TNF-α, level of MTDH and p65 of NF-κB subunit increases in the nucleus indicates their collective interaction for enhancement of the transcriptional activity of IL-8 (Emdad *et al*, 2006) (Sarkar *et al*, 2008). Studies has revealed that MTDH interacts with cAMP responsive element binding (CREB) protein (Sarkar *et al*, 2008), which is a universal co-activator for NF-κB and result in the IL-8 gene expression. Hence on activation by TNF-α, MTDH may act as a link between p65, NF-κB and CREB and result in activation of NF-κB related gene expression (Sarkar *et al*, 2008). There are 15 Astrocyte Elevated genes identified till date, and of which MTDH is found of great importance in cancer related studies. An increase expression of MTDH interact with the Ha-Ras oncogene leads to the tumor development and progression in melanocytes (Kang *et al*, 2005). At the same time hyper expression of MTDH results in NF-κB oriented gene expression which in turn leads to anchorage independent growth in cell lines (Emdad *et al*, 2006). The reason for which may be phosphorylation of p65 subunit of NF-κB at serine 536 and regulate the expression of MMP1 (Matrix metalo Peptidase 1), as the regulatory domain of NF-κB binds to the promoter of MMP1 (Wang *et al*., 2013). The HIV-1 or gp120 infection results in an increased expression of MTDH and decreased excitatory amino acid transporter 2 (EAAT2) gene expression by hindering its promoter activity. Thus TNF-α, HIV-1 or gp120 may directly result in the down regulating the EAAT-2 promoter activity by passing the PI-3k-Akt signalling pathway (Kang *et al*, 2005). TNF-α results in down regulation of EAAT-2 promoter is via activation of NF-κB, which binds to its promoter and inhibit its expression hence it may be hypothesised that MTDH hyper expression may also be involved in EAAT-2 inhibition by activating NF-κB as a result of TNF-α activation.

Tumor microarray analysis of hepatocellular carcinoma showed that there is an abundance of fusiform shaped tumorous cells at the margin tumorous tissue. As the fusiform shape is

characteristic of mesencymal cells, it was found that higher expression of MTDH results in lowering the level of E-cadherin and enhancing the N-cadherin and snail expression, signifying the Epithelial to mesenchymal transition. The reduction of E-cadherin results in localizing the β-catenin to the nucleus and thus functions in Wnt / β-catenin signalling pathway hence enhancing the cell motility. The MTDH hyper expression in hepatocellular carcinoma results in the up regulation of multiple genes and signalling pathways like activation of ERK 42/44 and p38 Mitogen Activated Protein Kinase (MAPK) signalling pathway along with it modulates the another cluster of gene belonging to the Wnt signalling pathway (Yoo *et al*., 2009). .

2.7 Expression of MTDH in different types of cancers

The expression of MTDH has been found in varied different cancers types where it is prominently associated with cancer metastasis. Several studies have showed its expression in different cancer types and responsible for the cancer progression which makes it a ubiquitously expressing oncogene in different cancers.

2.7.1 Breast cancer

Initially metadherin was identified by *in vivo* phage display screening, as a protein which was responsible for the breast to lung cancer metastasis possessing extracellular domain called lung homing domain. Immunostaining of 20 normal mammary gland tissue showed that 18 samples were devoid of metadherin expression which signifies that metadherin is selectively overexpressed in tumors (Brown *et al*). Another immunohistochemical study has shown that there is low or no expression of the metadherin in normal breast tissue, and it is highly expressed in ductal carcinoma in situ (DCIS). But the expression was found relatively low in invasive ductal carcinoma, which leads to a conclusion that metadherin may be responsible for the initiation of it (Su *et al*., 2010). Thus it is evident from the above instances that MTDH is expressed in cancerous tissue.

2.7.2 Prostate Cancer

In a study of 20 prostate cancer tissue samples 80 percent of the samples were found to be expressing the higher level of MTDH. When it was compared with normal prostate cell line RWPE-1 had an almost three time lower AEG -1 expression (Kikuno *et al*., 2007). When MTDH was knocked down in the cancerous cells a reduction in cell viability and enhanced

apoptosis was observed. It signifies that MTDH plays an important role in the prostate cancer progression (Kikuno *et al*., 2007).

2.7.3 Colorectal and Colon Cancer

The expression of MTDH was found enhanced in the sample of CRC (Colorectal Cancer) patient and 156 primary tumours whereas when compared with the normal mucosal cells the expression was significantly lower. When compared with the primary tumor cells to the lymph node metastasized cells an enhanced expression of MTDH was observed. It is evident that MTDH expression is enhanced in the colorectal cancer and it increases as the cell undergoes metastasis. Similarly in colon cancer in the colon cancer cell lines (SW480) and its lymph node metastasis cell line (SW620) and primary colon cancer (KM12)cell lines the MTDH expression was observed and in the metastatic cell lines a higher expression of the MTDH was observed (Gnosa *et al*., 2012). It further justifies that MTDH is highly expressed in colon and colorectal cancers.

2.7.4 Head and Neck cancer

Another study on 94 samples of head and neck cancer at different stages and 20 samples of head and neck squamous cell carcinoma has showed, that MTDH expression is elevated in these cancers. A siRNA knockdown of the MTDH resulted in reduced cell invasion and migration, and reduced AKT/PI3K signalling. It is clear that like other cancers the MTDH expression is important in progression of head and neck cancer. Similarly it was found that the combined effect of siRNA-375 and knockdown of MTDH is very prominent in the inhibition of the cancer progression (Hui *et al*., 2011).

2.7.5 Tongue and salivary Gland carcinoma

In tongue carcinoma an elevated expression of the MTDH expression was found to be enhanced in a study of 93 paraffin embedded tongue cancer tissue samples and a lower expression was found in normal tissue. The tissue sample adjacent to the cancerous tissue also showed the lower expression of the MTDH in comparison to the tumorous tissue (Ke *et al*., 2013). The hyper expression of MTDH is also involved in the progression of the salivary gland carcinoma. It was observed that patient with low MTDH expression had higher survival rate which signifies that expression of MTDH is important in cancer proliferation and metastasis. The normal tissue of the salivary gland and the tissue adjacent to the tumorous

tissue showed lower expression, which illustrate the expression is associated with different cancer stages (Liao., 2011).

2.7.6 Hepatocellular carcinoma

The role of MTDH is also prominent in this cancer type, the hyper expression of MTDH is responsible for cancer metastasis and progression through different pathways by modulating or inhibiting the various gene associated with the apoptosis, metastasis, invasion and senescence. The invasive and metastatic properties has been correlated with Epithelial to mesenchymal transition of the hepatocellular carcinoma as result of high MTDH expression (Zhu *et al*., 2011).

2.8 Role in chemo resistance and Radio resistance

MTDH has also found responsible for chemo resistance. The suppression of MTDH by DNA vaccine / siRNA lead to increased sensitivity for doxorubicin in breast cancer (Qian *et al*., 2011) and neuroblastoma cells (Liu *et al*., 2009). Similarly the hyper expression of the metadherin has found to be associated with the radio resistance in cervical cancer (Zhao *et al*., 2012). Hence the multiple role of the metadherin in the different cancer development and progression is making it a suitable candidate for the cancer vaccine development. Thus such vaccine can target the multiple cancers and at the same time can help in tackling the metastasis.

Chapter 3: OBJECTIVES

- Identification of peptide containing T and B cell epitopes of Metadherin using different immunoinformatics tools.
- Molecular docking of predicted peptides with the MHC class I and II structures and analyse their binding ability to the antigen binding pocket of MHC.
- *In vitro* analysis of potential of the predicted peptides to stimulate T-cell proliferation in PBMC culture.

Chapter 4. MATERIALS AND METHODS

4.1 Sequence

The sequence of metadherin (MTDH) was retrieved from Genbank bearing accession number Q86UE4.2. It is a 582 amino acid long lysine rich protein (Deborah E.Britt *et al*, 2004).

4.2 $CD8^+$ T-cell Epitope prediction

CD8+ cells recognise the peptides presented on class I MHC by the cells which are either tumorous in nature or viral infected. The target cells are killed by the release of perforins and granzymes by CD8+ cells. For the identification of CD8+ specific epitopes three different tools

4.2.1 BIMAS

BIMAS predict the potential epitopes for 33 HLA Class I alleles on the basis of half time of dissociation from the β2 microglobulin of Class I HLA molecules. The tool predicts the binding of the peptide to the HLA from the available peptide binding data of 152 peptides and depending upon their respective position on epitope. Eighty peptides with binding half-life of binding with β2 microglobulin more than 5 min has been included in the algorithm for the prediction of potential epitopes similarly the remaining 72 peptides which had half-life less than 5 min used to eliminate the false prediction (Parker *et al*., 1994). The cut off value 50, was used for BIMAS for the epitope prediction. It may be helpful for the prediction of stable and potential epitopes.

4.2.2 SYFPEITHI

SYFPEITHI makes the predictions for both class I and class II HLA alleles of human, mouse, rat, ape, cattle, and chicken. The algorithm identifies the anchor and auxiliary residues in the input sequence by comparing it with the database of naturally occurring epitopes. It assigns particular score to each residue depending upon their position and generate overall score for predicted epitope (Rammensee *et al*., 1999). The score is calculated by assigning amino acids of a certain peptide a specific value depending on the fact if they are anchor, auxiliary anchor or preferred residue. Anchors are given 10 points, unusual anchors 6-8 points, auxiliary anchors 4-6 and preferred residues 1-4 points. The amino acids which are responsible for

hindering the binding ability are assigned score between -1 and -3. The cut off value taken was 20 for the predicted epitopes.

4.2.3 NetCTL 1.2

NetCTL 1.2 leads to the epitope prediction on the basis of C-terminal proteasomal cleavage, TAP transport efficiency and MHC binding. It identifies the epitopes for 12 HLA-A and HLA-B super types (Larsen *et al*., 2007). The cut off value >0.75 was used for the epitope prediction. The cut off 0.75 signifies the 80% sensitivity and 0.97 specificity score of the predicted epitopes. The combined three scores results in the identification of the potential immunogenic epitopes from the given sequence.

The common epitopes predicted by all the three tools were selected. The selected epitopes were overlapped to generate the peptide fragments containing multiple epitopes.

4.3 CD4+ T-cell Epitope prediction

CD4+ cells recognise the antigen presented on MHC class II molecules and activate the CD8+ cells which in turn perform the cytotoxic action. To identify the CD4+ specific epitope three different tools Rankpep (Reche *et al*., 2007), Propred (Singh *et al*., 2001) and IEDB consensus (Wang *et al*., 2008) were used.

4.3.1 Rankpep

It predicts the Class II binding peptides from already known class II MHC binding aligned peptides by using PSSM (Position Specific Scoring Matrix). (Reche *et al*., 2007). The MHC binding score of the peptide is obtained by aligning the PSSM with the protein segment and assigning the score to the residue which matches to the profiles in the matrix. The threshold value for the epitope prediction was taken top 2% of the predicted peptides.

4.3.2 Propred

It uses the quantitative matrix based prediction algorithm for the prediction of epitopes by using co-efficient table created with previously reported epitopes from literature (Singh *et al*., 2001). The top 2% of the predicted epitopes were taken as the threshold for the epitope identification.

4.3.3 IEDB consensus

The IEDB consensus apply the multiple algorithms, NN-align, SMM-Align and combinatorial method for the prediction of available epitopes otherwise uses the netMHCpanII. The performance for MHC-II binding methods in are in order: Consensus > NetMHCIIpan > NN-align > SMM-align > Combinatorial Library. Based upon the epitope predicted by different algorithms a percentile rank is generated for the finally predicted epitopes (Wang *et al*., 2008). The threshold values taken for IEDB consensus were percentile rank of top 10% of predicted epitopes. The common epitopes predicted by all the three tools were selected.

4.4 Blast Screening

In order to avoid any similarity of the peptides with functional human protein other than metadherin, BLAST analysis (Altschul *et al*., 1990) was performed for epitopes predicted to bind MHC class I and II respectively. The peptides showing similarity in 7 out of 9 amino acids without gap or mismatch were eliminated (Tan *et al*., 2010) thus ruling out any possibility of autoimmune response against any human functional protein. The peptides screened after BLAST showing overlapped and merged together to generate single peptide fragment.

4.5 Population Coverage Analysis

MHC alleles are highly polymorphic and diversity of alleles were found in the global population. Different individuals of the worldwide population may respond in different way against particular antigen. The IEDB database enables predicting the possible world population capable of responding to particular immunogen. The database contain variety of MHC class I and II alleles frequently found in different world population of geographical areas. It compares the epitopes specific alleles with the alleles in database and find out percentage population capable of responding. In our study IEDB population coverage analysis tool was carried out for 16 different geographical areas distributed globally (Huynh-Hoa Bui et al., 2006).

4.6. Structure Prediction of metadherin and immunogenic peptide labelling

The structure of metadherin has not been resolved yet, the putative structure was predicted using online structure prediction server I-TASSAR (Zhang, 2008). The predicted peptide

enriched with CD8+ and CD4+ specific epitopes were labelled on the predicted structure with the help of discovery studio visualizer 4.0.

4.7 Epitope Docking

4.7.1 Peptides Structure Prediction

The structure of predicted epitopes common for class I and Class II MHC were predicted using peptide structure prediction server PEP-FOLD (Thevenet *et al*., 2012). The model generated by PEPFOLD are assorted either by using the coarse grain energy of the PEPFOLD or by predicted Tm score. For the peptides of length 36 residues the coarse grain energy of the PEPFOLD is used to sort the predicted models whereas Tm score for the Peptides with the residues more than 36.

4.7.2 Separation of naturally bound peptide from MHC molecule.

The structure of 22 peptide bound MHC class I and 6 MHC class II were retrieved. The Ligands and receptor of the PDB structure were separated using the discovery studio visualization tool 3.5. The generated peptide free MHC structures were further used for docking using Auto dock Vina (Trott *et al*., 2009) tool.

Ligand Separation

Open the PDB structure of HLA molecule in discovery studio visualization tool 3.5

⬇

Go to the scripts option and select the water molecules

⬇

Go to edit and delete

⬇

Go to scripts and select protein chains

⬇

Go to edit and delete

⬇

Save as ligand.pdb file

Receptor Separation

Open the PDB structure of HLA molecule in discovery studio visualization tool 3.5

Go to the scripts option and select the water molecules

Go to edit and delete

Go to scripts and select ligands

Go to edit and delete

Save as the receptor.pdb file

Receptor preparation for docking

Open the PDB structure of receptor molecule in Auto dock Vina 4.0

Go to the edit and add hydrogen atoms (polar only)

Go to grid and choose and select the macromolecule i.e. receptor molecule

Go to grid box and select the grid where the ligand should be bound

Save the file as receptor.pdbqt

Ligand preparation for docking

Open the ligand.pdb file in the Auto dock Vina 4.0

Go to ligand and choose the molecule

Select the ligand molecule

Go to output and save the molecule as ligand.pdbqt

4.7.3 Preparation of configuration file

To perform the docking in auto dock vina we are supposed to generate a text file named as config containing the following commands. The config file should be created for the each molecule separately depending upon their grid box selection.

receptor = receptor.pdbqt

ligand = ligand.pdbqt

out = out.pdbqt

center_x = 72.065 (molecules specific obtained by grid box selection)

center_y = -40.833(molecules specific obtained by grid box selection)

center_z = -10.862(molecules specific obtained by grid box selection)

size_x = 60(molecules specific obtained by grid box selection)

size_y = 72(molecules specific obtained by grid box selection)

size_z = 92(molecules specific obtained by grid box selection)

Exhaustiveness = 16.

4.7.4 Docking

The separated naturally bound peptides of the MHC molecules were redocked with the MHC peptide free MHC using AutoDock Vina (Trott *et al*., 2009) tool. The binding energy of these naturally bound ligands of MHC molecules was determined. The predicted epitopes were docked with the MHC class I and II structure and binding energy was compared with the

binding energy of the redocked naturally bound peptides. The docking was performed at exhaustiveness 16 with all the peptides and MHC structures.

4.8 Prediction of B-cell specific linear epitope

Antibodies generated in response to an immunogen on the cell surface leads to cell killing by mean of antibody dependent cell mediated cytotoxicity (ADCC). A humoral immune response against cancer antigen can help in cancer cell killing by mean of same mechanism. To identify the B-cell specific linear epitope IEDB Kolaskar and Tongaonkar (Kolaskar *et al*., 1990) method was used, using default threshold value 0.992 and window size 7. This method apply the semi empirical approach to identify the antigenic determinant on protein depending upon their on physicochemical properties and frequency of occurrence of amino acids on it based upon the previously known B-cell epitopes, The tool predicts the antigenic propensity of the predicted peptides. The fundamental of the determining the antigenicity is that if cysteine, leucine and valine occur on the surface of the protein they are more likely to act as B-cell antigenic site. The percentage accuracy of this method is found out be nearly 75% which is better than any other known B-cell epitope prediction method (Kolaskar *et al*., 1990). The predicted B-cell epitopes were further compared with the already predicted T-cell epitopes explained above, to find out if any similarity between T-cell and B-cell linear epitopes exists.

4.9 *Invitro* analysis of synthetic peptide for immunogenic response

4.9.1 Peptide synthesis

A predicted nonamer epitope FLLGYGWAA common for both class I and class II was synthesized from the Genscript (USA) having purity level more than 90%. The peptide was dissolved in 0.7% aqueous ammonia and used for further analysis.

4.9.2 MTT assay for lymphocyte proliferation against synthetic peptide

Peripheral blood mononuclear cells (PBMCs) were isolated by ficoll density gradient centrifugation. Blood was drawn from a healthy person with the help of vacutainer system (EDTA coated, Becton Dickinson). Blood was diluted in to three time of its volume with PBS. Then blood sample was layered carefully over Histopaque 1077 in the ratio 1:1 and it was centrifuged at 400xg for 30 minutes at 25^0 C in swinging bucket rotor. Plasma was removed and then the Buffy coat layer was transferred to a clean centrifuge tube. Buffy coat layer was washed twice with 3 volumes of PBS by centrifuging at 250xg for 10 minutes at

25^0C. Supernatant was discarded and pellet of PBMC was suspended in 1 ml of cell culture medium RPMI 1650 [RPMI + 10% FBS + Penicillin (100units/ml) + Streptomycin (100 µg/ml)] and cells were counted using haemocytometer. The cells were cultured in round bottom 96 well microtiter plate (1 lakh cells/well) with peptide concentration 10µg/ml, ConA (Concavilin A) 5µg/ml as positive control in RPMI 1640 media for 6 days at 37º C and 5% CO_2 concentration. After 6 days, 20µl of MTT (3-(4, 5-dimethythiazol-2-yl)-2, 5-diphenyl tetrazolium bromide) (5 mg/ml) was added to each well and an incubation of another 4 hours was given for reduction of MTT to formazan. The MTT enters the cells and passes into the mitochondria where it is reduced to an insoluble, coloured (dark purple) formazan product. The cells are then solubilised with an organic solvent and the released, solubilised formazan reagent is measured spectrophotometrically. Since reduction of MTT can only occur in metabolically active cells the level of activity is the measure of the viability of the cells. Therefore cell proliferation was tested using MTT assay. Media was removed carefully and purple formazan crystals were dissolved in 100µl of DMSO. Absorbance was recorded at 570 nm by microtiter plate reader.

.

5.1 $CD8^+$ T-cell specific epitope identification

Metadherin is primarily involved in metastasis of different cancers. Chemo resistance and radio resistance due to its expression indicates that it plays an important role in increasing the adversity of cancer. Targeting it as a vaccine design not only make it a favourable vaccine candidate for different cancers, but can also help in targeting metastasis. Since three different epitope prediction tools were used (NetCTL, BIMAS and SYFPEITHI), initially 33 CD8+ T-cell binding epitopes were predicted commonly by all the three tools used. After BLAST screening 2 out of 33 commonly predicted class I MHC restricted epitopes showed similarity with other human functional protein in 7 out of 9 amino acids, hence were eliminated. Out of 31 epitopes 22 showed the overlapping (Table 2) to generate 9 peptides fragments ranging from 9 to 20 amino acid containing one or more than one epitope (Table 3).

Table 2: CD8+ T-cell specific Epitopes screened after BLAST analysis showing overlapping

CD8+ specific T-cell Epitopes	**Amino Acid Position**
ARLREMLSV	19
LREMLSVGL	21
MLSVGLGFL	24
LRTELGLDL	32
TELGLDLGL	34
FLLGYGWAA	65
VPAAAPDDL	95
AAAPDDLAL	97
AAPDDLALL	98
KPKPNGRTV	125
KPNGRTVEV	127
VEVAEGEAV	133
HREKRQQRK	198
KRQQRKRDK	201
QQRKRDKVL	203
KVLTDSGSL	209
GRSWSDRSI	337
RSWSDRSIF	338
YIDDEWSGL	374
DEWSGLNGL	377
LEKEIREDL	465
KEIREDLPV	467

Table 3: Peptide containing overlapping CD8+ specific T-cell epitopes

Peptide Enriched CD8+ specific T-cell epitopes	Number of Epitopes
ARLREMLSVGLGFL	3
LRTELGLDLGL	2
FLLGYGWAA	1
VPAAAPDDLALL	3

KPKPNGRTVEVAEGEAV	3
HREKRQQRKRDKVLTDSGSL	4
GRSWSDRSIF	2
YIDDEWSGLNGL	2
LEKEIREDLPV	2

5.2 CD4^{+} T-cell specific epitope identification

CD4+ cells recognise the antigen presented on MHC class II molecules and activate the CD8+ cells which in turn perform the cytotoxic action. Thus both class I and class II specific epitopes are vital for the vaccine design as there can be more robust immune response against the antigens. Three tools used for Class II MHC (IEDB consensus, Rankpep and Propred) predicted 58 common epitopes specific for CD4+ T-cell. Four out of 58 epitopes were eliminated by BLAST screening because of their similarity with other human functional protein in 7 out of 9 amino acids. Out of these 54 amino acids 35 epitopes showed overlapping (Table 4) and gave rise to 11 peptide fragments ranging from 12 to 17 amino acids containing more than one epitope (Table 5).

Table 4: CD4+ specific T cell Epitopes showing overlapping after BLAST analysis

CD4+ specific T cell Epitopes			
Epitope	Position	Epitope	Position
LREMLSVGL	21	LNENLTVNG	274
MLSVGLGFL	24	WNSVSPASA	303
LSVGLGFLR	25	VSPASAGKR	306
VGLGFLRTE	27	IFSGIGSTA	344
LGFLRTELG	29	FSGIGSTAE	345
YPGWVILVG	47	IGSTAEPVS	348
WVILVGTGA	50	YIDDEWSGL	373
VILVGTGAL	51	WSGLNGLSS	378
ILVGTGALG	52	LNGLSSADP	381
LVGTGALGL	53	VLVKNSQPI	500
FLLGYGWAA	65	LVKNSQPIK	501
LLGYGWAAA	66	VKNSQPIKT	502
LGYGWAAAC	68	IKTLPPATS	508

WAAACAGAR	70
LALLKNLRS	102
LKNLRSEEQ	105
LRSEEQKKK	108
IENTITVTT	223
ITVTTEQLT	227
LNVQVSNFK	252
VSNFKSGKG	256
VSSGLNENL	270

Table 5: Peptide containing CD8+ specific T-cell epitopes

Peptides Containing CD4+ specific T-cell epitopes	Number of Epitopes
LREMLSVGLGFLRTELG	5
YPGWVILVGTGALGL	5
FLLGYGWAAACAGAR	4
LALLKNLRSEEQKKK	3
IENTITVTTEQLT	2
LNVQVSNFKSGKG	2
VSSGLNENLTVNG	2
WNSVSPASAGKR	2
IFSGIGSTAEPVS	3
YIDDEWSGLNGLSSADP	3
VLVKNSQPIKTLPPATS	4

5.3. Identification of common $CD4^+$ and $CD8^+$ specific epitope containing peptide fragments

An overlapping was observed between three CD4+ and CD8+ specific epitope enriched peptides fragments. It gave rise to three larger peptide fragments containing CD4+ and CD8+ specific epitope (Table 6) which contain six class I MHC and twelve Class II MHC specific nonamer epitopes (Table 7). Out of which there were four nonamer epitopes same for both class I and Class II MHC (Table 7).The peptide fragments were found to cover larger number of MHC class I and II alleles (Table 8).

5.4. Population coverage Analysis

It is important to find out the expected immune response generation in the global population by these predicted peptide fragments which was done by population coverage analysis. Three peptides containing CD4+ and CD8+ specific epitopes as shown in table 7 were considered for this analysis. These peptides and corresponding MHC restricted alleles obtained by prediction servers were the input data in the IEDB population coverage tool. Then the tool calculates the expected response of these peptides in different population based on these input data. The average population coverage for immunogenic response of predicted peptides were found out to be 64.46 % and 97.15 % for class I and Class II MHC specific respectively (Figure 1).

Table 6: Metadherin peptides containing overlapping CD8+ and CD4+ specific T-cell Epitopes

Peptide enriched CD8+ specific T-cell epitopes	No. of Epitopes	Peptides enriched CD4+ specific T-cell epitopes	No. of Epitopes	Peptide containing CD8+ and CD4+ specific T-cell epitopes
ARLREMLSVGLGFL	3	LREMLSVGLGFLRTELG	5	ARLREMLSVGLGFLRTELG
LRTELGLDLGL	2	YPGWVILVGTGALGL	5	
FLLGYGWAA	1	FLLGYGWAAACAGAR	4	FLLGYGWAAACAGAR
VPAAAPDDLALL	3	LALLKNLRSEEQKKK	3	
KPKPNGRTVEVAEGEAV	3	IENTITVTTEQLT	2	
HREKRQQRKRDKVLTDSGSL	4	LNVQVSNFKSGKG	2	
GRSWSDRSIF	2	VSSGLNENLTVNG	2	
YIDDEWSGLNGL	2	WNSVSPASAGKR	2	
LEKEIREDLPV	2	IFSGIGSTAEPVS	3	
		YIDDEWSGLNGLSSADP	3	YIDDEWSGLNGLSSADP
		VLVKNSQPIKTLPPATS	4	

Table 7: Final set of peptides containing both CD4+ and CD8+ T-cell specific epitopes and common nonamer epitopes for CD8+ and CD4+ T-cell

Peptides containing CD8+ and CD4+ specific T-cell epitopes	Position	CD8+ specific T-cell Epitopes	CD4+ specific T-cell Epitopes	Common CD8+ and CD4+ T-cell Specific epitopes
ARLREMLSVGLGFLRTELG(P1)	19-37	ARLREMLSV LREMLSVGL MLSVGLGFL	LREMLSVGL MLSVGLGFL LSVGLGFLR VGLGFLRTE LGFLRTELG	LREMLSVGL (M1) MLSVGLGFL (M2)
FLLGYGWAAACAGAR (P2)	65-78	FLLGYGWAA	FLLGYGWAA LLGYGWAAA LGYGWAAAC WAAACAGAR	FLLGYGWAA (M3)
YIDDEWSGLNGLSSADP (P3)	374-390	YIDDEWSGL DEWSGLNGL	WSGLNGLSS LNGLSSADP YIDDEWSGL	YIDDEWSGL(M4)

Table 8: Number of MHC class I and II alleles capable of responding against peptide fragment containing both CD4+ and CD8+ specific T-cell epitope

Peptide Fragment	Number of T- cell specific MHC Class I specific Alleles	Number of T- cell specific MHC Class II specific Alleles
ARLREMLSVGLGFLRTELG	97	182
FLLGYGWAAACAGAR	54	66
YIDDEWSGLNGLSSADP	121	57

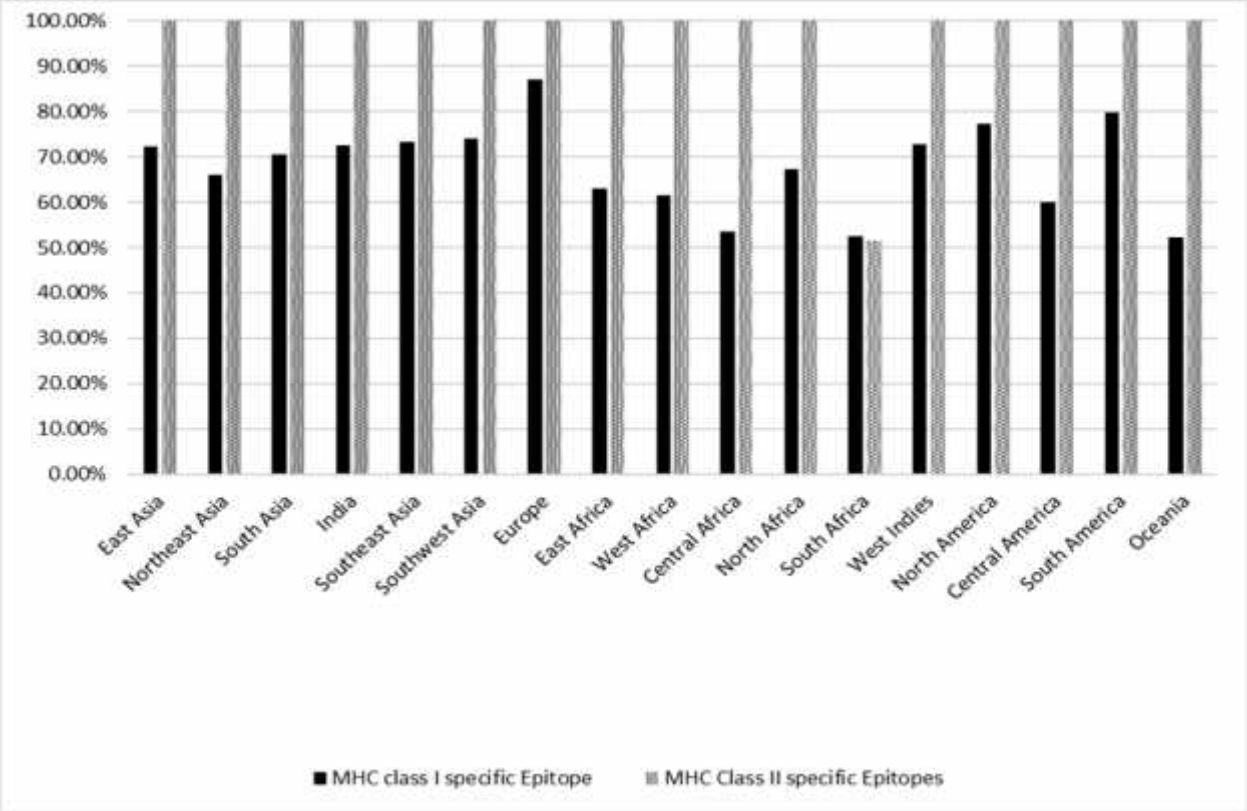
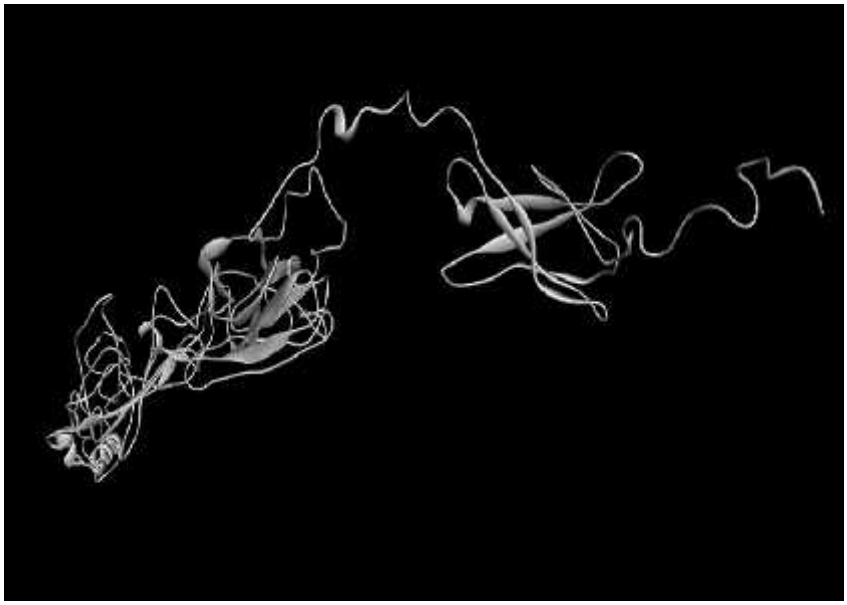


Figure 1: Population coverage analysis of peptides containing CD4+ and CD8+ T cell epitopes for class I and II MHC alleles in17 different geographic regions predefined in the IEDB database.

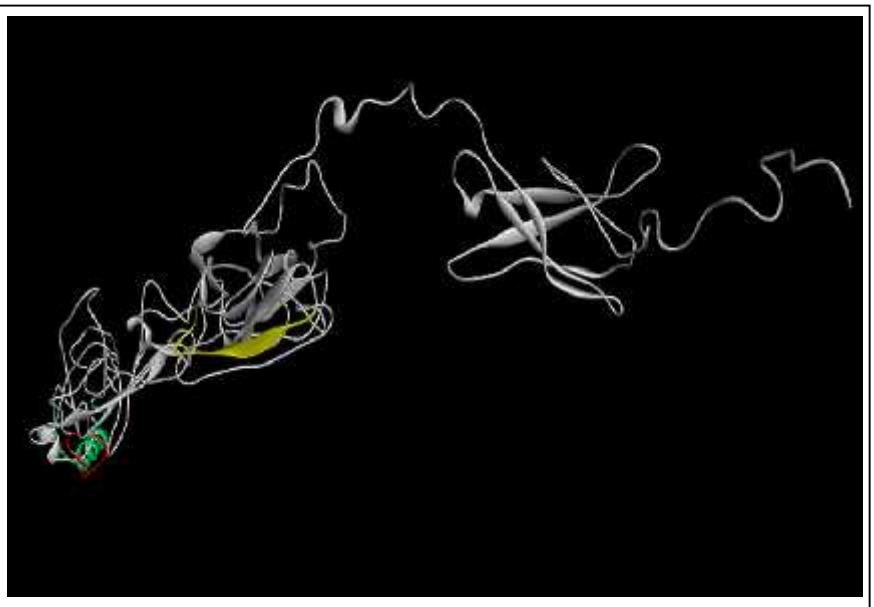
5.5. Structure Prediction of metadherin and immunogenic peptide labelling

I-TASSAR server predicts the structure of the protein on the basis of multiple threading alignments and iterative template fragment assembly simulations. It predicts the possible

functions of the protein by comparing the three dimensional model with the BioLIP protein function database (Zhang. 2008). The best one of five predicted structure was taken and three identified peptides containing CD4+ and CD8+ T cell epitopes as given in table 7 were colour labelled to visualize the position of immunogenic peptide fragments in protein. The labelled peptides are shown in figure 2.



(a)



(b)

Figure 2. (a) Predicted structure of Metadherin by I-TASSAR server (b) CD4+ and CD8+ specific T-cell epitope enriched peptide labelled on the peptide structure. Where green, red and yellow colour depicts P1, P2 and P3 respectively

5.6 Peptides Structure Prediction

PEP-FOLD predict the structure from the amino acid sequence. It is based upon structure alphabet letters to describe the structure of the given sequence. It predicts the structure of four consecutive amino acids and then combine the series of structural alphabet letter to generate the structure (Thevenet *et al*., 2012). The 5 high scoring models were generated by the server and the best model of each M1, M2, M3 and M4 peptide ((Figure 3) was taken and subjected for epitope docking.

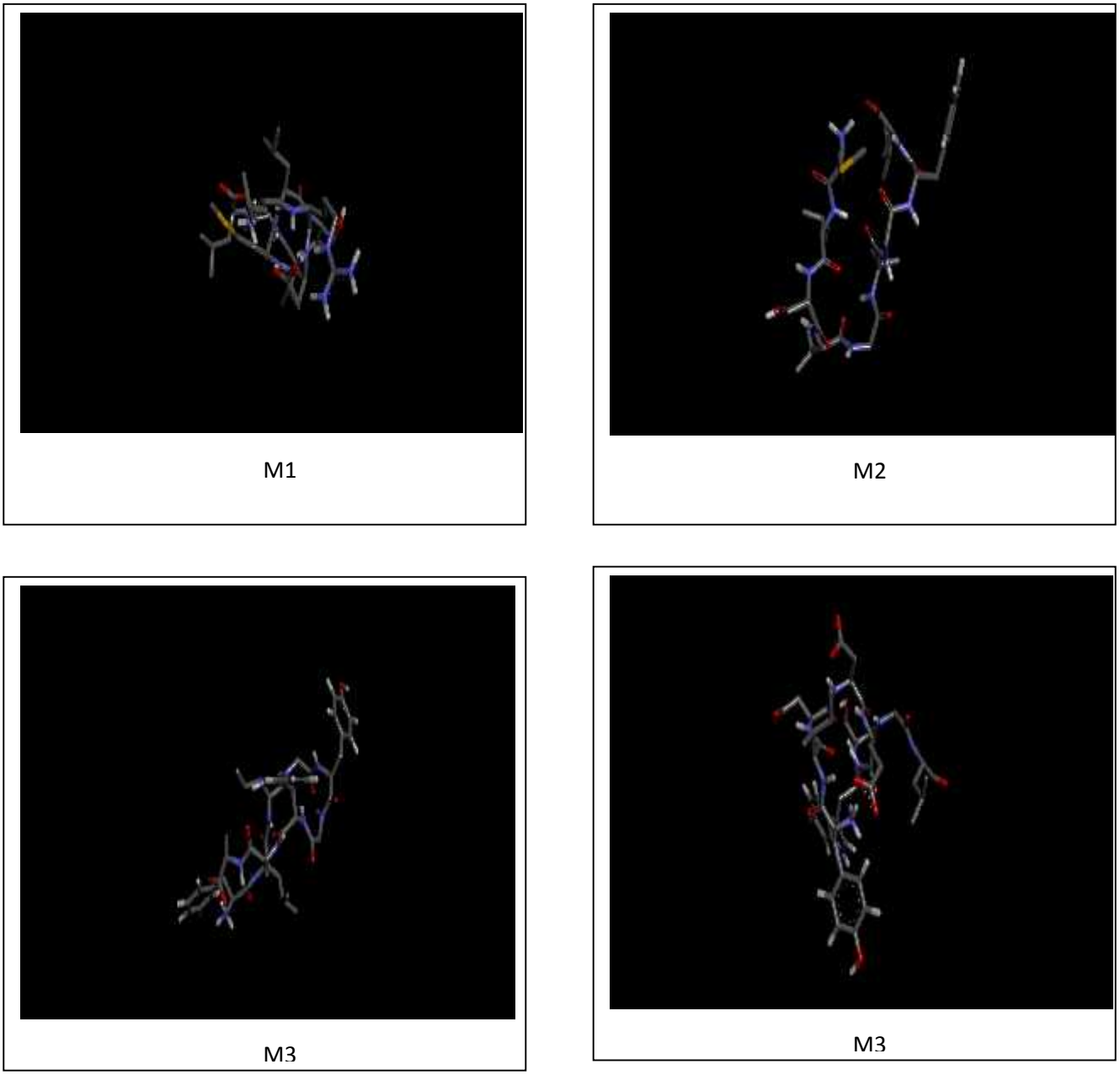


Figure 3: PEP-FOLD predicted structure of common CD8+ and CD4+ specific nonamer epitopes (M1, M2, M3 and M4)

5.7 Epitope Docking with Class I and II MHC molecules

The immunogenic potential of the predicted epitopes depends upon the fact that how efficiently it can bind to peptide binding pocket of the MHC molecule. The auto dock vina tool has been used to assess the binding affinity of the predicted epitope with MHC class I and class II molecules. Twenty two and six peptide MHC complex for Class I and II respectively were retrieved from PDB database for docking purpose. All peptides were extracted from the MHC complex and then redocked with corresponding MHC molecules using auto dock vina. Four epitopes (M1, M2, M3 and M4) (Figure 3) which are common for both CD4+ and CD8+ T cell epitopes as given in table 7 were docked with each MHC Class I molecule. The binding energies of the peptides capable of binding to MHC molecule were compared with the naturally bound peptide. The peptides bound outside the peptide binding motif were assigned nil binding energy (Table 9). The individual binding energies of different class I MHC molecules after docking with peptides were comparable to the binding energy after docking of the naturally bound peptides to MHC structure (Figure 4a. & 4b.)

Table 9: Binding Energies of epitopes common for both CD4+ and CD8+ T cell with Class I MHC alleles.

Class I MHC alleles	Binding Energy (Kcal/mol)				
	Redocked Naturally Bound Ligand	M1	M2	M3	M4
A1	8.8	6.1	6.5	8.8	7.5
A201	9.8	Nil*	Nil	Nil	Nil
A301	7.2	6.3	7.4	9.2	7.6
A1101	7.7	6.6	6.2	8.1	7.1
A2402	8.9	6.6	6.9	7.3	6.9
B8	10.5	8.3	8.5	10.6	7.1
B1402	10.3	Nil	7.9	8.7	8.1
B1501	9	6.7	6.1	9	8.3
B1801	8.8	8.3	7.7	8.9	8.5
B2705	7.7	7.4	7.1	8.6	8
B2709	9.7	7.2	6.6	8.9	7.4
B3501	7.4	8	7.5	9.4	8.7
B3508	14.7	8.2	8.3	10.4	9.5

B4103	7.5	Nil	Nil	9.4	Nil
B4104	7.5	6.6	5.7	8.8	8.5
B4402	6.3	4.9	5.8	6.6	7.2
B4403	7.6	8.2	8.6	10.4	9.7
B4405	7.6	6.8	8.4	Nil	7.4
B5703	8.1	7.4	7.5	9.5	9
B5701	8.2	8.1	7.4	10.8	9.5
Cw3	8.1	8.2	Nil	9.6	7.4
Cw4	8	8.5	7.6	10.1	9.5

*Peptide bound outside the epitope binding pocket of MHC

Similarly the binding energies of the epitopes (M1, M2, M3 and M4) which are common for both CD4+ and CD8+ T cell epitopes as given in table 7 were docked with each MHC Class II molecule. The peptide capable to bind in MHC pocket were screened and binding energies were obtained (Table 10). Similar comparable pattern of the binding energies was observed for class II MHC molecule between docking of naturally bound peptides and M1, M2, M3 and M4 (Figure 5) respectively.

Table 10: Binding Energies of epitopes common for both CD4+ and CD8+ T cell with Class II MHC alleles.

	Binding Energy (Kcal/mol)				
HLA Class II Molecule	Redocked Naturally Bound Ligand	M1	M2	M3	M4
HLA DP2	7.5	Nil*	Nil*	8.5	7.9
HLA DQ2	9.5	7	7.4	9.4	8.8
HLA DR1*01:01	7.2	6.6	6.7	9	8.5
HLA DR1	6.6	6.5	6.2	9.9	7.6
HLA DR2	6.8	6.7	7.5	9	7.9
HLA DR4	9.3	8.3	7.3	8.1	8

*Peptide bound outside the epitope binding pocket of MHC

The average binding energies of these four epitopes were found to be within the energy range of the redocked peptides for both MHC class (Table 11 & 12).

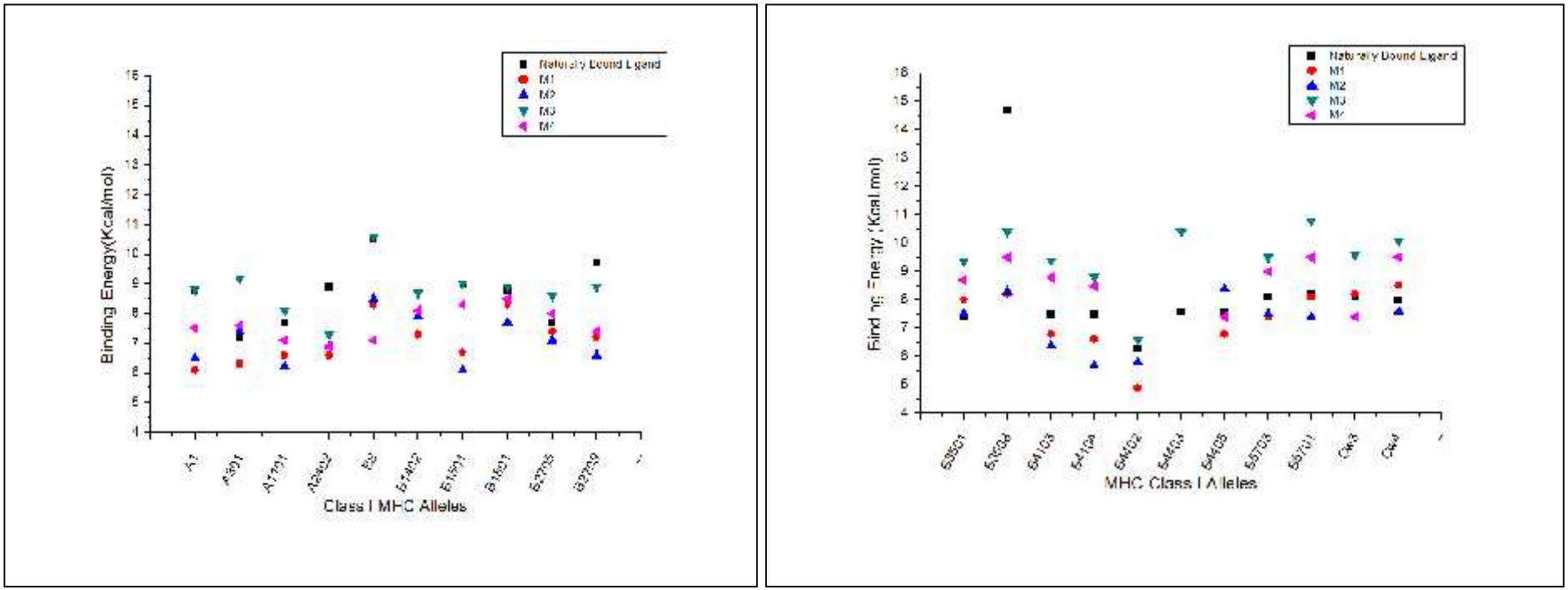
Table 11: Average binding energies of the predicted peptides and naturally bound peptides after docking with MHC class I structures by using auto dock vina.

Binding Energy (Kcal/mol)			
	Average	Standard Deviation	Range
Naturally bound peptide	8.6	1.7	7.2-14.7
M1	7.3	1	6.1-8.8
M2	7.3	0.9	5.7-8.7
M3	9.3	1.1	6.6-10.8
M4	8.2	0.9	6.9-9.7

Table 12: Average binding energies of the predicted peptides and naturally bound peptides after docking with MHC class II structures by using auto dock vina.

	Binding Energy (Kcal/mol)		
	Average	Standard Deviation	Range
Natural bound peptide	7.8	1.1	6.6-9.5
M1	6.9	0.6	6.4-8.3
M2	7.0	0.4	6.2-7.5
M3	8.8	0.7	7.6-9.9
M4	8.1	0.4	7.6-8.8

The docking poses of the bound and unbound peptides are given in figure 6 and 7 for selective MHC class I alleles to demonstrate the docking. Although the comparable docking energies of the non-binding peptide were obtained but the peptides not binding in the pocket of MHC alleles were eliminated despite of having comparable binding energy. Only the bind peptides were considered for further evaluation.



(a) (b)

Figure 4. Binding energy of naturally bound peptides and M1, M2, M3 and M4 epitopes obtained after docking to 21 class I MHC alleles.

(a) First eleven out of 10 class I MHC alleles (b) Rest eleven class I MHC alleles

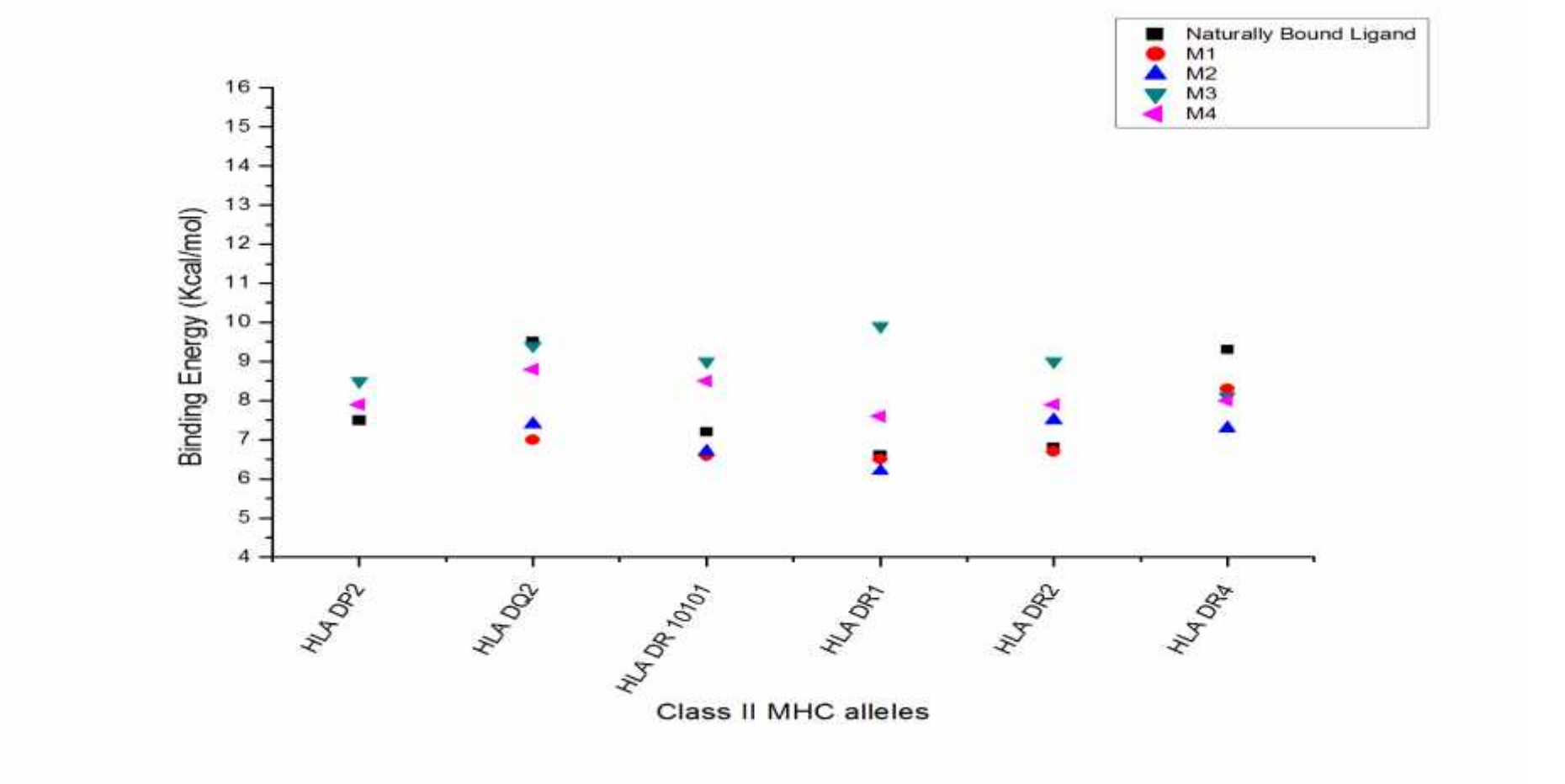
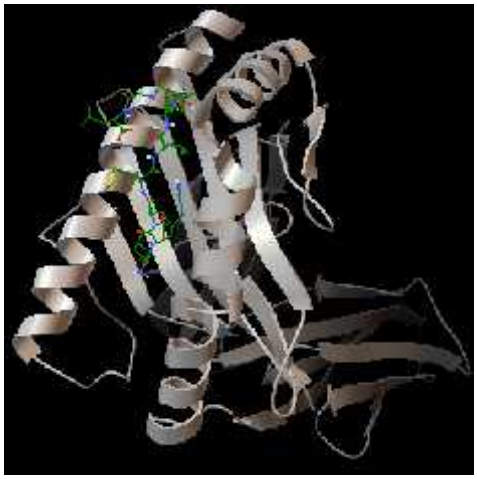


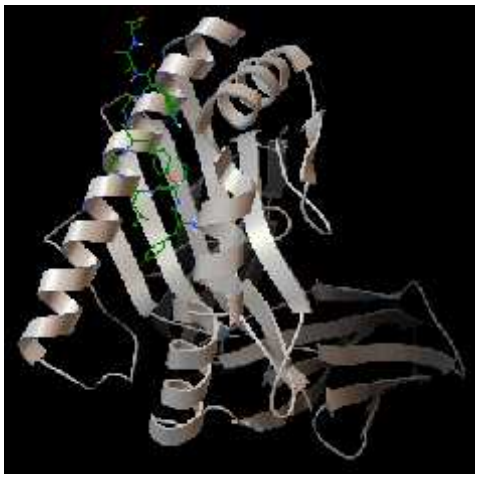
Figure 5. Binding energy of naturally bound peptides and M1, M2, M3 and M4 epitopes obtained after docking to six class II MHC alleles

Docking Poses of HLA A2402

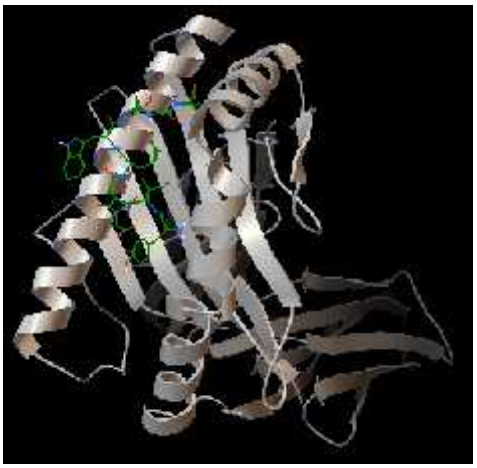


M1

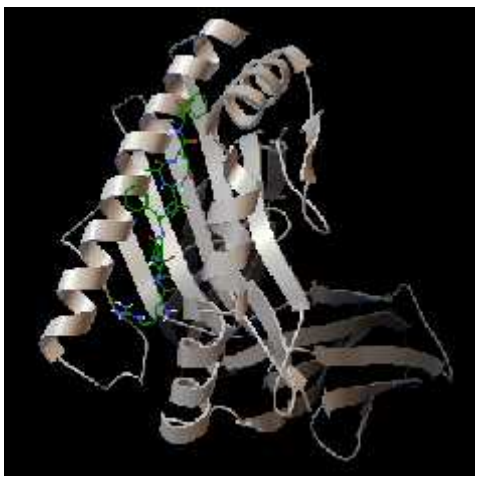
M2



M3



M4

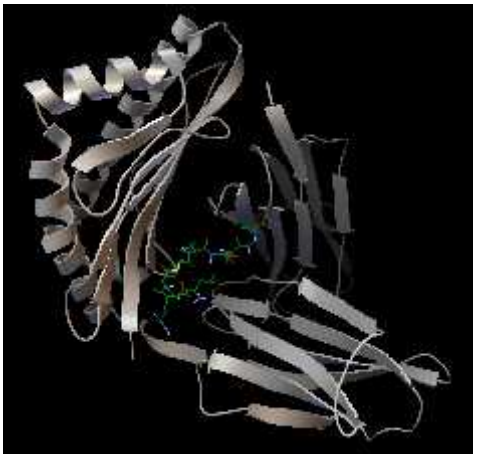


Redocked Naturally Bound Lignad

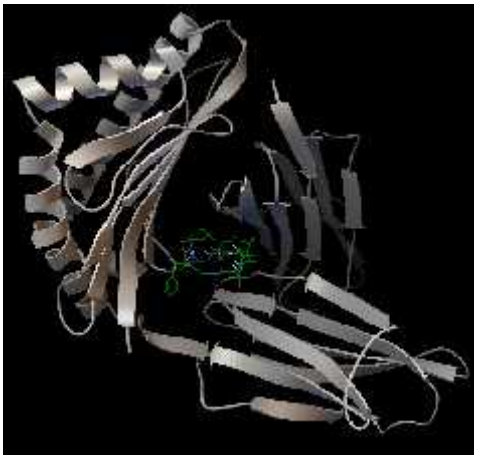
Figure 6: The peptide M1, M2, M3 and M4 are docked to the peptide binding pocket of HLA A2402.

As that of naturally bound ligand showing binding ability of M1, M2, M3 and M4 towards HLA A2402

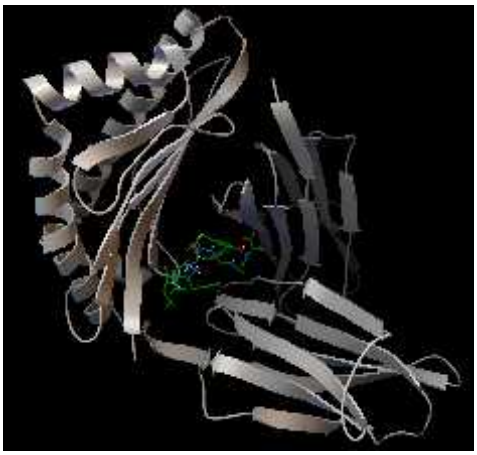
Docking Poses of HLA A 201



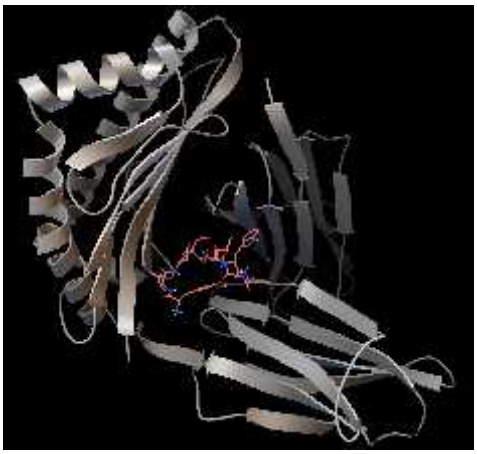
M1



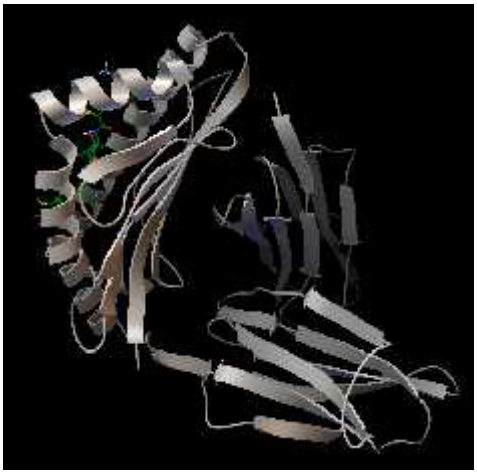
M2



M3



M4



Redocked Naturally Bound Ligand

Figure 7: The peptide M1, M2, M3 and M4 are docked at the site peptide binding pocket of HLA A201 whereas of naturally bound ligand is docked into epitope binding pocket showing non-binding ability of M1, M2, M3 and M4 towards HLA A201

5.8 Identification of B-cell epitope

One of the major arm of immune response is humoral immune response i.e. antibody mediated response. B cell produces antigen specific antibodies hence it will be interesting if the predicted peptide have the potential to generate humoral immune response. Linear B cell epitope prediction of metadherin protein was carried out by IEDB Kolaskar and Tongaonkar tool. There were twenty one linear B-cell specific immunogenic peptides found ranging from six to thirty one amino acid long. After comparing these epitope with finally selected peptides containing T cell epitopes, two of the predicted immunogenic peptides were found to be common for both T and B cell epitope ((Table 13).

5.9 MTT assay for lymphocyte proliferation against synthetic peptide

The peptide M3 was found to stimulate PBMCs cell division in comparison to peptide solvent control (0.7 % aqueous ammonia), which signifies its potential to generate the immune response. The control peptide N1 (LQIGNIINSIW), which is an epitope of neuraminidase protein of H1N1 virus predicted by immunoinformatics (Lohia *et al*., 2014) showed relatively more proliferation comparison to solvent control (10 % DMSO control) and M1. Whereas cells with positive control Con A showed a very sharp rise in cell number. In comparison to the cell control the cell number of stimulated peptide was found to be reduced which indicates the effect of solvent on cell number (Figure 8). Hence in 6 day culture the peptide is found to be stimulate the PBMCs signifying their immunogenic potential.

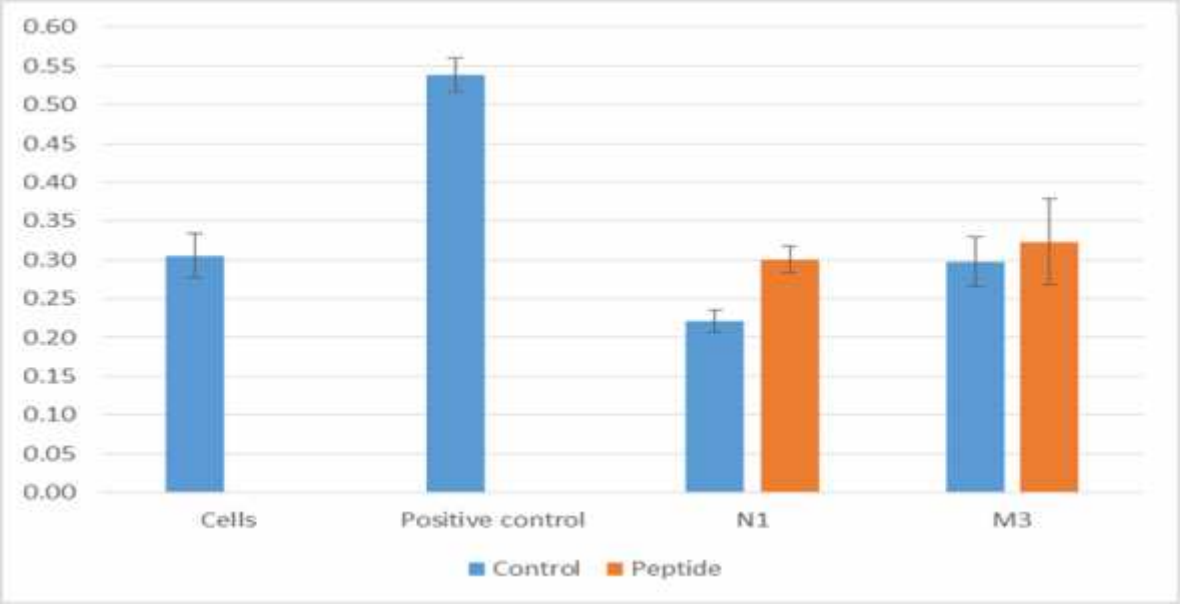


Figure 8: Proliferation assay of peripheral blood mononuclear cells (PBMCs) against peptide M3 using MTT assay.

Table 13: Predicted B-cell epitopes and peptides showing both T and B cell epitopes in metadhein protein

Peptide	**Peptides containing CD4+ and CD8+ T cell epitopes**	**Peptides containing both T and B cell epitopes**
DELAQQ		
MLSVGLGFLRTE	ARLREMLSVGLGFLRTEL	ARLREMLSVGLGFLRTEL
GLDLGLEPK		
PGWVILVGTGALGLLLLFLLGYGWAAACAGA	FLLGYGWAAACAGAR	PGWVILVGTGALGLLLLFLLGYGWAAACAGA
AAAVPAAAPDDLALLKN		
VRTPQSVTAK		
AKAVQNS		
DKVLTDS		
TTASFPVGS		
HLNVQVSNF		
TLQVSSG		
SVKLSSQIS		
SVSPAS		
AEPVSQS		
ASLLKSQEP		
DLPVNT		
AFSLKTI		
PAEVLVKNS		
IKTLPPA		
EPSVILSKS		
SSQVPPILQE		

Chapter 6. DISCUSSION

Bioinformatics driven vaccine design research is emerging as a new way to lessen time in vaccine design process**.** To bypass the exhaustive and expensive conventional vaccine design approach, *in silico* vaccine design algorithms provides a practical alternative for it. Vaccine development against cancer has taken some advancement in last decade an *in vivo* study for HER-2/neu specific peptide based vaccine has shown immune response in the patients with breast, ovarian and small cell lung carcinoma which overexpresses HER-2/neu (Disis *et al*., 2002). Similarly *in vivo* study against NY-ESO-1, a cancer testis antigen expressed in epithelial ovarian cancer has shown immune response against cancer cells (Odunsi *et al*., 2007). At the same time *in silico* studies has also shown the significant results *in vitro* as well as *in vivo,* identification of HER/neu specific epitope predicted by SYPEITHI has shown immunogenic response *in vitro* and *in vivo* (Gritzapis *et al*., 2010). There are other reports which promises the immunogenic response of *in silico* predicted peptide already confirmed in *in vitro* and *in vivo* tests. But majority of these studies have used one (Gritzapis *et al*., 2010) or two (Yuan *et al*., 2012) tool for the epitope prediction. Application of six different algorithms (different tools) in our study was to ensure the immunogenic potential of the predicted epitopes. Also the same epitopes predicted by different algorithms strengthen the potential of putative epitopes for immunogenicity. Further the selection of common epitopes enriched peptide fragments for both class I and Class II MHC may give rise to overall T-cell specific immune response. At the same time similarity between T-cell and B-cell linear epitopes may be helpful in choosing vaccine candidates capable of generating overall immune response. The extracellular lung homing domain in metadherin from amino acid 378-440 is responsible for the breast to lung metastasis (Brown *et al*, 2004). One of the predicted common class I and class II peptide fragment YIDDEWSGLNGLSSADP is located at amino acid position 374-390 which is part of lung homing domain. If in further analysis it will be found immunogenic it may be helpful in preventing breast to lung metastasis. There are reports which shows the presence of metadherin specific auto antibodies in the serum of cancer patients (Chen *et al*., 2012) dictating the immunogenic property of metadherin. Which again may validate our claim to choose the metadherin as a potential candidate for cancer vaccine design. Other than its role in cancer migration and proliferation, its role in chemo resistance (Liu *et al*., 2009) and radio resistance (Zhao *et al*., 2012) signifies that immunotherapy in combination with chemotherapy and radiotherapy can produce better

results for varied cancer treatments. There are less evidences of its expression and role in normal tissue which is favourable to the fact of its candidature for vaccine research. Also to avoid any chance of auto immune response and eliminating the chances of tolerance by the predicted peptides BLAST analysis is of great importance. Thus BLAST screening eliminated the 6 class I and II MHC specific epitopes hence ruled out any chance of auto immune response.

Molecular docking is extensively used for computer aided drug design but in recent years it is also being used for peptide MHC interaction study. Four nonamer peptides which are common to both CD4+ and CD8+ T cell epitopes were considered for docking studies because core binding region of peptide to MHC molecule is nonamer. These peptides were found to have similar binding energy with natural bound peptides for both class I and II MHC. Interestingly, FLLGYGWAA peptide has more average binding energy in compare to natural bound peptide for both class I and II MHC. Hence, these peptides will be strongly presented by MHC molecules present on antigen presenting cells to induce T cell response. Similar kind of studies has been reported for nucleoprotein of H5N1 virus where they compare the binding energy with predicted epitope and bound peptides for only class I MHC molecules (Hou *et al.*, 2012). In contrast to this report, we have done docking studies for both class I and II MHC. Peptide MHC docking studies has been also reported for different tumor antigen like melanoma antigen E and carcino embryonic antigen (CEA) (Akiyama *et al.*, 2012, Nakamura *et al.*, 2011).

Another additive advantage of computational epitope prediction is that the epitope predictions are made specific to MHC allele. Thus it enabled us to select those putative epitopes which are recognized by large number of MHC class I and class II allele. The population coverage analysis is the result of it that on an average approximately 65% of the global population is expected to show the MHC class I specific immune response where as it is approximately 98% in case of the class II. Hence the approach is making the prediction process more promising for the vaccine design covering multiple parameters for each predicted epitope. Thus applying bioinformatics approach for peptide based vaccine design in more precise, rational and time saving. The major limitation in current cancer treatments is tackling the metastasis and involvement of metadherin in it may be helpful in this direction. Targeting it for vaccine design can give a viable alternative to overcome the limitations of current vaccine strategies for cancer.

B cell epitope identification gives an additional advantages as it will give the idea of peptide having potential to induce antibody mediated immune response. B cell epitope can be discontinuous and continuous or linear. Prediction of discontinuous epitope is complicated as it is based on the three dimensional structure of protein (Ponomarenko *et al* 2009). Structure of metadherin protein is not resolved experimentally hence linear epitope prediction was carried out. Kolaskar and Tongaonkar method were used because it involves multiple factors such as hydrophobicity, accessibility, flexibility and antigenic propensity to predict the B cell epitope and its accuracy is 75% (Ponomarenko *et al* 2009). Interestingly, two peptide fragments were generated which contains both T and B cell epitopes (Table 7). Thus these peptides have the potential to induce both cellular and humoral immune response which suggest to be effective vaccine candidate

As there is not much evidences in clinical studies where efficient results with vaccine against cancer has been observed. But a combinatorial approach for the cancer immunotherapy *i.e.* larger peptide fragments containing both class I and class II specific immunogenic peptides, peptide vaccines with cytokines and agonist antibodies as co stimulatory molecules can be more efficient in producing good results in generating immune response (Arens *et al*., 2012). In peptide based cancer vaccine development more advancement has been made against melanoma. Recent studies with melanoma had demonstrated that gp100 based peptide vaccines along with interleukin-2 in phase III clinical trials had given promising results in generating immune response in patients with advanced stages of disease (Schwartzentruber *et al*., 2011). The ability of the peptide to stimulate the growth of lymphocytes signifies their immunogenic potential. As in our study a small surge in the lymphocyte count was observed against the FLLGYGWAA which indicates the immunogenic potential of the predicted peptides. It also validates the use multiple different epitope prediction algorithms for the peptide based vaccine design. Hence the immunoinformatics tools for the vaccine design can act as time and resources saver to select the lead peptides which can accelerate the process vaccine development.

Thus, it can be concluded that the multiple roles of metadherin especially metastasis in varied cancers provide us with a suitable option for vaccine design against cancer. The predicted peptides in our study if show positive results *in vivo* further studies may provide with a valuable candidate for multiple cancer vaccine. At the same time combination of vaccine based immunotherapy along with other treatments may produce more amusing outcomes in cancer treatment.

7. Chapter. SUMMARY

The role of metadherin in cancer progression and metastasis is well reported. It is one of the most widely over expressed protein in different cancers, which can be used for diagnosis and treatment. In our present study it was attempted to find out the immunogenic peptides of metadherin protein. The combination of *in vitro* and *in silico* approaches was applied to find out the potential immunogenic peptides for cancer vaccine. Six different epitope prediction tools were used for class I and II specific epitope prediction. Three peptide fragments containing both CD4+ and CD8+ specific T-cell epitopes were identified. There were four nonameric epitopes found amongst these predicted peptide fragments which were having affinity towards both class I and II MHC. To find out the immunogenic potential of these four epitopes, their docking was done with 22 and 6 MHC class I and II molecules respectively using auto dock vina 4.0. All the four peptides were found to show binding with 5 of 6 MHC class II molecules whereas the M1 and M2 showed no binding with the HLA DP2. Similarly all the four peptides showed binding to 17 out of 22 class I molecules alleles. Whereas none amongst four showed binding with HLA A201 and HLA B4103 was bound to only M3. HLA B1402, B4405 and Cw3 showed no binding with M1, M3 and M2 respectively. The binding energy of all the MHC binding peptides were found to be comparable with the docked naturally bound ligand. It signifies the good binding affinity of the peptides with MHC molecules. When the linear B-cell epitope prediction was two of 21 peptide fragment showed the similarity with the CD8+ and CD4+ enriched peptide fragment. The average population coverage for immunogenic response of predicted peptides were found out to be 64.46 % and 97.15 % for class I and Class II MHC specific respectively To further validate the claim of these four nonamers to be immunogenic, *in vitro* study was carried out for peptide M3 to find out its capability to stimulate peripheral blood mononuclear cell (PBMCs) using MTT assay. The peptide was found to stimulate the PBMCs in a 6 days culture at the concentration of 10µg/ml, which further validate the claim of immunogenic peptide identification of metadherin protein. Hence we can say that the combination of different computational and *in vitro* approaches can help in reducing the vaccine design process.

Chapter 8. REFERENCES

[1] Akiyama Y, Komiyama M, Nakamura Y, Iizuka A, Oshita C, Kume A, et al. Identification of novel MAGE-A6- and MAGE-A12-derived HLA-A24-restricted cytotoxic T lymphocyte epitopes using an in silico peptide-docking assay. Cancer Immunol Immunother. 2012; 61(12):2311-9.

[2] Altschul SF, Gish W, Miller W, Myers EW, and Lipman D. Basic local alignment search tool. J Mol Biol 1990; 215:403–10.

[3] Anderson P. H, Neilsen M, Lund O. Prediction of residues in discontinuous B-cell epitopes using protein 3D structure. Protein Science 2006; 15:2558-67.

[4] Britt DE, Yang DF, Yang DQ, Flanagan D, Callanan H, Lim YP et al. Identification of a novel protein, LYRIC, localized to tight junctions of polarized epithelial cells. Experimental Cell Research 2004; 300:134– 48.

[5] Brown DM, Ruoslahti E. Metadherin, a cell surface protein in breast tumors that mediates lung metastasis. Cancer Cell 2004; 5: 365-74.

[6] Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting population coverage of T-cell epitope based diagnostics and vaccines. BMC Bioinformatics 2006; 7: 153.

[7] Chen X, Dong K, Long M, Lin F, Wang X, Wei J et al. Serum anti-AEG-1 auto-antibody is a potential novel biomarker for malignant tumors. Oncology Letters 2012; 4:319-23.

[8] Cormier JN, Salgaller ML, Prevette T, Barracchini KC, Rivoltini L, Restifo N, et al. Enhancement of cellular immunity in melanoma patients immunized with peptides from MART-1 Melan A. Cancer Journal from Scientific American 1997; 3:37–44

[9] Davies M N, Flower D R. Harnessing the bioinformatics to discover new vaccines. Drug Discovery Today 2007; 12 (9-10):389-395.

[10] Disis ML, Gooley TA, Rinn K, Davis D, Piepkorn M, Cheever MA. Generation of T-Cell Immunity to the HER-2/neu Protein after Active Immunization with HER-2/neu Peptide–Based Vaccines. Journal of Clinical Oncology 2002; 20:2624-32.

[11] Emdad L, Das SK, Dasgupta S, Hu B, Sarkar D, Fisher PB. AEG-1/MTDH/LYRIC: Signaling Pathways, Downstream Genes, Interacting Proteins, and Regulation of Tumor Angiogenesis. Advances in Cancer Research. 2013; 120:75–11.

[12] Emdad L, Sarkar D, Su ZZ, Randolph A, Boukerche H, Valerie K, et al. Activation of the Nuclear Factor KB Pathway by Astrocyte Elevated Gene-1: Implications for Tumor Progression and Metastasis. Cancer Research 2006; 66:1509-16

[13] Emens Leisha A. Cancer vaccines: on the threshold of success. Expert Opinion on Emerging Drugs 2008; 13(2):295–8.

[14] Gjesrtan MK, Bjornheim J, Saeterdel I, Myklebust J, Gaudernack G. Cytotoxic CD4+ CD8+ T lymphocytes, generated by mutant p21-ras (12VAL) peptide vaccination of a patient, recognize 12VAL-dependent nested epitopes present within the vaccine peptide and kill autologous tumor cells carrying this mutation. International Journal of Cancer 1997; 72:784–90.

[15] Gnosa S, Shen YM, Wang CJ, Zhang H, Stratmann J, Arbman G. Expression of AEG-1 mRNA and protein in colorectal cancer patients and colon cancer cell lines. Journal of Translational Medicine 2012; 10:109.

[16] Gritzapis AG, Fridman A, Perez SA, Monica NL, Papamichail M, Aurisicchio L, et al. HER-2/neu (657-665) represents an immunogenic epitope of HER-2/neu oncoprotein with potent antitumor properties. Vaccine 2010; 28:162–70.

[17] Guo C, Manjili MH, Subjeck JR, Sarkar D, Fisher PB, Wang XY. Therapeutic Cancer Vaccines: Past, Present and Future. Advances in Cancer Research. 2013; 119:421–75.

[18] Hou Y, Guo Y, Wu C, Shen N, Jiang Y, Wang J. Prediction and Identification of T Cell Epitopes in the H5N1 Influenza Virus Nucleoprotein in Chicken. PLoS ONE 2012; 7(6): e39344.

[19] Hu G, Wei Y, Kang Y. The multifaceted role of MTDH/AEG-1 in cancer progression. Clinical Cancer Research 2009; 15(18):5615–20.

[20] Hui BAY, Bruce JP, Alajez NM, Shi W, Yue S, Ordonez BP et al. Significance of Dysregulated Metadherin and MicroRNA-375 in Head and Neck Cancer. Clinical Cancer Research *2011*; 17:7539-7550.

[21] Jagger E, Gnjatic S, Nagata Y, Stokert E, Jager D, karbach J, et al. Induction of primary NY-ESO-1 immunity: CD8+ T lymphocyte and antibody responses in peptide-vaccinated patients with NY-ESO-1+ cancers ” Proceedings of National Academy of Science 2000; 97:12198–203.

[22] Kang DC, Su ZZ, Sarkar D, Emdad L, Volsky DJ. Cloning and characterization of HIV-1-inducible astrocyte elevated gene-1, AEG-1. Gene 2005; 353:8 – 15

[23] Ke ZF, He S, Li S, Luo D, Feng C, Zhou W. Expression characteristics of astrocyte elevated gene-1 (AEG-1) in tongue carcinoma and its correlation with poor prognosis. Cancer Epidemiology 2013; 37:179–185

[24] Kikuno N, Shiina H, Urakami S, Kawamoto K, Hirata H, Tanaka Y, et al. Knockdown of astrocyte-elevated gene-1 inhibits prostate cancer progression through up regulation of FOXO3a activity. Oncogene 2007; 26:7647–55

[25] Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Letters 1990; 276:172-74.

[26] Kulkarni K, Bhosle S, Kolaskar AS. CEP: a conformational epitope prediction server. Nucleic acid research 2005; 33:168-171.

[27] Larsen MV, Lundegaard C, Lamberth K, Buus S, Lund O, Nielsen M. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. BMC Bioinformatics 2007; 8:424.

[28] Larson J.E.P, Lund O, Neilson M. Improved methods for predicting the B-cell epitope. Immune research 2006; 2:2.

[29] Lee P, Wang F, Kuniyoski J, Rubio V, Stuges T, Groshen S, et al. Effect of interleukin-12 on immune response to multi peptide vaccine for restricted metastatic melanoma. Journal of Clinical Oncology 2001; 19:3836–47.

[30] Liao WT, Guo L, Zhong Y, Wu YH, Li J , Song LB. Astrocyte elevated gene-1 (AEG-1) is a marker for aggressive salivary gland carcinoma. Journal of Translational Medicine 2011; 9:205

[31] Liu H, Song X, Liu C, Xie L, Wei L, Sun R. Knockdown of astrocyte elevated gene-1 inhibits proliferation and enhancing chemo-sensitivity to cisplatin or doxorubicin in neuroblastoma cells. Journal of Experimental & Clinical Cancer Research 2009; 28:19-8.

[32] Lohia N, Baranwal M. Conserved Peptides Containing Overlapping CD4 + and CD8 + T-Cell Epitopes in the H1N1 Influenza Virus: An Immunoinformatics Approach. Viral Immunology 2014; 27(5):225-34

[33] Marchand M, van Baren N, Weynants P, Brichard V, Dreno B, Tesssier MH, et al. "T regressions observed in patients with metastatic melanoma treated with an antigenic peptide encoded by gene MAGE-3 and presented by HLA-A1" International Journal of Cancer 1999; 80:219–30.

[34] Mizukoshi E, Nakamoto E, Arai K, Yamashita T, Sakai A, Sakai Y, et al. Comparative analysis of various tumor-associated antigen-specific t-cell responses in patients with hepatocellular carcinoma. Hepatology 2011; 53:1206-16.

[35] Muderspach L, Wilczy k S, Roman L, Bade L, Felix J, Small LA, et al. "A phase l trial of human papillomavirus (HPV) peptide vaccine for women with high-grade cervical and vulvar intraepithelial neoplasia who are HPV 16 positive. Clinical Cancer Research 2000; 6:3406–16.

[36] Nakamura Y, Tai S, Oshita C, Iizuka A, Ashizawa T, Saito S, et al. Analysis of HLA-A24-restricted peptides of carcinoembryonic antigen using a novel structure-based peptide-HLA docking algorithm. Cancer Science 2011; 102(4):690-6.

[37] Odunsi K, Qian F, Matsuzaki J, Fauceglia PM, Andrews C, Hoffman EW. Vaccination with an NY-ESO-1 peptide of HLA class I/II specificities induces integrated humoral and T cell responses in ovarian cancer. PNAS 2007; 104(31):12837–42

[38] Parker KC, Bednarek MA, Coligan JE. Scheme for Ranking Potential HLA-A2 Binding Peptides Based on Independent Binding of Individual Peptide Side Chains. Journal of Immunology 1994; 152:163-75.

[39] Ponomarenko JV, van Regenmortel MHV. B-Cell Epitope Prediction. In: Gu J, Bourne PE, editors. Structural Bioinformatics, second edition: John Wiley & Sons, Inc; 2009, p 849-79.

[40] Qian BJ, Yan F, Li N, Liu QL, Lin YH, Liu CM, et al. MTDH/AEG-1-based DNA vaccine suppresses lung metastasis and enhances chemo sensitivity to doxorubicin in breast cancer. Cancer Immunology and Immunotherapy 2011; 60:883–93

[41] Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanovic´ S. SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics 1999; 50: 213–19.

[42] Raval RR, Sharabi AB, Walker AJ, Drake CG, Sharma P. Tumor immunology and cancer immunotherapy: summary of the 2013 SITC primer. Journal for Immunotherapy of Cancer 2014; 2:14.

[43] Reche PA, Reinherz EL .Prediction of peptide-MHC binding using profiles. Methods in Molecular Biology. 2007; 409:185-00.

[44] Rosenberg SA, Yang JC, Schwartzentruber DJ, Hwu P, Marincola FM, Topalian SL, et al. Immunologic and therapeutic evaluation of synthetic peptide vaccine for the patients with metastatic melanoma. Nature Medicine 1998; 4:21–7

[45] Rosenberg SA, Yang JC, Schwartzentruber DJ, Hwu P, Marincola FM, Topalian SL, et al. Impact of cytokine administration on generation of antitumor reactivity in patients with metastatic melanoma receiving peptide vaccine. Journal of Immunology 1999; 163: 1690–5.

[46] Saha S, Raghava GPS. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Protein 2006; 65:40-8.

[47] Sarkar D, Park SE ,Emdad L, Lee SG, Su ZZ, Fisher PB. Molecular Basis of Nuclear Factor-KB Activation by Astrocyte Elevated Gene-1. Cancer Research. 2008; 68:1478-84.

[48] Scheibenboge C, Schmittal A, Keilholz U, Allaguer T, Hofman U, Max R, et al. Phase 2 trial of vaccination with tyrosinase peptides and granulocyte-macrophage colony stimulating factor in patients with metastatic melanoma. Journal of Immunotherapy. 2000; 23:275–81.

[49] Schwartzentruber DJ, Lawson DH, Richards JM, Conry RM, Miller DM, Treisman J, et al. gp100 Peptide Vaccine and Interleukin-2 in Patients with Advanced Melanoma. The New England journal of medicine 2011; 364:2119-27.

[50] Singh H, and Raghava GP. ProPred: prediction of HLA-DR binding sites. Bioinformatics 2001; 17:1236–37.

[51] Su P, Zhang Q, Yang Q. Immunohistochemical analysis of Metadherin in proliferative and cancerous breast tissue. Diagnostic Pathology 2010; **5**:38.

[52] Thevenet P, Shen Y, Maupetit J, Guyon F, Derreumaux P, Tuffery P. Nucleic Acids Research, 2012 ; 40:288–93

[53] Thirkettle HJ, Girling J, Warren AY, Mills IG, Sahadevan K, Leung H. LYRIC/AEG-1Is Targeted to Different Subcellular Compartments by Ubiquitinylation and Intrinsic Nuclear Localization Signals. Clinical Cancer Research 2009; 15:3003-13.

[54] Tiriveedhi V, Sarma NJ, Subramanian B, Fleming T P, Gillanders WE, Mohanakumar T. Identification of HLA-A24-restricted CD8cytotoxic T-cell epitopes derived from mammaglobin-A, a human breast cancer–associated antigen. Human Immunology 2012; 73:11-6.

[55] Trott O, Olson AJ. Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. Journal of Computational Chemistry 2010; 31:455–61.

[56] Vigneron N, Stroobant V, Van den Eynde BJ, van der Bruggen P. Database of T cell-defined human tumor antigens: the 2013 update. Cancer Immunity 2013; 13:15-20.

[57] Wang F, Bade E, Kuniyoski c, Spears L, Jeffery G, Marty V et al. peptide vaccine with incomplete Freund's adjuvant for resected high risk melanoma. Clinical Cancer Research 1999; 5:2756–65

[58] Wang P, Sidney J, Dow C, Mothe´ B, Sette A, Peters B. A Systematic Assessment of MHC Class II Peptide Binding Predictions and Evaluation of a Consensus Approach. PLoS Computational Biology 2008; 4(4) : e1000048

[59] Wang YP, Liu IJ, Chiang CP, Wu HC. Astrocyte elevated gene-1 is associated with metastasis in head and neck squamous cell carcinoma through p65 phosphorylation and up regulation of MMP1. Molecular Cancer 2013;12:109

[60] Yoo BK, Emdad L, Lee SG, Su ZZ, Santhekadur P, Chen D, et al. Astrocyte Elevated Gene-1 (AEG-1): a multifunctional regulator of normal and abnormal physiology. Pharmacology and Therapeutics. 2011; 130(1):1–8.

[61] Yuan B, Zhao L, Xian R, Zhao G. Identification of novel HLA-A⁄0201-restricted CTL Epitopes from Pokémon. Cellular Immunology 2012; 274:54-60.

[62] Zhao Y, Moran M.S, Yang Q, Liu Q, Yuan C, Hong S et al. Metadherin regulates radio resistance in cervical cancer cells. Oncology Reports 2012; 27:1520-26.

[63] Zhu K, Dai Z, Pan Q, Wang Z, Yang GH, Yu L. Metadherin Promotes Hepatocellular Carcinoma Metastasis through Induction of Epithelial–Mesenchymal Transition. Clinical Cancer Research 2011; 17:7294-02.

[64] Sela M, Arnon R, Schechter B. Therapeutic vaccines: realities of today and hopes for the future. Therapeutic Focus 2002;7 (12) : 664-673

[65] Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 2008:9:40.