Echinococcosis disease is a cosmopolitan zoonosis caused by metacestode stages of tapeworms of the genus Echinococcus (family Taeniidae). Within this genus, four species, *Echinococcus granulosus, E. multilocularis, E. vogeli*, and *E. oligarthrus*, are recognized which all may establish and develop in the human host. Among them, the clinically most relevant species are *E. granulosus* and *E. multilocularis* which are responsible for cystic echinococcosis (CE) and alveolar echinococcosis (AE), respectively. CE is recognized as one of the world’s major zoonosis (McManus and Smyth, 1986). The adult worm lives in the small intestine of a carnivore (definitive host), and the intermediate larval stage can infect a wide range of mammalian species - including humans - that acquire the infection through accidental ingestion of eggs. The domestic life cycle is maintained through dogs and ungulates, mainly sheep and cattle. The disease constitutes an emerging public health problem, with a considerable impact in both human and animal health and causing important socio-economic consequences in endemic areas (Schantz et al., 1995), especially in regions with extensive livestock husbandry and non-supervised slaughter (Romig, 2003). The annual incidence varies from 1-200/10^5 persons.

The incidence of human CE is closely related to the prevalence of the disease in domestic animals. Humans become exposed to eggs of tapeworm after close contact with dog or its contaminated environment. Typically larvae that pass through the liver are trapped in pulmonary arterial capillaries and develop into hydatid cysts. Chronic hydatid or pulmonale can occur due to obstruction of pulmonary artery. Secondary bronchogenic hydatidosis causes disseminated lesions, which can be lead to hydatidoptysis – vomiting of small entire hydatids (Gottstein et al., 2006).

Most human patients harbor a solitary cyst within a single organ. In 70% of patients, infection occurs in the liver, the lung being affected in 20% of patients. Other sites of infection include kidney, spleen, muscle, skin, abdominal and pelvic cavity, mediastinum, heart, brain, spinal cord and others (Kern, 2003). Within fertile metacestodes, brood capsules are formed out of distinct cells originating from the germinal layer, which themselves contain protoscolces. Infertile metacestodes only contain the germinal layer. Often the metacestode contains one or more daughter cysts, which give rise to the septet appearance in diagnostic imaging (Von Sinner, 1991).
minimum time required for the development of protoscolesces in cysts in humans is not exactly known, but based on data from animals, it is expected to be 10 months or longer after infection (Ammann et al., 1995; Pawlowski, 1997). Protoscolesces can be already formed in small cysts of 0.5 cm-2.0 cm diameter. In the same patient, fertile (with protoscolesces) and sterile (without protoscolesces) cysts may coexist. The cysts slowly expand over several years, often without being noticed. Spread of infection into other sites, such as the peritoneal cavity, has been described, and this can be caused by cyst rupture and spillage of protoscolesces. These have the potential to develop into metacestodes in the intermediate host, and into adult parasites when ingested by the final host (Siles-Lucas and Hemphill, 2002).
LIFE CYCLE OF ECHINOCOCCUS GRANULOSUS

1. Embryonated eggs expelled in canid feces...
2. Ungulates like deer, elk or sheep ingest eggs as they browse...
3. Ocosnphore hatches, penetrates intestine/numen wall...
4. Ocosnphore travels in bloodstream to become lodged in capillaries beds...
5. Hydatid cyst is formed, usually in liver or lungs...
6. Canids like wolves, coyotes or dogs eat hydatid cysts in organs...
7. Adult tapeworm in small intestine produces eggs...

Definitive Host

Scolex attaches to intestine, grows into tapeworm...

Domestic dogs are known, and very likely, the most important source of this parasite to human populations worldwide...

Humans can ingest eggs by failing to follow sanitary hygienic practices like wearing protective barriers like gloves and washing their hands well with soap and water after possible exposure to canid feces...

Figure 2.1: Life Cycle of E. granulosus (Source: WHO 2002)
2.1. EPIDEMIOLOGY OF CYSTIC ECHINOCOCCOSIS

*Echinococcus granulosus* has a world-wide geographic distribution and occurs in almost all parts of the world with only a very few countries regarded being free of CE (Craig *et al.*, 1996, Schantz *et al.*, 1995). The highest prevalence of the parasite is found in parts of Eurasia, Russian Federation, Africa, China, Australia and South America (Fig. 2.3.).

![Figure 2.2: Approximate geographic distribution of *Echinococcus granulosus* (1999)
(Source: Eckert and Deplazes, 2004)](image)

2.1.1. *Echinococcus granulosus* strains distribution

*Echinococcus granulosus* is a complex of species/genotypes which differ in terms of life-cycle patterns and hosts. Molecular methods have allowed discrimination of different genotypes (G1-10 and the 'lion strain'), some of which are now considered separate species. *Echinococcus* has been divided into two genotypic clusters G1-3 (*E. granulosus sensu stricto*) and G6-10 (*Echinococcus canadensis*). Based on mitochondrial DNA (mtDNA) analysis, the *E. granulosus* complex has been split into *E. granulosus*
sensu stricto (genotypes G1, G2 and G3), *E. equinus* (G4), *E. ortleppi* (G5) and the still controversial taxon *E. canadensis* (G6-G10) (Alvarez et al., 2014).

**G1 strain:** It is the most frequent strain associated with human CE. This strain appears to be widely distributed in all continents. The communities involved in extensive sheep farming are reported to possess higher rates of infection and many epidemiological studies suggest that this genetic variant is the principal strain infecting humans (Eckert *et al*., 2001, Thompson and McManus, 2001, 2002; Dinkel *et al*., 2004). Consequently, its presence coincides with areas which have high prevalence of human CE such as in Morocco, Tunisia, Kenya, Kazakhstan, western China and Argentina.

**G2 strain:** The G2 strain is known to be transmitted among sheep and infect humans also, but genetic differences biologically distinguish it from the G1 strain, conferring a different life cycle (McManus and Bryant, 1995). It has been found in Australia Tasmania.

**G3 strain:** The G3 strain is reported among buffalos has been recorded in South Asia (Macpherson and Wachira, 1997) and transmitted by water, but no susceptibility among humans has been found.

**G4 strain:** The G4 strain, formerly known as *Echinococcus equinus*, appears to infect exclusively equines as intermediate hosts and no human cases have been documented (Thompson *et al*., 1992; Thompson and McManus, 2001). It is known to be diffused in the Mediterranean regions of Spain, Italy, Lebanon, and Syria, as well as in South Africa.

**G5 strain:** The former cattle strain (G5), known as *Echinococcus ortleppi*, is transmitted by cattle in Europe, Asia, parts of Africa and South America and only one case in humans has been isolated in past years (Bowles *et al*., 1992), suggesting a less pathogenic risk for humans than the sheep strain of *E. granulosus*.

**G6 strain:** The G6 strain is known to principally affect camels and goats. Animal infection is diffused in the Middle East, Africa, southern Asia and South America (Thompson *et al*., 2001) and cases of human infection have been found in Nepal, Iran, Mauritania, Kenya and Argentina (Thompson *et al*., 2002; Dinkel *et al*., 2004).

**G7 strain:** The G7 strain is transmitted by domestic pigs in Europe (Spain and Italy), Asia and South America, as well as the closely related genetic variant G9 that has been
documented to affect Polish patients (Scott et al., 1997) although the animal reservoir is unknown.

**G8 strain:** The G8 strains are known to be transmitted between wolves and wild cervids in the northern regions of Europe, Asia and North America. Few cases of human infection have been documented with a lower severity of the disease than CE caused by other forms of *E. granulosus* (Wilson et al., 1968). However, transmission between humans of this genetic variant seems to be low and further data is needed to better assess its pathogenicity.

### 2.1.2. Worldwide distribution

**America:** The cervid strain (G8) and the sheep strain (G1) of *E. granulosus* is reported more in North America. The cervid strain is diffused in wildlife mainly in Canada, Alaska and Minnesota (Peterson, 1997). The wildlife reservoir was found to be largely diffused among cervids and wolves, coyotes and domestic dogs (Webster and Cameron, 1967). Sporadic autochthonous transmission among humans of the sheep strain in the western States of North America such as Arizona, California, New Mexico and Utah has been documented in reports from the 1960s (Schantz et al., 1995). The source of these *E. granulosus* infections was Australian sheep dogs imported into Utah in 1938 when the parasite diffused among sheep of this area as well as adjoining states through trading of live sheep (Crellin et al., 1982).

In the United States, all genetic variants of the *E. granulosus* complex have been introduced into South America with domestic animals imported from other regions, such as Europe. The principal strain of *E. granulosus* is the sheep strain (G1), widely diffused in Peru, Chile, Argentina and Brazil (Eckert et al., 2001).

Chile is an endemic area for *E. granulosus* infection. According to Panamerican Health Organization report (2000), the prevalence of bovine, sheep and canine hydatidosis for the entire country decreased to 22.3%, 6.3% and 11%, respectively (Alvarez et al., 2002; Serra et al., 1996). A major endemic area for CE is the southern part of Chile where the population residing in this area of Chile are on risk (Apt et al., 2000).
**Australia:** The most common strain currently found in Australia is the G1, while the G2 strain was previously also found in Tasmania (Pearson *et al.*, 2002, McManus and Thompson, 2003, Jenkins, 2005). This G2 strain probably evolved as a genetically modified variant after a Tasmanian hydatid control campaign aimed to strictly control helminthic diffusion among dogs. Thus, this genetic variant became dominant because of the limited gene pool on an island (McManus and Bryant, 1995). In Australia, several areas have been documented at high risk of transmission of *E. granulosus*. The definitive hosts most commonly involved in transmission of CE in south eastern Australia are wild dog (Grainger and Jenkins, 1996), while the most common intermediate hosts are grey kangaroos and wallabies (Jenkins and Morris, 2003).

**Western and Central Asia:** The G1 strain, infecting sheep, goats, cattle and camels, is the most common genetic variant documented in Iran. However, the G6 strain has also been reported from camels, sheep and cattle. Both these have also been reported in humans (Harandi *et al.*, 2002).

Hydatidosis is a serious public health problem in Turkey where *E. granulosus* infection in dogs ranges between 0.32% and 40% (Altintas, 2003). The predominant genotype of *E. granulosus* in Turkey is the G1 strain with a prevalence infection rate in farm animals ranging from 26.6% to 50.9% in sheep, from 13.3% to 35.68% in cattle, and reaching 22.1% in goats, 44.31% in cows and 24.39% in bulls in the most endemic areas.

**India:** According to WHO report (2002), *E. granulosus* has a wide geographic distribution, as indicated by reports on parasite prevalence in livestock from different parts of the country, High prevalence of Hydatidosis has been reported from Kashmir, Andhra Pradesh, Tamil Nadu and Central India (Nepalia *et al.*, 2006; Craig *et al.*, 2007; Mandal *et al.*, 2011; Sharma *et al.*, 2013). A study in southern Indian states like Karnataka, Maharashtra, Kerala, Tamil Nadu, Pondicherry and Goa revealed overall prevalences of Hydatid cysts 7.0% in sheep, 7.1% in cattle, 9.4% in water buffalo, and 11.5% in pigs (Hafeez, 1997). Another study carried out from 1995 to 1997 in Pondicherry reported higher prevalence rates of 37.8% in sheep and of 47.6% in goats (Das & Sreekrishnan, 1998). While in Uttar Pradesh, the prevalence of CE in sheep were
2.9%, in goats 1.4%, and in pigs 0.9% (Deka & Gaur, 1998). All these studies indicate that CE is endemic in many states of India.

**China:** China is one of the most important endemic regions of CE (Eckert et al., 2002). The sheep strain (G1) and the camel strain (G6) are the only two strains of *E. granulosus* found in China (Zhang et al., 2005), both of them infectious to human beings (Bart et al., 2006). The hyper-endemic areas for *Echinococcus* spp. have been recognized as the provinces and autonomous regions stretching from western Xinjiang, Ningxia and Inner Mongolia, with the highest prevalence rates occurring in pastoral communities of the eastern Tibetan plateau (Tiaoying et al., 2005, Chai, 1995) (south western Qinghai and north western Sichuan) and the Tibetan autonomous area of south Gansu (Craig et al., 2004), located in western and northwestern China (Bart et al., 2006, Giraudoux et al., 2006, Yang et al., 2005; Tang et al., 2006). The definitive host is mainly represented by canids, predominantly the domestic dog. Indeed, they are kept in large populations in northwestern China for pastoralism and cultural reasons (Tiaoying et al., 2005).

**Africa:** The limited information is available about the prevalence of CE in Africa due to lack of epidemiological studies (Ibrahem and Gusbi, 1997). The most common strain prevalent is the G1, highly diffused in the North and East African sheep raising areas. Moreover, the exclusive presence of the camel strain (G6) has been documented. In addition, wild strains such as the *E. equinus* (the “horse strain”) (Kumaratilake et al., 1986) and the “lion strain” (Verstr et al., 1965) have been found in South Africa. However, the nature of *Echinococcus* in African wildlife is poorly documented.

**Europe and the Mediterranean Basin:** *E. granulosus* is present in most countries of Europe with the exception of Ireland, Iceland and Denmark.

The Mediterranean regions of European countries are hyper-endemic region of CE where annual incidence rates for human CE ranged from 4-8 per 100 000 (Eckert et al., 2001).

### 2.2. Genotyping of *E. granulosus*

Genotyping is useful tool to understand the transmission patterns of the parasite for epidemiological purposes and the human susceptibility to a particular genotype of *E. granulosus*. Various molecular techniques have been applied for the genotyping of
Echinococcus species like Restriction fragment length polymorphism (RFLP) of ribosomal DNA (Bowles & McManus, 1993; Siles-Lucas et al., 1994), PCR of the mitochondrial cox-1 gene, nd-1 genes (Bowles et al., 1992, Bowles & McManus, 1993, Zhang et al., 1998), RAPD-PCR (Scott & McManus, 1994, Siles-Lucas et al., 1994), and SSCP (Gasser et al., 1998).

E. granulosus is a complex of species/strains which exhibits diversity in their life cycle patterns and host range (Thompson et al., 2002). To date, 10 genotypes of E. granulosus have been identified by molecular genetic analysis using gene specific PCR of mtDNA sequences (Bowles et al., 1992; Scott et al., 1997; Lavikainen et al., 2003) (Table 2.1).

2.2.1. Cytochrome oxidase subunit-1

Cytochrome c oxidase I (COX1) also known as mitochondrially encoded cytochrome c oxidase I (MT-CO1) is an enzyme encoded by mt-co1 gene. In other eukaryotes, the gene is called COX1, CO1, or COI (Kosakyan et al., 2012). Cytochrome c oxidase I is the main subunit of the cytochrome c oxidase complex.

The gene has been widely used for genotyping of E. granulosus (Jabbar et al., 2011; Yanagida et al., 2012; Piccoli et al., 2012; Sharma et al., 2013). The gene is often used as a DNA barcode to identify animal species because its mutation rate is often fast enough to distinguish closely related species and also because its sequence is conserved among conspecifics. (Hebert et al., 2012).

2.2.2. NADH dehydrogenase 1 (nd-1)

NADH dehydrogenase (ubiquinone) is an enzyme of the respiratory chains of myriad organisms from bacteria to humans, located in the inner mitochondrial membrane. It catalyzes the transfer of electrons from NADH to coenzyme Q10 (CoQ10). It is one of the "entry enzymes" of cellular respiration or oxidative phosphorylation in the mitochondria (Nakamaru-Ogiso et al., 2012). The gene remains conserved among conspecifics, thus is widely used in genotyping of different organisms. The gene has already been widely used for genotyping of E. granulosus by different workers (Jabbar et al., 2011; Piccoli et al., 2012; Sanchez et al., 2012).
2.3. IN-SILICO STUDY IN IMMUNOLOGY

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 computation immunology, and a number of immunology-focused resources and software, which help in understanding the properties of the whole immune system, have become available (Brusic et al., 2005). This has given rise to a new field, called immunoinformatics.

Immunoinformatics can be described as a branch of bioinformatics concerned with in silico analysis and modeling 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 analyze the sequence areas with potential binding sites, which in turn leads to the development of new vaccines. The methodology of analyzing the pathogen genome to identify potential antigenic proteins is known as ‘reverse vaccinology’ (Tomar et al, 2010). This is mainly beneficial because 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 revealing virulence genes and surface-associated proteins.

2.4. B cell epitopes

Antibodies are proteins produced by B-cells in response to immunogenic substances such as viruses, allergens, and vaccines. B-cell, or antibody, epitopes are the molecular structures within an antigen that make specific contacts with residues of soluble and membrane-bound antibody molecules. For various immunological applications, a computational prediction of epitopes in an antigen is highly desirable. Many such tools exist that mainly base their predictions on shared amino acid characteristics of known B-cell epitopes in protein antigens. The identification and characterization of B-cell epitopes play an important role in vaccine design, immunodiagnostic tests, and antibody production.

The main objective of epitope prediction is to design a molecule that can replace an antigen in the process of either antibody production or antibody detection. Such a molecule can be synthesized or, in case of a protein, its gene can be cloned into an expression vector. Designed molecules are preferable to use because they are inexpensive and noninfectious in contrast to viruses or bacteria, which may be harmful to a researcher or experimental animal.

Diagnostic immunoassays utilize peptides or antipeptide antibodies for detection, isolation, and characterization of molecules associated with various disease stages. Synthetic peptides can be used for detection of antibodies produced as a result of infections, allergies, autoimmune diseases, or cancers. To be used as a diagnostic tool, a peptide should be antigenic, that is, able to bind a specific antibody. The prediction of such peptides is one of the goals of the epitope prediction methods. Another goal is the prediction of immunogenic peptides that can be used for production of antipeptide antibodies. Antibodies can detect proteins and various disease marker molecules, including viral proteins and bacterial lipopolysaccharides, present at the early stages of infections. Antipeptide antibodies used in diagnostic immunoassays can be obtained either in vivo by immunizing animals or in vitro by developing hybridoma cell lines, that
is, engineered cells designed to produce high volumes of antibodies, or using combinatorial libraries.

2.4.1. Types of B cell epitopes:

Conformational/Discontinuous B cell epitopes:

A conformational epitope is a sequence of sub-units (usually, amino acids) composing an antigen that come in direct contact with a receptor of the immune system. Proteins are composed of repeating nitrogen-containing subunits called amino acids that in nature do not exist as straight chains called primary structure, but as folded whorls with complex loops. The latter is known as the tertiary structure of a protein. So, whenever a receptor interacts with an undigested antigen, the surface of amino acids that come in contact may not be continuous with each other if the protein is unwound. Such discontinuous amino acids that come together in three-dimensional conformation and interact with the receptor's paratope are called conformational epitopes (Table 2.3).

Continuous/Linear B cell epitopes:

A linear or a sequential epitope is an epitope that is recognized by antibodies by its linear sequence of amino acids, or primary structure. In contrast, most antibodies recognize a conformational epitope that has a specific three-dimensional shape and its protein structure. The linear sequence of amino acids that compose a protein is called its primary structure, which does not exist in nature. But, when an antigen is broken down in a lysosome, it yields small peptides, which can be recognized through the amino acids that lie continuously in a line, and hence are called linear epitopes (Table 2.4).
Figure 2.3: Recognition of conformational epitopes by B cells. Note how the segments widely separated in the primary structure have come in contact in the three-dimensional tertiary structure forming part of the same epitope. (Source: Goldsby et al., 2003).
While performing molecular assays involving use of antibodies such as in the Western blot, immunohistochemistry, and ELISA, one should carefully choose antibodies that recognize linear or conformational epitopes.

For instance, if a protein sample is boiled, treated with beta-mercaptoethanol, and run in SDS-PAGE for the Western blot, the proteins are essentially denatured and therefore cannot assume their natural three-dimensional conformations. Therefore, antibodies that recognize linear epitopes instead of conformational epitopes are chosen.

Figure 2.4: Recognition of epitopes in a linear fashion. Note: the same (colored) segment of protein can be a part of more than one epitopes. (Source: Goldsby et al., 2003).
for immunodetection. In contrast, in immunohistochemistry where protein structure is preserved, antibodies are preferred that recognize conformational epitopes.

2.4.2. Programs/ Server used for B-Cell epitopes prediction

**ABCpred:**

ABCpred program is an artificial neutral network which used to predict B-cell epitopes present within protein sequences. ABCpred is a first server which was developed on the basis of artificial neutral network (machine based technique). It is using fixed length patterns for the prediction of epitopes (Saha and Raghava, 2006). The system was trained on 700 epitopes from the Bcipep database and 700 randomly selected peptides represented by amino acid sequences of lengths varying between 10 and 20 amino acids. Employing fivefold cross-validation on this dataset, the method achieved a maximum accuracy of 66%.

**BCPREDs**

The current implementation of BCPREDs allows the user to select among three prediction methods: (i) implementation of AAP method (Chen et al., 2007); (ii) BCPred (EL-Manzalawy et al., 2008); (iii) FBCPred (EL-Manzalawy et al., 2008b). In BCPREDs the classifier specificity was set at 80% and the epitope length was set 20.

**BepiPred 1.0**

BepiPred predicts the location of linear B-cell epitopes using a combination of a hidden Markov model and a propensity scale method (Larsen et al., 2006). The residues with scores above the threshold (default 0.35) are predicted to be part of an epitope. In the present study, the threshold for epitope assignment was set at 0.35 using BepiPred server.

**Ellipro**

Ellipro predicts linear and discontinuous antibody epitopes based on a protein antigen's 3D structure (Ponomarenkoet al., 2008). Ellipro first runs the MODELLER software to model the protein 3D structure via homology modeling. For the prediction of
epitopes, the prediction was made at a minimum level of 0.5 to the most stringent level of 1.0 and the maximum distance for residue clustering was kept at 6.0 Å. Selection of diagnostic proteins as B cell epitope candidates was based on the number of epitopes predicted with a minimum cutoff score of 0.8.

**Discotope-2.0**

Discotope server predicts discontinuous B cell epitopes from protein’s three dimensional structures. The method utilizes calculation of surface accessibility (estimated in terms of contact numbers) and a novel epitope propensity amino acid score. The final scores are calculated by combining the propensity scores of residues in spatial proximity and the contact numbers (Jens VK et al., 2012). This innovative algorithm combines scores from this matrix with a measure of the surface area to predict the location of discontinuous epitope residues and does so with a fair degree of accuracy (Greenbaum et al., 2007).

**2.5. T-cell epitopes**

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 (Huang and Honda, 2006). 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 (Madden, 1995). MHC class I molecules usually present peptides between 8 and 11 amino acids in length, whereas the peptides binding to MHC class II may have length from 12 to 25 amino acids (Jardetzky et al., 1996). 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 (Figure 2.5). 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 (Janeway, 2001).
The recognition of epitopes by T cells and the induction of immune response have a key role for the individual’s immune system. Even the slightest deviation from their normal functioning can have a grave impact on the organism. Knowledge about the peptide’s epitopes has a key role for manufacturing epitope-based vaccines, which, injected into the recipient, can induce 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. MHCs are among the most polymorphic proteins in higher vertebrates, with more than 6000 class I and class II MHC molecules listed in IMGT/HLA (Robinson et al., 2011). 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) (Liu et al., 2006, Wan et al., 2006), hidden Markov models (HMMs) (Zhang et al., 2006; Noguchi et al., 2002; Mamitsuka, 1998), quantitative structure–activity relationship (QSAR) analysis (Doytchinova and Flower, 2003; Doytchinova and Walshe, 2001), and structure-based approaches (Wan et al., 2005).
Figure 2.5: Antigen-processing pathways for MHC Class I and II molecules (Source: rsob.royalsocietypublishing.org)

**Left:** intracellular pathway. Protein is cleaved into oligopeptides in the proteasome, the peptides enter the endoplasmic reticulum (ER) via TAP protein and bind to MHC class I, and the complex of peptide–MHC protein so formed is presented to Cytotoxic T cell.

**Right:** extracellular pathway. Protein is endocytozed, cleaved into oligopeptides in the endosome, bound to MHC class II protein and presented on the cell surface. In the ER, MHC class II molecules are adjoined to a specific peptide, known as invariant chain (Ii). It blocks the binding cleft of the MHC molecule, thereby preventing the binding of endogenous peptides. In the endosome, the Ii is initially cleaved to CLIP peptide, and is
then replaced by an exogenous peptide. The process is facilitated by the HLA-DM molecule.

2.5.1. Servers used for T cell epitope prediction:

**NetMHCpan 2.4**

NetMHCpan server predicts binding of peptides to any known MHC molecule using artificial neural networks (ANNs) (Hoof et al., 2009). The method is trained on more than 115,000 quantitative binding data covering more than 120 different MHC molecules. Predictions can be made for HLA-A, B, C, E and G alleles, as well as for non-human primates, mouse, Cattle and pig. Predictions can be made for 8-11 mer peptides.

**NetMHC3.4**

NetMHC 3.4 server also predicts binding of peptides to a number of different HLA alleles using artificial neural networks (ANNs) (Nielsen et al., 2003). Predictions can be made for lengths between 8 and 14 for all alleles using a novel approximation algorithm using ANNs trained on 9mer peptides. Probably because of the limited amount of available 10mer data this method has a better predictive value than ANN strained on 10mer data.

**IEDB-NN**

IEDB-NN align are artificial neutral network–based alignment algorithm for the prediction of MHC class-I binding peptides. NN-align is trained using a novel training algorithm that allows for correction of unfairness in the training data due to redundant binding core representation. Incorporation of information about the residues flanking the peptide-binding core is shown to significantly improve the prediction accuracy. The method is evaluated on a large-scale benchmark consisting of six independent data sets covering 14 human MHC class II alleles, and is demonstrated to outperform other state-of-the-art MHC class II prediction methods (Nielsen et al., 2009).
IEDB-SMM

IEDB SMM-align is a freely available server utilizing a large set of quantitative MHC class II peptide binding data which has been made publicly available on the IEDB databases (Toseland et al., 2005). The data set contains peptide data with IC$_{50}$ binding affinities with more than 14 HLA (human MHC) and also with several mouse MHC class II alleles. SMM-align is a novel method for quantitative MHC class II binding predictions. The method is an extension of the stabilization matrix method proposed by Peters et al. (Swets, 1988). The SMM-align method seeks to identify a weight matrix that optimally reproduces the measured IC$_{50}$ values for each peptide in the training set. The method allows for identification the MHC class II binding motif in terms of a position specific weight matrix. The output of the SMM-align method is IC50 binding affinity values, enabling direct readout of the peptide: MHC binding affinity (Nielsen et al., 2007; Peters et al., 2005).

NetMHCIIpan 3.0

NetMHCIIpan 3.0 server predicts binding of peptides to MHC class II molecules (Karosiene et al., 2013). The method was evaluated in several benchmarks and demonstrates a significant improvement over molecule-specific methods as well as the ability to predict peptide binding of previously uncharacterized MHCII molecules. The predictions are available for all three human MHC class II isotypes, HLA-DR, HLA-DP and HLA-DQ, as well as mouse molecules. NetMHCIIpan 3.0 server can produce predictions for peptides of 9 - 19 amino acids in length along with the IC$_{50}$ value.

ProPred

The aim of this server is to predict MHC Class-II binding regions in an antigen sequence, using quantitative matrices derived from published literature by Sturniolo et al., (1999). The server assists in locating promiscuous binding regions that are useful in selecting vaccine candidates. This server is a useful tool in locating the promiscuous binding regions that can bind to a total of 51 alleles belonging to nine serologically defined HLA-DR molecules encoded by the DRB1 and DRB5 genes (Singh and Raghava, 2001).
**MultiPred2**

MULTIPRED2 is a computational system for large-scale screening of peptide binding to multiple alleles belonging to human leukocyte antigen (HLA) class I and class II DR supertypes as well as to alleles belonging to an individual's genotype. netMHCpan and netMHCIIpan are used as prediction engines (Zhang et al., 2011). In total MULTIPRED2 performs binding predictions on 1077 alleles belonging to 26 HLA supertypes. Built upon our previous developments - MULTIPRED and PEPVAC, MULTIPRED2 is for mapping of promiscuous T-cell epitopes as well as the regions of high concentration of these targets - termed T-cell epitope hotspots.

**2.6. Peptide vaccines**

The improved knowledge of antigen recognition at molecular level has contributed to the development of rationally designed peptide vaccines. The general idea 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 by Jackob et al., (1985). They introduced recombinant DNA and express epitopes against cholera in *Escherichia coli*. Epitope-based vaccines can be constructed for T and B lymphocytes (Dermime et al., 2004, Meloen et al., 2001). The T-cell epitopes are typically peptide fragments, whereas the B-cell epitopes can be proteins, lipids, nucleic acids or carbohydrates (Dermime et al., 2004, Meloen et al., 2001, Sundaram et al., 2004, Mahler et al., 2003, Lehner et al., 1990). Peptides have become desirable vaccine candidates owing 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 (Patronov and Doytchinova, 2013).

**2.7. Structure-based methods**

The structure-based methods do not solely rely on binding data and sequence information, but rather use the structural information, and use computational methods.
developed in the field of structural biology for prediction of potentially good binders. For the MHC molecule to recognize antigenic peptides, geometric and electrostatic complementarities between the receptor and ligand are essential for the formation of a stable complex. By aligning the sequences known to bind to a given MHC molecule, residues favoring the binding could be identified along the peptide. The synthesis of this knowledge together with that obtained from crystallographic studies has led to better understanding of the basic principles that guide peptide–MHC recognition (Kellam et al., 2008; Schonbach et al., 2008).

2.8. Docking of peptides and screening of peptide libraries

In the field of molecular modeling, docking is a method which predicts the preferred orientation of one molecule to a second when bound to each other to form a stable complex (Lengauer et al., 1996). Over recent years, many techniques and methods, such as combinatorial peptide library screening and ligand docking, commonly used in the drug design field, have found their application for the purposes of bioinformatics. Davenport et al., (1995) generated MHC class II models by evaluating the contribution of a given amino acid to the overall peptide affinity. They took into account how frequently the amino acid is present at a certain position. New peptides exhibiting affinity towards DRB1*0101 were found based on relationships derived from peptide libraries (Fleckenstein et al., 1996). Screening of peptide libraries was also applied for studying other MHC alleles. Computer-simulated ligand docking is a quick and powerful technique for investigating intermolecular interactions. In general, the purpose of docking simulation is twofold: to find the most probable translational, rotational and conformational juxtaposition of a given ligand–receptor pair and to evaluate the relative binding affinity of the ligand towards its receptor. Docking is mostly known for its wide application in computer-aided drug design (Vajda et al., 2004). However, this approach found its application for designing novel peptides exhibiting binding affinity towards MHC molecules. Initially, the docking studies were mainly used for investigation of peptides that bind MHC class I molecules. Sezerman et al., 1993; Rosenfeld et al., 1995. Zeng et al., 2001 used residues with different properties (polar, hydrophobic, charged, etc.) by docking them to different positions of the binding groove of the receptor, thus
evaluating the most acceptable residues’ properties for each position of the potential epitope. The interaction between the T-cell receptor and the MHC–ligand complex were also studied via docking (Wu et al., 2002; Buslepp et al., 2003). Tong et al., 2004 develop a novel docking approach that consists of three steps: (i) anchor residue docking; (ii) positioning of the peptide backbone in the binding groove; and (iii) adjustment of the overall positioning of the peptide backbone and the side chains. This approach showed improved accuracy in comparison with the other methods. Liu et al., 2004 take into account the flexibility of the MHC proteins during the docking simulation. Immunoinformatics can effectively leverage computational techniques to deliver effective and utilitarian advantage in the search of new vaccines. It is considered to contribute to vaccine design as the computational chemistry contributes to drug design. Immunoinformatics-based vaccine design is able to achieve effective, cost-efficient development of vaccines or vaccine components.

Figure 2.6: General scheme applied for in silico analysis in the present study.
2.9. DIAGNOSIS

Early diagnosis of the disease is important, because radical removal of metacestode tissue by surgery is considered as the most successful treatment. However, diagnosis very often is delayed because cysts develop slowly and little specific symptoms, such as abdominal pain, nausea, jaundice and feeling of exhaustion occurs. Up to 50% of AE and CE cases may remain asymptomatic and parasite lesions would incidentally be detected during examinations for other diseases. Current routine diagnosis of human echinococcosis is based on imaging procedures (ultrasound, x-ray, computed tomography and magnetic resonance imaging) (Siles-Lucas and Gottstein, 2001). Immunodiagnostic techniques such as ELISA and immunoblotting are currently applied to confirm the presence of an *Echinococcus* cyst (Eckert and Deplazes, 2004). Due to the easy availability of IgG-ELISA, this technique is probably the best overall choice for use in immunodiagnosis for human CE (Orduna et al., 1997).

Hospital based studies and case reports revealed that the disease is endemic in many parts of India, particularly in sheep raising states (Sibal et al., 2013, Malik et al., 2010). Himachal Pradesh and Jammu Kashmir are sheep rearing states in North India with high endemicity for human hydatidosis (Sharma et al., 2013). The present study was therefore designed to investigate the epidemiology of echinococcosis in Himachal Pradesh by determining the seroprevalence of *Echinococcus granulosus* antibodies among occupational risk group of Solan District. For analyzing the seroprevalence of hydatidosis, IgG and its subclasses (IgG1, 2, 3 and 4) were taken and tested using Indirect ELISA.

Hydatid cyst fluid has been used as a main antigenic source for the primary immunodiagnosis of human cystic echinococcosis (Ortona et al., 2003). ELISA using crude hydatid cyst fluid (EgHF) has a high sensitivity (75 - 95%), but its specificity is often unsatisfactory (Keller, 2002). The main problem is its cross-reactivity with sera from individuals infected by other helminths, mainly *E. multilocularis* and *Taenia solium* (Leggatt et al., 1992).

Nonetheless, to date there is no standard, highly sensitive and specific test available for immunodiagnosis of human CE (Li et al., 2003). It was also shown that carbohydrate and lipid content varies between fertile and non-fertile cysts. Thus, there is
somehow a lack of standardization in using a native parasite material as main source of antigen due to varying content of its components (Siles-Lucas, 1998). Thus source and fertility of cysts seem to be critical for test outcome. Despite of important progress made in the improvement of sensitivity of various diagnostic methods, the variability of the source of HCF Ag as well as the fact that HCF can contain host serum components remains the two major problems of effective diagnosis. Even if highly sensitive tests are used, such as the IgG-ELISA, antibodies may not be detectable in a certain proportion of patients with echinococcosis (false-negative). Cysts in the brain or eye and calcified cysts often induce low or no antibody titers. Antibody response may also be low in certain human population groups and in young children. False positive results may also occur, especially in patients with other helminthic diseases (WHO, 2002). This stresses the need for a standardized antigen such as a defined E. granulosus gene product.

It has been suggested that serodiagnosis of CE may be improved by use of recombinant proteins, synthetic peptides or combinations of well-defined antigens that enhance diagnostic specificity (Zhang et al., 2003). Through constant improvements of synthesis protocols, reagents and instruments, it has become possible to synthesize peptides of varying lengths (Corradin et al., 2007). Already shorter peptides of about 20-25 amino acids in length were proven suitable antigens for the diagnosis of various infectious diseases, not only of viral origin (Zrein et al., 1993; Alcaro et al., 2003; Chan et al., 2004,) but also of parasitic diseases (Kong et al., 2003; Intapan et al., 2005; Madhumathi et al., 2010). Benefits of synthetic peptides are unlimited availability, stability and reproducibility. It has earlier been reported the specificity of protein epitope(s) of E. granulosus antigen 5 defined by a mouse monoclonal antibody was recognized by human hydatid sera (Chamekh et al., 1990). Santivanez et al., (2012) reported that the sensitivity and specificity of peptide derived from Ag B was 78.69% and 96.88%, respectively for the serodiagnosis of human lung cystic hydatid disease. Similarly, Rouhani et al., (2013) reported that the specificity was increased from 72.46% to 97.65% using specific synthetic peptide derived AgB of E. granulosus. Gao et al., (2012) carried out serological evaluation of recombinant synthetic peptides of E. granulosus antigens for detecting cystic echinococcosis by ELISA. The specificity presented by these antigens was 96.0%, 98.0% and 97.0% respectively.
Thus these findings suggest that the specificity of diagnosis of CE could be increased using the recombinant peptides and may prevent the misdiagnosis encountered so far. Considering the need for better and accurate diagnostic protocol, six antigens, Antigen 5, Antigen B, Heat shock proteins- Hsp-8, Hsp-90, Phosphoenol pyruvate carboxy kinase (PEPC), and Tetraspanin-1 (Tsp-1) were selected for prediction of their immunogenic peptides which could be used for rapid diagnosis using in-silico tools.

2.9.1. **Antigen-5:**

Ag5 (Capron et al., 1967), one of the main component of HCF, is a thermo labile and highly immunogenic protein which is composed of 57 and 67 kDa components (Di Felice et al., 1986), that under reducing conditions dissociates into 38 and 22–24 kDa subunits (Lightowlers et al., 1989). The biological role of this molecule is almost completely unknown, although its elevated concentration in HCF suggests a relevant function in the development of the metacestode. The clonal sequence of both the subunits indicates that the 38 subunit is closely related to serine proteases of the trypsin family and the 22 kDa subunit has heparin sulphate proteoglycans and calcium binding sites, suggesting that the molecule provides interactions with cell surfaces and the extracellular matrix (Lorenzo et al., 2003).

2.9.2. **Antigen B:**

Antigen B is one of the major diagnostically relevant antigens found in HCFA. It is a polymeric lipoprotein of 120-160 kDa that aggregates from 8 kDa subunits which are encoded by a multigene family. A recent study showed that this multigene family comprises at least five subfamilies and identified 10 unique family members which were differentially expressed throughout the life-cycle (Zhang et al., 2010). This study also showed that antigen B genes were highly conserved even in isolates from distant geographical locations or from different hosts (Zhang et al., 2010). These findings contradicted previous reports claiming significant variation in antigen B family members isolated from different hosts (Frosch et al., 1994; Chemale et al., 2001; Kamenetzky et al., 2005). The implication of this recently detected sequence conservation is that antigen B subunits can be considered valuable antigens for immunodiagnosis. The biological
function of antigen B is not yet fully understood. A possible role of antigen B could be in the evasion of host immune response. Antigen B was shown to inhibit neutrophil chemotaxis (Shepherd et al., 1991) and to elicit a non-protective Th2 immune response (Rigano et al., 2001).

Immunoblot detection of the smallest subunit (8 kDa) of AgB provides a good diagnostic tool for cystic hydatid disease (Maddison et al. 1989, Legatt et al. 1992). Nonetheless, 18% of sera from cystic hydatid patients did not react and 39% of sera from alveolar hydatid patients cross-reacted with that subunit (Maddison et al. 1989). The available data on the diagnostic use of recombinant AgB (Helbig et al., 1993) suggests that they do not produce a relevant improvement of diagnostic sensitivity and/or specificity, as compared with the corresponding native antigens previously used. Thus it is possible that both antigens may contain epitopes recognized by sera from cystic hydatidosis patients as well as from patients with other parasitic diseases. Nonetheless, the possible occurrence in the same antigenic molecule of some epitopes solely recognized by antibodies in sera from cystic hydatidosis patients cannot be ruled out. Therefore, the identification of *E. granulosus*-specific epitopes in each antigen and the design of the appropriate synthetic peptides to mimic them are relevant research aims in this area.

2.9.3. **Heat shock proteins:**

Heat shock proteins, as a class, are among the most highly expressed cellular proteins across all species (Csermely et al., 1998). Heat shock protein 90 (Hsp90) is one of the most common of the heat-related proteins. The "90" comes from the fact that it weighs roughly 90 kiloDaltons. A 90 kDa protein is considered fairly large for a non-fibrous protein. Hsp90 is found in bacteria and all branches of eukarya, but it is apparently absent in archaea (Chen et al., 2006). Heat shock proteins are well known to play an important role in generating an immune response against malaria, schistosomiasis, trypanosomiasis, leishmaniasis, filariasis, syphilis, tuberculosis, leprosy, legionnaires’ disease, Lyme disease, Q fever etc (Shinnick et al., 1990), thus indicating them to be an important marker in vaccine development and diagnosis.
2.9.4. Tetraspanins:

The tetraspanins constitute a diverse super family of transmembrane protein, with 33 members in mammalian cells (Garcian-Espana et al., 2008). Tetraspanins are involved in basic cell functions such as range of physiological processes including sperm egg fusion, antigen presentation, and tissue differentiation. Tetraspanins are relatively small membrane proteins (200-350 a.a.) and share common structural features: four transmembrane domains, two extracellular loops (EC1 and EC2), a small intracellular loop and two short intra cytoplasmic termini (Stipp et al., 2003; Boucheix et al., 2001). There are several studies reporting the tetraspanins as potential candidates for diagnosis and vaccine development in other helminthic parasites like Schistosoma mansoni (Tran et al., 2006; Da’dara et al., 2001) and E. multilocularis (Dand et al., 2009).

Recently, many studies have focused on the tetraspanins, a transmembrane protein family, for their multiple functions involved in the coordination of intracellular and intercellular processes, including signal transduction, cell proliferation, adhesion, migration, cell fusion, and host–parasite interactions (Levy and Shoham, 2005). The host–parasite interactions are thought to be associated with immune evasion (Tran et al., 2006; Da’dara et al., 2008), which has resulted in the use of tetraspanins as vaccines interfering with the schistosome survival strategy. Several tetraspanins, TSP1, TSP2 and Sm23 of Schistosoma mansoni, and SJ23 of S. japonicum (Da’dara et al., 2008; Zhu et al., 2004) have been reported as potential vaccine candidates against schistosomiasis. It is notable that these tetraspanins caused varying reductions in different parasitic stages: adult worms (TSP1, TSP2, Sm23 and SJ23), liver eggs/granulomas (TSP1, TSP2 and SJ23) and intestinal/fecal eggs (TSP2 and SJ23).

Although such successes have been achieved using tetraspanins to protect mice against different stages of schistosome infections no studies have been made on tetraspanins in Echinococcus, especially with a view to exploiting their vaccine potential and diagnosis.
2.9.5. **Phosphoenol Pyruvate Carboxykinase (PEPCK)**

Phosphoenol pyruvate carboxykinase (PEPCK) is an enzyme in the lyase family used in the metabolic pathway of gluconeogenesis. It converts oxaloacetate into phosphoenol pyruvate and carbon dioxide. It was reported to be highly expressed in all the life cycle stages of *E. granulosus* and suggested to be an important marker for diagnosis of CE (Zheng *et al.*, 2013).

2.10. **VACCINATION**

Immunization of intermediate hosts with *E. granulosus* recombinant proteins may provide protection against the infection of onchosphere larvae. With the development of bioinformatic technology, the epitope based vaccines (Ben-Yedidia *et al.*, 2005) have been shown to be potential therapeutic agents in immune prevention. A major challenge in the development of epitope-based vaccines is to establish the immunogenic sites of the antigen that exhibit the greatest efficacy (Pfaff *et al.*, 1988).

Pfaff *et al.*, (1988) demonstrated that certain amino acid peptides in the foot-and-mouth disease virus (FMDV) contained immunological epitopes for the neutralization of the virus and can be used for the preparation of epitope based vaccines. In addition, Kouguchi *et al.*, (2011) revealed that the Emy162 recombinant antigen induced a 74.3% protective immune effect in rats. Similarly, Katoh *et al.*, (2008) cloned the Em95 antigen, and generated a vaccine based on this antigen in order to protect against the larval-stage infection of *Echinococcus multilocularis*. These observations suggested that the prevention of hydatid disease is quite feasible by a molecular vaccine. Keeping these points in view the present study was aimed to identify the T and B cell epitopes targeting the potential vaccine candidates of *E. granulosus* like Low Density Lipoprotein (LDL)-Receptor, Tetraspanin, Reticulon-4, Eg-95 and Glutathione-S-transferase (GST).

2.10.1. **Low Density Lipoprotein Receptor:**

Relationship of serum cholesterol levels with parasitic infection status in man has drawn the attention of various workers. Recent studies have shown elevated levels of lipoproteins like high density lipoprotein (HDL), low density lipoprotein (LDL) and total
cholesterol in patients suffering from parasitic infections (Faucher et al., 2002). However, such changes in the lipid profile of the patients are still not clear. The whole genome study of *E. granulosus* indicated that it does not generate cholesterol *de novo* (Zheng et al., 2013), as it lacks the essential enzymes like squalene synthase, squalene monooxygenase required for the cholesterol synthesis, thus depending completely on the host for it just like Schistosomes (Berriman et al., 2009) and *Toxoplasma gondii* (Coppens et al., 2000). Cholesterol ester provided by the host is the only source of cholesterol for *E. granulosus*. This notion is supported by the presence of sequences encoding cytoplasmic sterol O-acyltransferase and transmembrane cholesterol esterasein *E. granulosus* (Zheng et al., 2013). Thus targeting the LDL receptor of *E. granulosus* for vaccine development could be very effective. Several vaccine candidates of *Echinococcus granulosus* have been identified, among which the gene encoding Low-Density Lipoprotein (LDL) receptor is highly expressed during all stages of its life cycle including adult worms, oncospheres, scolesces and hydatid cyst membranes (Zheng et al., 2013). Thus keeping these facts in view, the present study was designed to predict T and B cell epitopes of Low-Density Lipoprotein (LDL) receptor of *E. granulosus*.

### 2.10.2 Glutathione S transferase (GST):

Glutathione (GSH) transferases are widely distributed as isoenzymes in the plant and animal kingdoms (Mannervik, 1985). They are multifunctional proteins that can function as enzymes catalyzing the conjugation of the GSH thiolate anion with a multitude of second substrates, or they can function as non-covalent binding proteins for a range of hydrophobic ligands (Ketterer, Meyer & Clark, 1989). GSH transferases can detoxify exogenously derived toxic compounds (xenobiotics) and endogenously derived toxic compounds, including the secondary products of lipid peroxidation, using both their enzymatic and binding capabilities (Ketterer et al., 1989; Mannervik, 1985). A number of GSH transferases may have more specific endogenous functions in metabolism, including a role in leukotriene/prostaglandin synthesis (Mannervik, 1985) and as transport proteins for ligands such as heme, steroids and bile acids. The native mammalian enzyme is either a homo- or hetero-dimer. Helminths have limited detoxification enzymes and appear to lack the important cytochrome P-450- dependent detoxification reactions (Precious &
Barrett, 1989). GSH transferase has been detected in a range of helminths (Brophy, 1988) where it may be one of the major detoxification enzymes.

In contrast to its roles in the development of drug resistance and protection against lipid peroxidation GSH transferases have been identified as important protective antigens in schistosomiasis (Tiu et al., 1988; Mitchell, 1989). Two groups have reported the presence of GSH transferase in both the parenchyma and tegument of S. mansoni (Balloul et al., 1987; Taylor et al., 1988) while a third group has presented evidence for the absence of tegumental S. mansoni GSH transferase (Holy et al., 1989). There are also biochemical and physical differences, such as variation in catalytic activity and native molecular mass, between the major GSH transferase isolated from S. mansoni by two of the research groups (Taylor et al., 1988; O'Leary & Tracy, 1988). It is possible that the enzyme becomes exposed during surface turnover, and exposure may be increased by drug- or immune induced surface damage. Immune responses to intracellular proteins have previously been reported, for example paramyosin is a protective antigen in murine schistosomiasis (Sher & Pearce, 1988), but in this case is thought to elicit a T lymphocyte dependent, cell-mediated response rather than acting as a target for host antibody. There is no evidence that GSH transferases are important antigens in parasites other than schistosomes (Sher et al., 1989).

GST is one of the highly expressed protein during all life cycle stages of E. granulosus, thus seems to be critical for parasite survival, additionally also predicted to be potent vaccine target for cystic echinococcosis (Zheng et al., 2013). Such facts were the keen of interest for us, thus we studied the GST protein of Echinococcus and identified T and B cell epitopes targeting the GST protein.

2.10.3. Reticulon-4:

Reticulon-4 is a transmembrane protein which is associated with endoplasmic reticulum (van de Velde et al., 1994). Though its function is not clearly understood but recently it is reported to be an important protein in malarial (Manjurano et al., 2012) and Schistosomal (Prosdocimi et al., 2002) infections. The expression of Reticulon-4 protein was also found high in larval stages of Schistosoma (Prosdocimi et al., 2002). Several
vaccine candidates of *E. granulosus* have been identified, among which the gene encoding Reticulon-4 protein is one of the highly expressed protein during all stages of life cycle - adult worms, oncospheres, scoleces and hydatid cyst membranes (Zheng *et al.*, 2013).

2.10.4. **Eg95:**

The EG95 vaccine antigen is a secreted glycosylphosphatydilinositol (GPI)-anchored protein containing a fibronectin type III domain, which is ubiquitous in modular proteins involved in cell adhesion. EG95 is highly expressed in oncospheres, the parasite life cycle stage which actively invades the intermediate hosts.

The interest on EG95 comes largely from its high potential as a vaccine to protect against *Echinococcus* infection. Related antigens from Taenia (45W, 18K and 16K), showing a similar molecular structure and expression pattern (Waterkeyn *et al.*, 1997) provide similar levels of protection against cysticercosis caused by *Taenia solium*. The gene encoding Eg95 protein is one of the highly expressed proteins during all stages of life cycle - adult worms, oncospheres, scoleces and hydatid cyst membranes (Zheng *et al.*, 2013).

2.11. CHEMOTHERAPY

Treatment of human CE by chemotherapy has been developed following a series of animal studies, which demonstrated the efficacy of benzimidazolecarbamates against *E. granulosus* and *E. multilocularis* (Saimot *et al.*, 1983). Albendazole and mebendazole are considered to be equally effective. When evaluated up to 12 months after initiation of benzimidazole treatment, 10%–30% of patients exhibit cyst disappearance, 50%–70% show degeneration of cysts and 20%–30% show no morphological changes in cyst appearance (Horton, 1997). Continuous or intermittent treatment with albendazole is recommended for a period of 6 months. Praziquantel, shown to exhibit protoscolicidal activity, has been reported to enhance the effect of benzimidazoles, especially when cyst spillage was considered to pose a problem. In about 75% of patients, metacestodes exhibit clear signs of degenerative changes (Horton, 1997). However, adverse side effects have been observed with both benzimidazoles, including abnormalities in liver function,
leucopenia and alopecia, and albendazole has been shown to be teratogenic in rats and rabbits. It should therefore be avoided during pregnancy and lactation (Kern, 2003; Horton, 1997). As a consequence, the development of alternative means of treatment of CE is necessary.

Therefore, the search for novel drug candidates for the treatment of the disease has to be pursued. Improved drug treatments with lesser side effects are needed (Hemphill and Muller, 2009; Reuter et al., 2004). Thus considering the importance for intervention of new drugs, the present study focused on the ability of amides for their antiechinococcal activity using in silico and in vitro approach. Certain amide derivatives have previously been reported to possess potent antiparasitic activity. Walker et al., (2008) reported the potent antiechinococcal activity of nitazoxanide (an amide derivative) against protoscolec es and metacestodes of Echinococcus granulosus. Similarly, Qtaitat et al., (2014) reported the antimalarial, antileishmanial and antischistosomal activity of novel metronidazole-piperazine amides. Draper et al., (2014) reported Pyridyl ethyl amides as potential antitrypanosomal agents, and arylimidamide-azole hybrids as potential antileishmanial agents. Smit et al., (2014) reported the antimalarial activity of novel 4-aminoquinolinyl-chalcone amides. Dornbush et al., (2010) reported the antiparasitic activity of 3,4-dichloroaniline amides against Trichomonas vaginalis. Such studies suggest that the treatment of Echinococcosis is feasible by amide derivatives as alternative drugs.