Prediction of vaccine candidates for dental caries using Immunoinformatics

A Major Project dissertation submitted

in partial fulfilment of the requirement for the degree of

Master of Technology

In

Bioinformatics

Submitted by

Meenakshi Yadav

(2K10/BIN/07)

Delhi Technological University, Delhi, India

Under the supervision of

Dr. Asmita Das
Department of Biotechnology
Delhi Technological University
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, New Bawana Road, Delhi-110042
INDIA

ACKNOWLEDGEMENT

“All intelligent thoughts have already been thought; what is necessary is only to try to think them again.”

I would like to express my heartfelt gratitude to all those who have contributed directly or indirectly towards obtaining completion of this project. I am grateful to my esteemed supervisor, **Dr. Asmita Das**, who has guided me through thick and thin. Many other people like **Ms. Pooja Kesari** who have provided me with the courage to complete this task on time. I would also like to acknowledge the many helpful comments received from my Teachers of the biotechnology and bioinformatics department. A special thanks to the person, **Dr. Pawan Yadav** who suggested this topic and helped me by all possible means.

I am obliged to all those who provided reviews & suggestion for improving the result and the topics covered in my project, and extend my apologies to anyone if I may have failed to mention.

Meenakshi Yadav

Roll No. 2K10/BIN/07

CERTIFICATE

This is to certify that the M. Tech. dissertation entitled **"PREDICTION OF VACCINE CANDIDATES FOR DENTAL CARIES USING IMMUNOINFORMATICS",** submitted by **MEENAKSHI YADAV (2K10/BIN/07)** in partial fulfillment of the requirement for the award of the degree of Master of Engineering, Delhi Technological University (Formerly Delhi College of Engineering, University of Delhi), is an authentic record of the candidate's own work carried out by her under my guidance.

The information and data enclosed in this dissertation is original and has not been submitted elsewhere for honouring of any other degree.

Date: 31/07/2014

Dr. ASMITA DAS
(Project Mentor)
Department of Bio-Technology
Delhi Technological University
(Formerly Delhi College of Engineering, University of Delhi)

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

S.mutans	*Streptococcus mutans*
HLA	Human Leukocyte Antigen
GI	GenInfo Identifier
NCBI	National Centre for Biotechnology Information
UniProtKB	UniProt KnowledgeBase
HOMD	Human Oral Microbiome Database
BCPREDS	B-cell Epitope PREDiction Server
MHC	Major Histocompatibilty Complex
IgG	Immunoglobulin G
IgE	Immunoglobulin E
IgA	Immunoglobulin A
CPP	Casein phosphopeptide
CPP-ACP	Casein phosphopeptide-amorphous calcium phosphate
GTF	Glucosyltransferase
GBP	Glucan Binding Protein
VIP	Virulence Associated Immnumodulatory Extracellular Protein
AAP	Amino acid pair

CONTENTS

PREDICTION OF VACCINE CANDIDATES FOR DENTAL CARIES USING IMMUNOINFORMATICS

MEENAKSHI YADAV

Delhi Technological University, Delhi, India

1. ABSTRACT

Despite the existing preventive measures, dental caries remains one of the most pervasive infections in humans. Epidemiological studies imply that there is a recent increase in dental caries in children. Conventionally, *Streptococcus mutans* has been considered as the earliest colonizer of oral cavity and the main etiological agent of dental caries. However, there are reports where other microbes have resulted in dental caries even in the absence of *S.mutans*. These studies support the poly-microbial nature of this disease and the use of immunization strategies using vaccine targets shared by different pathogens involved in the process of tooth decay. Till date, almost all the efforts put in to immunize people against caries have targeted *S.mutans* only. Only a few studies have tried to develop vaccines targeting both *S.mutans* and *S.sobrinus* with little or no success.

This study, aimed at preventing dental caries, provides the basis for identification of vaccine candidates for developing a dental caries vaccine that can elicit both B-cell and T-cell mediated immune response against multiple cariogenic microorganisms. A novel strategy has been used to predict such antigenic B-cell epitopes which contain T-cell epitopes also. The predicted T-epitopes bind to the maximum number of HLA-DR alleles and bind with high affinity to the most frequently occurring HLA-DR alleles in the human population and hence have worldwide protective effects.

Keywords: Dental caries, cariogens, *Streptococcus mutans*, vaccine, epitope, B-cell, T-cell, HLA alleles, antigenicity, immune response.

2. INTRODUCTION

Dental caries, one of the most prevalent chronic diseases (Selwitz *et al*., 2007), is an infection that can lead to pain and tooth loss. Despite the measures taken to prevent caries, these are still with us. It is one of the most common infections in humans (Smith DJ, 2002). The possible reasons of prevalence of dental caries are lack of proper dental care, the microbial flora in oral cavity, higher uptake of dietary sugars, inadequate salivary flow, insufficient fluoride exposure, poor oral hygiene, inappropriate methods of feeding infants etc (Selwitz *et al*., 2007). Moreover, in developing countries economics and infrastructure are responsible for deprivation of people to good dental health care.

To manage infection that has reached an epidemic level and is of recurring nature, the best strategy is that of vaccination as it augments the host's natural defense system. Though the main etiological agent identified in dental caries is *S.mutans* (Loesche WJ, 1986), the role of other bacteria in this infection cannot be neglected. Becker *et al* in 2001 identified *Actinomyces gerencseriae, Bifidobacterium, S.mutans, Veillonella, S.salivarius, S.constellatus, S.parasanguinis* and *Lactobacillus fermentum* as cariogenic bacteria associated with childhood caries (Becker *et al*., 2002). Kleinberg described a mixed-bacteria ecological approach for understanding the role of different microorganisms in dental caries causation (Kleinberg I., 2002). The mixed-bacteria approach gives the idea of developing drugs or vaccines that are directed to a number of microorganisms and not just *S.mutans*. Further, Okada *et al* concluded that the presence of both *S.mutans* and *S.sobrinus* results in higher incidence of dental caries than with *S.mutans* alone (Okada *et al*., 2005). In some other studies, dental caries have been observed even in the absence of *S.mutans* (Gross *et al*., 2012).

As the role of microorganisms other than *S.mutans* cannot be negated in the causation of dental caries, anti-caries agents targeting as many cariogenic microbes as possible are needed to combat this chronic disease. To develop a vaccine for this disease, the proteins that are conserved in majority of these microorganisms should be targeted. Such a broad range vaccine is expected to inhibit the growth of caries associated bacteria found responsible for initiation and subsequent progression of caries.

To provide broad range protection, a vaccine should be able to elicit both humoral and cell mediated immune response. A number of studies have used the strategy of predicting B-cell epitopes and T-cell epitopes originating from the said B-cell epitopes only (Barh *et al*., 2010). These B-cell epitopes are capable of inducing both B-cell and T-cell mediated immune responses.

In the present study, vaccine candidates targeting a number of cariogens and capable of evoking both branches of adaptive immunity, thereby providing the widest possible protection from dental caries have been predicted.

3. REVIEW OF LITERATURE

3.1 DENTAL CARIES

Dental caries is the result of localized destruction of susceptible dental hard tissues by acidic by-products from bacterial fermentation of dietary carbohydrates (Marsh *et al*., 1999). It is one of the most common avertable pediatric diseases (Selwitz *et al*., 2007). Considering the US population, 90% of adolescents and young adults have been diagnosed with dental caries and 94% of dentate adults have medical history of treated or untreated coronal caries (Featherstone JD, 2000; Featherstone JD, 1999). It is a disease which results in the damage of whole tooth as it progresses. Dental caries is the result of complex interaction between the host, his/her diet and the microflora on the tooth surface bounded by the time factor (Kruger *et al*., 2004) which leads to the dissolution of inorganic and destruction of organic matter of the tooth.

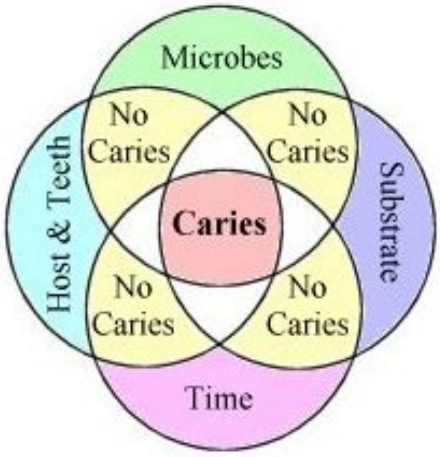


Figure 3.1.1: Showing the factors responsible for dental caries

(Source: http://www.ncl.ac.uk/dental/oralbiol/oralenv/mcqs/oralmicro/mutans2.htm)

3.1.1 PREVALENCE OF CARIES

A survey by Public Health England in 2013 reported that 27% of 5 year olds have tooth decay (available at: https://www.gov.uk/government/news/survey-finds-27-of-5-year-olds-have-tooth-decay). According to the WHO, more than 5 billion people have experienced tooth decay (Smith DJ, 2002). In industrialized nations, 60-90% school aged children have been diagnosed with dental caries (Petersen *et al*., 2005). This data indicates the high preponderance rate of dental caries.

The prevalence of dental caries in India was reported as 55.5% in 1940 and 68% in 1960. In 2004, a National Oral Health Survey by National Council of India reported dental caries in 51.9% of the 5-year old children, 53.8% in 12 year old children 63.1% in 15 year old teenagers. In a study conducted in the Dakshinpuri area of New Delhi, the prevalence of dental caries in the age-groups 35-44 years and ≥60 years was found to be 82.4% and 91.9% respectively (Patro *et al*., 2008). There are various other regional studies assessing the prevalence of dental caries in children of different age groups (Basha *et al*., 2012; Moses *et al*., 2011).

3.1.2 MICROBIOLOGY OF DENTAL CARIES

Previously, *S. mutans* has been implicated as a causative organism of dental caries (Smith DJ, 2002). *S. mutans* accounts for seven distinct species isolated from animals and humans; *Streptococcus cricetus, Streptococcus ferus, Streptococcus macacae, Streptococcus rattus, Streptococcus downey, S. mutans,* and *Streptococcus sobrinus*. *S. mutans* and *Streptococcus sobrinus* are exclusively isolated from humans and *S. mutans* is the most prevalent species (Graner *et al*., 2004).

In further studies, other microorganisms capable of initiating caries were also identified. Elevated levels of *S. salivarius, S. sobrinus*, *S. parasanguinis and Veillonella* have been associated with caries even in the absence of *S. mutans* in the subjects. Krithika *et al.* identified *Streptococcus mutans*, *Lactobacillus acidophilus*, and *Actinomyces viscosus* as the main pathogenic species involved in the initiation and development of dental caries (Krithika *et al*., 2004; Russel *et al*., 2004). *Actinomyces gerencseriae, Bifidobacterium, S.mutans, Veillonella, S.salivarius, S.constellatus, S.parasanguinis* and *Lactobacillus fermentum* have been associated with childhood caries (Becker *et al*., 2002). Scientific literature is replete with studies demonstrating the cariogenic potential of microorganisms other than *S.mutans* (Torlakovic *et al*., 2012; Preza *et al*., 2008; Brailsford *et al*, 1999).

These studies prove the existence of multiple pathogens in the causation of dental caries and suggest that a strategy targeting multiple microorganisms, i.e., mixed-bacterial approach is needed to prevent caries (Gross *et al*., 2012; Kleinberg I., 2002).

3.1.3 INITIATION AND PROGRESSION OF CARIES

The adherence of bacteria on tooth surface produces dental plaque. Dental plaque is a pale yellow colored biofilm which develops naturally on the teeth. Mature dental plaque is a complex multispecies biofilm that grows on the tooth surface and is embedded in a protective matrix and bacterial polymers (polysaccharides, proteins, and DNA) secreted by the cells (Marsh *et al*., 1999). The protective matrix provides protection from desiccation, host defenses and predators and provides enhanced resistance to antimicrobial agents (Weatherell JA, 1975). Bacteria in the biofilm (dental plaque) utilize dietary carbohydrates to produce organic acids as metabolic by-products. These acids cause a decrease in local pH and when the pH falls below a critical value, demineralization of the tooth tissue occurs (Weatherell JA, 1975). Specifically speaking, the production of lactate by the acidogenic oral microflora causes demineralization of calcium and phosphate present in the crystal form of hydoxyapatite, which comprises the enamel of the teeth (Weatherell JA, 1975). When the frequency and rate of acid production exceeds the natural re-mineralization activity of the teeth, demineralization occurs and results in the subsequent progression of cavitations, provided the pH remains below a 'critical' value of approximately 5.5–5.3 for a sufficient amount of time (Takahashi *et al*., 2008).

In the demineralization process, the organic acids produced by bacteria diffuse into the tooth surface through the water amongst the hydroxyapatite crystals. When a susceptible site, formed due to impurities and inclusions of other ions like carbonate in the hydroxyapatite crystals, comes in contact with these diffused acids, dissolution of calcium and phosphate into the surrounding aqueous phase between the crystals occurs (Featherstone JD, 2004). When the diffusion of calcium, phosphate, and carbonate out of the tooth occurs without proper re-mineralization, cavitation takes place (Featherstone JD, 2008). Remineralization is the body's natural repair process for subsurface non-cavitated carious lesions (incipient lesions). In re-mineralization process, calcium and phosphate either from saliva or other topical sources diffuses into the tooth and with the help of fluoride, re-builds the existing crystal remnants (Featherstone JD, 2000). The rebuilt crystalline surface is composed of a coating of well-formed mineral similar to fluorapatite due to the presence of fluoride, and is much more resistant to acid attack than the original structure (Selwitz *et al*., 2007). Demineralization and re-mineralization occur simultaneously in the oral cavity. The progress in development of dental caries depends on the balance between the process of demineralization and re-mineralization. Hence, any factor that can maintain this balance towards re-mineralization can be utilized to combat dental caries.

3.1.4 PREVENTION OF DENTAL CARIES

A number of aids like fluoride gel and varnish, chlorhexidine and sealant are available for preventing dental caries (Rozier RG, 2001)**.** These aids can be grouped into physical and medical aids.

3.1.4.1 Physical Aids

Physical aids help in mechanical debridement of tooth surface. These include toothbrush, dental floss, ultrasonic's for cleaning tooth surface, dentifrices etc (Daly *et al*., 2009).

3.1.4.2 Medical Aids

Medicals aids help in prevention of plaque build-up, decrease bacterial count and shift the equilibrium towards re-mineralization.

The available medical aids to control dental caries can be classified into two groups: one that prevent demineralization of tooth and the other one that promote the re-mineralization process (Chen *et al*., 2004).

i. Therapeutics to prevent the demineralization of tooth

This category of therapeutics utilizes strategies involving: killing of bacteria, preventing their attachment to the tooth surface or detaching them. This category includes chemoprophylactic agents, antimicrobial peptides, sugar substitutes, vaccines, probiotics and replacement therapy.

Chemoprophylactic agents include antibiotics like penicillin, cationic, anionic and non-ionic agents, plant extracts such as sanguinaria extract. For chemoprophylactic agents to be effective, it is necessary to maintain their minimum inhibitory concentration in the oral cavity because the concentration of a drug decreases almost immediately after the delivery of drug. Further, the chemoprophylactic agents do not provide protection for all the strains of a particular microorganism. Though antimicrobial peptides (Brogden *et al*., 2005) have resolved this problem, but their use is limited due to the difficulty in synthesizing these peptides. These peptides are also susceptible to proteolytic cleavage, which may make them ineffective. (Marr *et al*.,2006**).**

As vaccines do not need to be maintained in minimum inhibitory concentrations like chemoprophylactic agents and utilize body's own defense mechanism to prevent a disease, they are better candidates for preventing demineralization of tooth surface.

In oral cavity, there are some bacteria which are beneficial for preventing dental caries. One strategy is the reduction in growth of pathogenic microorganisms and the promotion of growth of beneficial bacteria, i.e., probiotics.

A novel strategy is replacement therapy which is based on the concept of replacing pathogenic strains of microorganisms with non-pathogenic ones. A recombinant *S.mutans* strain BCS3-L1 incapable of metabolizing fermentable sugars has been developed (Hillman *et al*., 2000). This strain produces an antibiotic called mutacin 1140, which kills other *S. mutans* strains.

ii. Therapeutics to promote the remineralizatio process

The remineralization process is an inorganic chemistry process in which the calcium and phosphate from saliva are recrystallized on the remnants of crystals on the dentine or enamel surface. If fluoride is incorporated within the crystal during remineralization, the mineral formed is much more resistant to acids than the original enamel or dentine mineral (Featherstone JD, 1999). This fact is supported by the marked decrease in the levels of dental caries in individuals using fluoride-containing toothpastes (Nabi *et al*., 1990). This may be reason for incorporation of fluoride salts in water and mouthwashes. The salivary fluoride concentration is 0.02ppm, but the currently used fluoride delivery methods are not able to deliver this much amount of fluorine. A device was developed maintaining the salivary concentration of fluoride. This device called "glass device" dissolves slowly when it comes in contact with saliva and releases fluoride (Pessan *et al*., 2008). Further, Xu *et al*. developed a dicalcium phosphate anhydrous (DCPA) nanocomposite capable of slowly releasing $CaPO_4$ as a restoration material. In 2010, another nanocomposite based on CaF_2 was developed. This nanocomposite has greater fluoride releasing and stress bearing capabilities (Xu *et al*., 2010).

Milk, milk concentrates and cheese have been identified as non-carigenic or cariostatic because of the presence of casein in them (Cross *et al*., 2007). In 1987, Reynolds observed the incorporation of CPPs in intra-oral appliance plaque (Reynolds EC, 1998). CPPs increased the concentrations of calcium and phosphate in plaque. CPPs form an amorphous complex with calcium phosphate called CPP-ACP, which binds strongly to hydroxyapatite and is retained in dental plaque. This complex buffers the acid produced by microorganisms and increases the levels of calcium phosphate in plaque or close proximity to the tooth surface, thereby inhibiting enamel demineralization and promoting remineralization. CPP-ACP, trademarked as Recaldent™, is currently used in sugar-free gum, mints and in dental professional product Tooth Mousse™.

3.2 IMMUNITY

Immunity is the defense mechanism of body against pathogens. Human immune system comprises of two branches: innate and adaptive immunity. Innate immunity is natural resistance of body and is non-specific in nature. It protects the body through physical (skin, epithelial layers), chemical (interferon's) and cellular (phagocytosis) means. Adaptive or acquired immunity is acquired after birth and can be acquired naturally or artificially (vaccination). Adaptive immunity comprises of humoral (B-cells) and cell mediated (T-cells) immune responses. Figure 3.2.1 illustrates the interactions of various components of immune system.

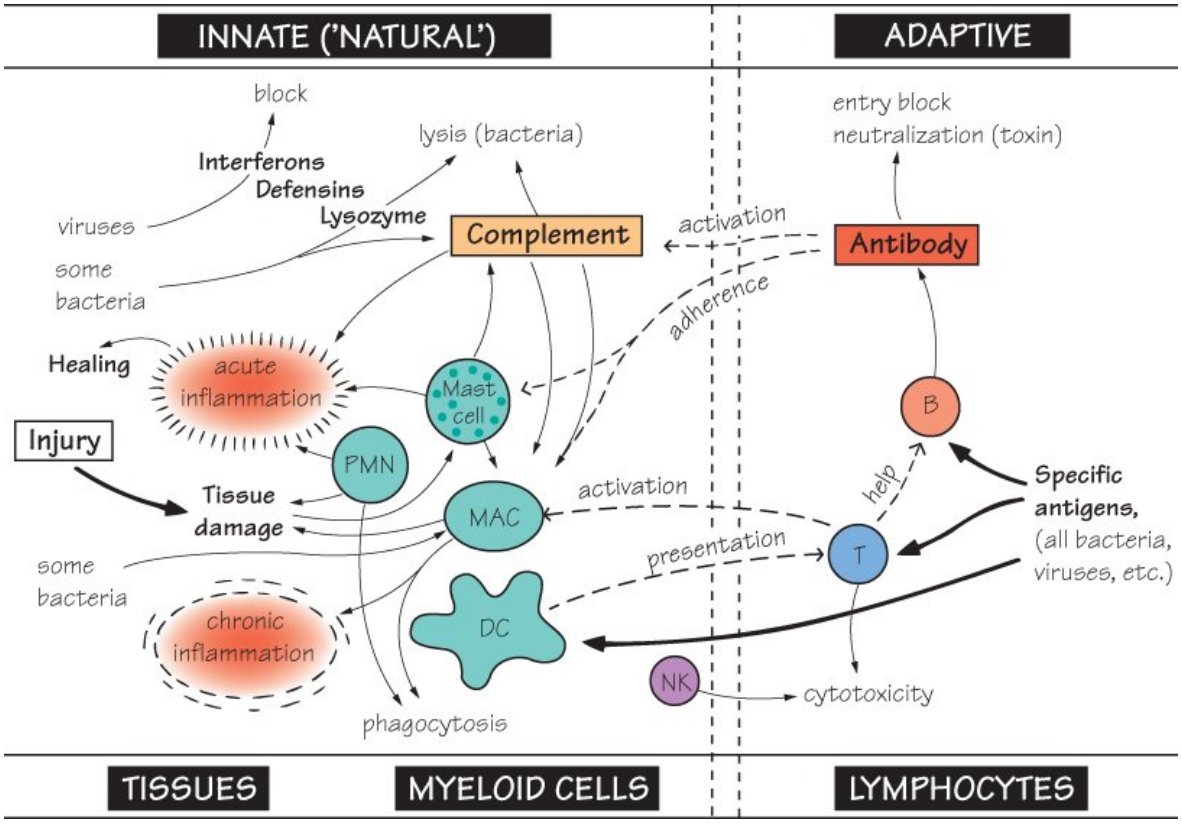


Figure 3.2.1: Showing components of immune system [Source: Immunology at a Glance, 10th Edition]

3.2.1 IMMUNE RESPONSE IN ORAL CAVITY

In the oral cavity, humoral immune response is predominant. The major immunoglobulin present in saliva is secretory IgA. Saliva also contains IgG and IgM from the gingival sulcular fluid. Lymphocytes, macrophages and neutrophils, which are the components of cell mediated immunity, are also present in gingival sulcus (Setia *et al*., 2012). The following immune mechanisms are responsible for eliciting immune responses in oral cavity:

- Agglutination: Antibodies in saliva may interact with bacterial cell surface receptors. S-IgAs do not activate complement system but if IgG interacts with the antigens, complement system may be activated. The antibodies in saliva may prevent interaction of bacteria with teeth by specifically binding to bacterial surface receptors. Further, antibodies may inhibit the activity of enzymes like glucosyltransferase and prevent plaque formation.
- Sensitization of B-cells: If the gut associated lymphoid tissue (GALT) is immunized, sensitization of B-cells occurs and these sensitized B-cells then move to the salivary glands, which in turn secrete IgA. The secreted IgA may then prevent bacterial adhesion to the tooth surface (Lehner T. *et al*., 1992).
- CD4+ cells sensitization: The gingival crevicular mechanism engages both the humoral and cellular components of the immune system. When *S.mutans* is administered subcutaneously, it is phagocytosed by the antigen presenting cells which present the processed antigens on MHC-class II molecules. The antigen-MHC class II complex is recognized by CD4+ T-helper cells which then activate B-cells. The activated B-cells then produce antibodies (IgG) against the bacterial antigen (Walker DM, 2004).

Cell-mediated immune responses may be elicited in animals after immunization with *S. mutans*, these responses may modify the humoral response via helper and suppressor functions of T-cells and may also cause gingival inflammation but they are not known to play a significant role in the immunology of caries (Setia *et al*., 2012).

3.3 ANTI-CARIES VACCINE

Vaccine is an immuno-biological substance designed to produce specific protection against a given disease. It stimulates the immune system of the host. It may stimulate humoral or cell-mediated immune response. Vaccines are prepared from live modified organisms, inactivated or killed organisms, extracted cellular fractions, toxoids, or a combination thereof.

3.3.1 VACCINE CANDIDATES IN *S.MUTANS*

Cell surface proteins or substances of *S. mutans* have been used as vaccines in a number of studies. These cell surfaces substances include adhesins, GTFs, glucan binding proteins (GBP) and dextranases. Most of the recent experimental studies for finding a vaccine against *S.mutans* have been utilizing these cell surface proteins as vaccine candidates [45].

Adhesins: Adhesins (Antigen I/II, PAC or P1 and Spa-a from *S.mutans* and *S.sobrinus* have been purified and used for vaccine preparation. Antigens I/II are found in both the culture supernatant and on *S. mutans* cell surface. An antibody specific for the Antigen I/II molecule or to its salivary binding domain was successful in blocking the adherence of *S. mutans* to saliva-coated hydroxyapatite. Synthetic peptide comprising of residues 301-319 of Antigen I/II was effective in reducing tooth colonization by *S.mutans* (Smith DJ, 2002).

Glucosyltransferase

S. mutans has the following three forms of glucosyltransferases (GTFs):

- Water insoluble glucan synthesizing enzyme: GTF-I
- Water insoluble and water-soluble glucan synthesizing enzymes: GTF-S-I
- Water-soluble glucan synthesizing enzymes: GTF-S

The genes encoding GTF-I, GTF-SI, AND GTF-S are called the *gtf-b, gtf-c, and gtf-d* genes, respectively. These three *gtf* genes have been found to be important in smooth surface caries formation in the pathogen-free rat model system. *Streptococcus sobrinus* also produces a water insoluble glucan-synthesizing enzyme *gtf-s*. *S.mutans* and *Streptococcus sobrinus* both synthesize a number of GTFs (Luo *et al*., 1988).

Glucan binding protein (GBP): *S. mutans* secretes at least three glucan binding proteins: GBP-A, GBP-B, AND GBP-C. Of the three *S. mutans* GBPs, only GBP-B has the potential to induce a protective immune response against experimental dental caries.

Dextranases: Dextran, an important constituent of early dental plaque, is degraded by the enzyme dextranase produced by *S.mutans*. Due to dextranase, *S.mutans* has the capability of invading early dental plaque. Dextranase, when used as an antigen, can prevent the colonization of the organism in early dental plaque (Krithika *et al*., 2004).

Some other antigens have also been explored for designing a vaccine against dental caries. These antigens include the virulence-associated immunomodulatory extracellular proteins (VIP), secreted by *S.mutans* and *S.sobrinus.* These VIPs evade the host immune system by inducing the production of IL-10, which suppresses the host's immune system response against bacteria. Vaccines developed from VIPs have been able to induce immune-neutralization of VIP induced immunomodulatory effects (Gomes *et al*., 2009). Vaccines have also developed using Glucans. Glucans have been found to be less antigenic.

3.3.2 VACCINE DEVELOPMENT APPROACHES

3.3.2.1 Immunoinformatics based approach

Recently, a study predicting B-cell and T-cell vaccine candidates from *S.mutans* GtfD using Immunoinformatics was published (Bower *et al*., 2014). In another study, the antigenic potential of the catalytic region (CAT) and glucan-binding domain (GBD) of glucosyltransferase B (GtfB) from Streptococcus mutans has been evaluated using *in-silico* approaches followed by *in-vitro* and *in-vivo* experiments (Hoshino *et al*., 2011).

3.3.2.2 Laboratory (Wet-lab) based approach

I. Animal studies

Most of the studies evaluating the effectiveness of dental caries vaccines use rats and monkeys as animal models. The use of purified components of *S.mutans* is still limited. GTF, when used as a vaccine provided protection against dental caries but for getting positive results 5-15 injections of GTF are required. Whereas cell wall antigen, Antigen I/II utilizes only one subcutaneous injection with adjuvant. A reduction of 70% was observed in both smooth surface and fissure caries when whole cells of *S. mutans* or purified I/II antigen was used as vaccine. In gnotobiotic rats, ingestion of whole *S.mutans* selectively produces s-IgA. The appearance of s-IgA has been correlated with a reduced incidence of the caries vaccine (Bowen WH, 2002).

Bowen for the first time reported successful immunization of monkeys against caries (Bowen WH, 1996) by injecting whole cells of *S.mutans* into *Macaca fascicularis*

monkeys and observed successful prevention of dental caries. Experimental design of almost all the studies consists of the fundamental steps which are:

(a) Immunizing animal with an antigen from *S.mutans* along with an adjuvant,

(b) Repeated immunization in order to achieve high antibody levels,

(c) Inoculating the immunized organism with the pathogen (*S.mutans*) and maintaining a high sucrose diet.

Positive results of successful immunization of rats and hamsters with GTF as anti-caries vaccine have been reported, but no such observation is found in monkeys.

Another study reported creation of a fusion DNA vaccine pGLUA-Pby cloning the GLU region of GTF into a DNA vaccine, pCIA-P, encoding two highly conservative regions of Pac and was found more efficient in preventing carious lesions in rat as compared to pCIA-P (Guo *et al*., 2004). Another DNA fusion vaccine by the same researchers could elicit protective effects against S.mutans colonization but not against S.sobrinus. The said DNA fusion vaccine was further modified to incorporate the CAT fragment of the of the *S. sobrinus* OMZ176gtf-I (Niu *et al*., 2009). Mice immunized with the modified DNA vaccine demonstrated protection against S.sobrinus infection. It was the first study that demonstrated a vaccine targeting both S.mutans and S.sobrinus.

II. Human Studies

The possibility of preventing dental caries by vaccination has been instituted due to it infectious nature. The idea is that immunization with *S.mutans* should induce an immune response that can prevent the colonization of surfaces of teeth by *S.mutans* which can surely prevent dental caries. Administration of vaccine at the age of 6 months (before the eruption of deciduous/ primary/ milk teeth) would prevent the caries in children who show the highest incidence of caries. Thereafter, booster doses can be given at regular intervals. Immunization could be done using existing delivery systems.

The effects of immunization with *S.mutans* or *S.sobrinus* proteins in humans have been documented by few studies (Smith *et al*., 1987). These small-scale human trials of dental caries vaccines have reported increase in levels of S-IgA. Oral administration of GTF from *S. sobrinus* combined with aluminum phosphate in capsule form to 14 subjects resulted in an increase in salivary IgA antibody response when combined with an aluminum based adjuvant (Smith *et al*., 1987). In another study, GTF from *S.sobrinus* administered topically onto the lower lips of young adults stimulated local antibody production in the minor salivary glands and also resulted in delayed oral re-colonization with mutans streptococci (Smith *et al*., 1990). Oral immunization of 7 adult volunteers with an enteric coated capsule containing 500 micrograms of GTF from *S.mutans* also resulted in elevating salivary IgA antibodies to the administered antigenic preparation (Childers *et al*., 1994). Levels of salivary antibodies were also elevated when similar

preparations were administered intranasal or by topical application to the tonsils, either in soluble form or incorporated in liposome's (Li *et al*., 2003).

Despite the abundance of experimental evidences for the effectiveness of dental caries vaccine, not a single one is yet available for human use. Also, all these studies are oriented to the protection of colonization of oral cavity by S.mutans or S.sobrinus. None of the studies considers the polymicrobial nature of dental caries (Belda-Ferre *et al*., 2012; Kleinberg I., 2002). In the present study, an attempt has been made to design vaccines keeping in view the polymicrobial nature of this disease.

4. METHODOLOGY

This flow chart illustrates the project workflow:

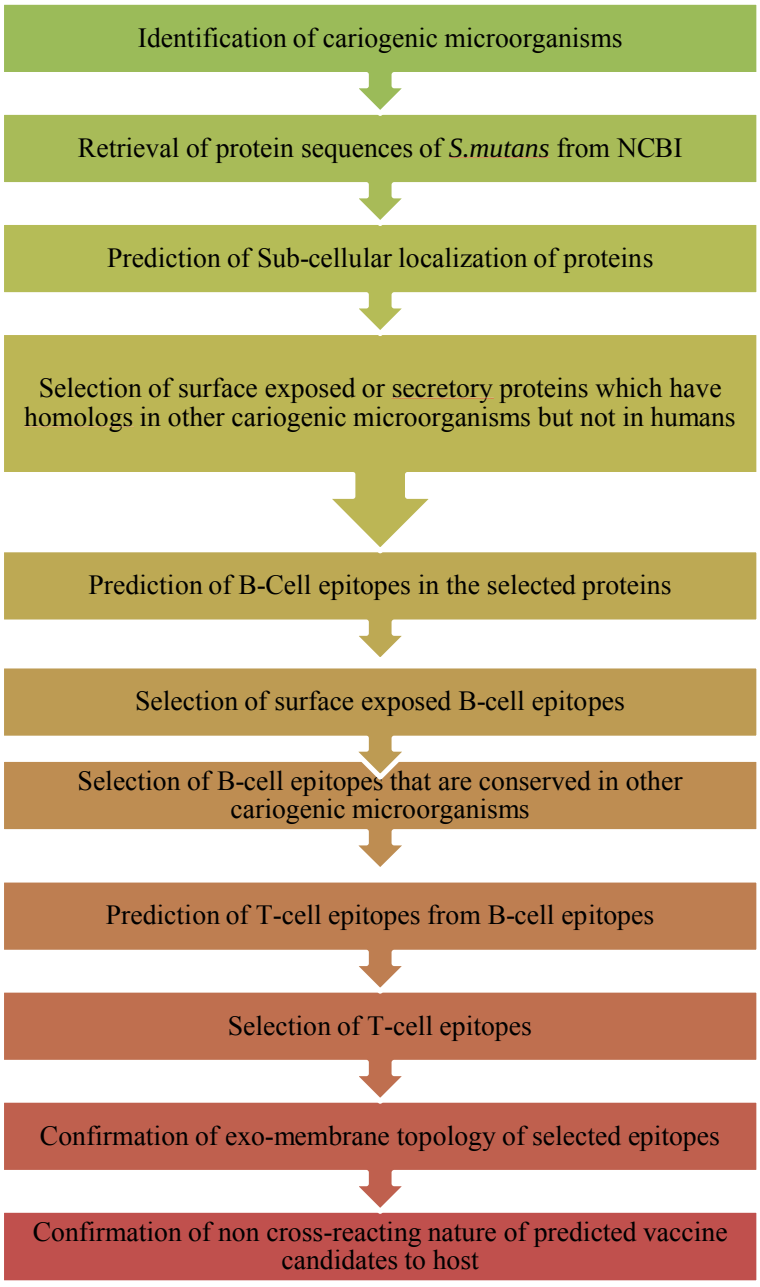
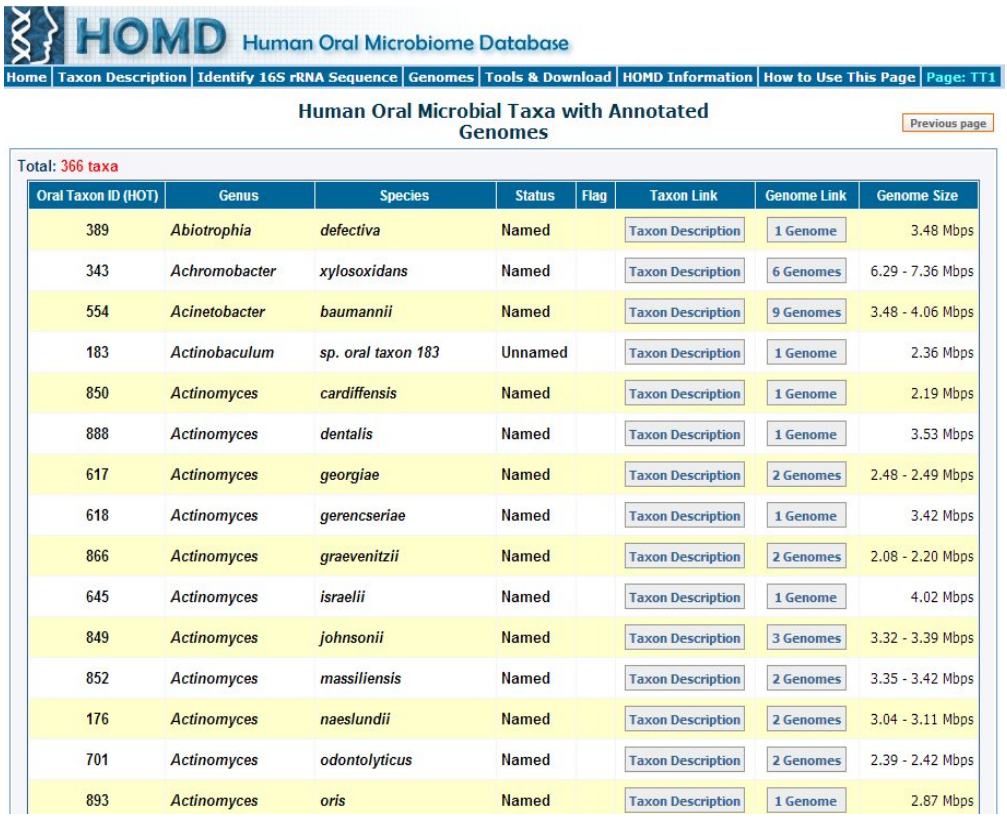


Figure 4.1: Showing the project workflow.

Many microorganisms are associated with the causation of dental caries and this study tries to predict vaccine candidates that would provide protective immunity against a number of microbes. The literature search has to be performed so as to cover as many microorganisms as possible for this study. So, the initial step was to identify as many cariogenic microorganisms reported in literature. These microorganisms may be present in the oral cavity at any phase of the dental caries, i.e., early stage or late stage.

4.1 IDENTIFICATION OF CARIOGENIC MICROORGANISMS

- Human Oral Microbiome Database (HOMD) (available at: http://www.homd.org/) provides a comprehensive list of microorganisms found in the oral cavity. Literature search was performed to select microorganisms associated with dental caries.
- Keywords like microorganism_name AND caries, microorganism_name AND cariogen etc were used to mine the PubMed database (available at: http://www.ncbi.nlm.nih.gov/pubmed/).



HOMD Human Oral Microbiome Database

Home | Taxon Description | Identify 16S rRNA Sequence | Genomes | Tools & Download | HOMD Information | How to Use This Page | Page: TT1

Human Oral Microbial Taxa with Annotated Genomes

Previous page

Total: 366 taxa

Oral Taxon ID (HOT)	Genus	Species	Status	Flag	Taxon Link	Genome Link	Genome Size
389	*Abiotrophia*	*defectiva*	Named		Taxon Description	1 Genome	3.48 Mbps
343	*Achromobacter*	*xylosoxidans*	Named		Taxon Description	6 Genomes	6.29 - 7.36 Mbps
554	*Acinetobacter*	*baumannii*	Named		Taxon Description	9 Genomes	3.48 - 4.06 Mbps
183	*Actinobaculum*	*sp. oral taxon 183*	Unnamed		Taxon Description	1 Genome	2.36 Mbps
850	*Actinomyces*	*cardiffensis*	Named		Taxon Description	1 Genome	2.19 Mbps
888	*Actinomyces*	*dentalis*	Named		Taxon Description	1 Genome	3.53 Mbps
617	*Actinomyces*	*georgiae*	Named		Taxon Description	2 Genomes	2.48 - 2.49 Mbps
618	*Actinomyces*	*gerencseriae*	Named		Taxon Description	1 Genome	3.42 Mbps
866	*Actinomyces*	*graevenitzii*	Named		Taxon Description	2 Genomes	2.08 - 2.20 Mbps
645	*Actinomyces*	*israelii*	Named		Taxon Description	1 Genome	4.02 Mbps
849	*Actinomyces*	*johnsonii*	Named		Taxon Description	3 Genomes	3.32 - 3.39 Mbps
852	*Actinomyces*	*massiliensis*	Named		Taxon Description	2 Genomes	3.35 - 3.42 Mbps
176	*Actinomyces*	*naeslundii*	Named		Taxon Description	2 Genomes	3.04 - 3.11 Mbps
701	*Actinomyces*	*odontolyticus*	Named		Taxon Description	2 Genomes	2.39 - 2.42 Mbps
893	*Actinomyces*	*oris*	Named		Taxon Description	1 Genome	2.87 Mbps

Figure 4.1.1: Showing Human oral microbial taxa with annotated genomes.

As *S.mutans* is the main etiological agent in dental caries (Loesche WJ, 1986), so we made sure that each vaccine candidates predicted in this study must elicit immune response against *S.mutans.* Consequently, the genome of *S. mutans* strain UA159 was taken as the reference genome; *S. mutans* strain UA159 is the first *S.mutans* genome to be sequenced (Song *et al*., 2013).

4.2 RETRIEVAL OF PROTEIN SEQUENCES OF *S.MUTANS* FROM NCBI

- NCBI's GENOME database (http://www.ncbi.nlm.nih.gov/genome/) was searched with the keyword "*Streptococcus mutans*".
- The genome information of *S.mutans* strain UA159 with the accession number NC_004350 was downloaded.
- The sequences of all the proteins of *S.mutans* were downloaded in batch from NCBI using the BATCH ENTREZ tool (http://www.ncbi.nlm.nih.gov/sites/batchentrez).

Since proteins/peptides which are either surface exposed or secreted by the cell are potentially immunogenic, the proteins which are localized either on surface or secreted by *S.mutans* were selected. This step was performed by predicting sub-cellular localization of all the proteins.

4.3 PREDICTION OF SUB-CELLULAR LOCALIZATION OF PROTEINS

- PSORTb (http://www.psort.org/psortb/), CELLO (http://cello.life.nctu.edu.tw/) and Gpos-mPLoc (http://www.csbio.sjtu.edu.cn/bioinf/Gpos-multi/) were used to predict the sub-cellular location of *S.mutans* proteins retrieved from NCBI.
- The proteins predicted to be extracellular or surface exposed by all the three servers were selected for the next step.
- Steps Performed:

1. Protein sequence in fasta format was given as input in PSORTb, CELLO and Gpos-mPLoc.
2. Prediction parameters specific for gram-positive bacteria were selected, e.g., in PSORTb, gram stain was selected as Positive.
3. The protein sequence was submitted for sub-cellular location prediction.

4.3.1 Protein Localization prediction by PSORTb

PSORTb, an SVM based classifier, predicts the localization score of a protein in four different locations, namely cytoplasmic, cytoplasmic membrane, cell wall, extracellular (Yu *et al*., 2010). The location having the highest score is the predicted localization of a protein.

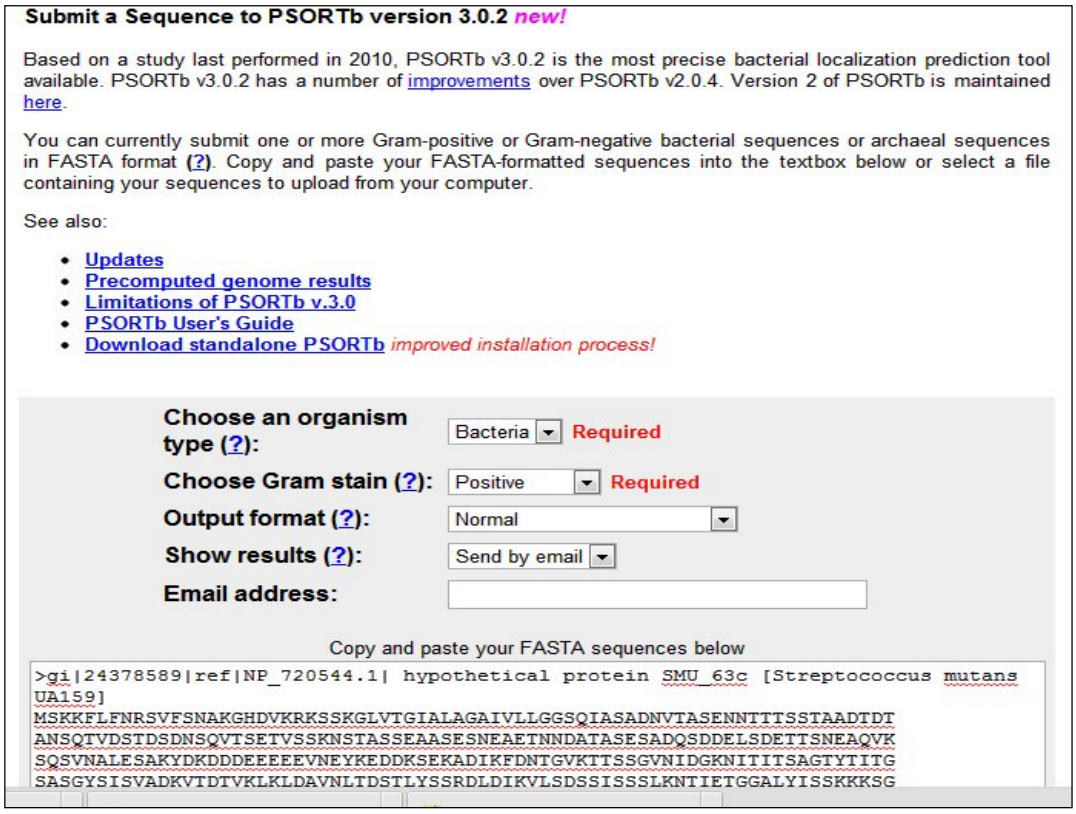


Figure 4.3.1 : Showing protein sequence submission in PSORTb .

4.3.2 Protein Localization prediction by CELLO

CELLO, a two-level SVM based tool, predicts the subcellular localization of a protein/DNA sequence based on the reliability scores of each sequence for five different features and uses these scores in second level SVM classifiers to generate probability distribution of decisions for possible localizations (Yu *et al*., 2006).

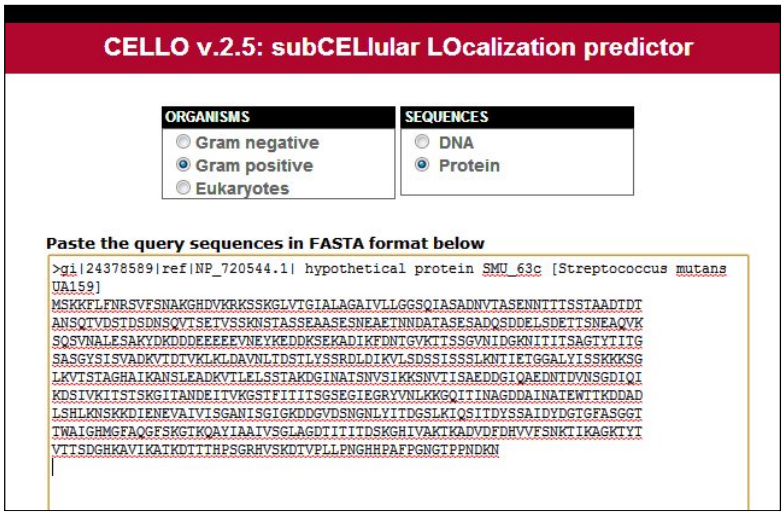


Figure 4.3.2 : Showing protein sequence submission in CELLO.

4.3.3 Protein Localization prediction by Gpos-mPLoc

Gpos-mPLoc predicts the sub-cellular localization of Gram positive bacterial proteins by fusing the information of gene ontology, as well as the functional domain information and sequential evolution information (Shen *et al*., 2009).
Like PSORTb, it also accepts only protein sequences in fasta format.

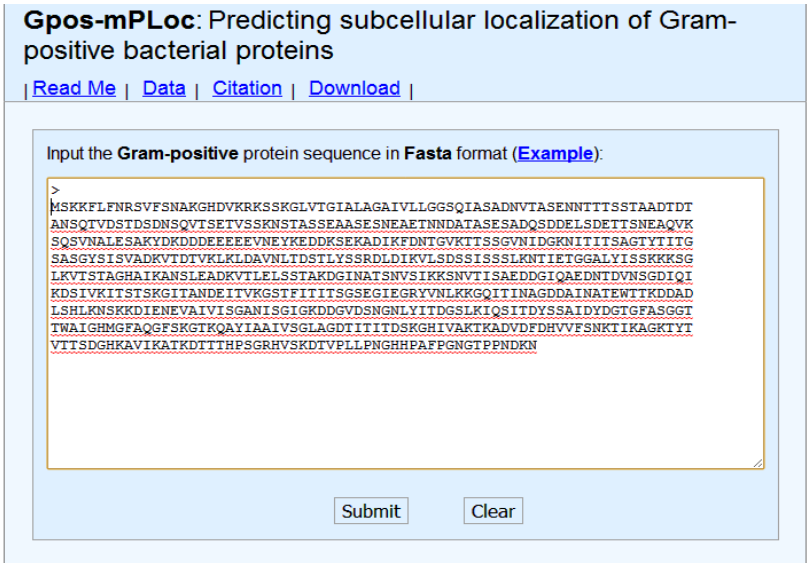


Figure 4.3.3 : Showing protein sequence submission in Gpos-mPLoc.

Proteins which do not have homologs in other cariogenic microorganisms are of no significance in this study as the principal aim of the study is to ensure broad spectrum effectiveness of vaccine. Also the proteins which have homologs in humans cannot be used as vaccines due to cross-reactivity. Consequently, only the proteins which are conseved in cariogenic microorganisms and do not have homologs in humans were selected.

4.4 SELECTION OF SURFACE EXPOSED OR SECRETED PROTEINS WHICH HAVE HOMOLOGS IN OTHER CARIOGENIC MICROORGANISMS BUT NOT IN HUMANS

4.4.1 Selection of *S.mutans* proteins having regions conserved in other cariogenic microorganisms.

- To predict vaccine candidates which can elicit immune response against a number of microorganisms, the selected antigens should be well conserved in all the cariogenic microorganisms. To select such conserved antigens, the homolog's of *S.mutans* proteins selected in Step 4.3 were predicted in other cariogenic bacteria.

- Blast search (available at: http://www.ncbi.nlm.nih.gov/BLAST) was performed using the BLASTP program to find homolog's of the selected *S.mutans* proteins in the microorganisms listed in Appendix I.

4.4.2 Prediction of homolog's in humans

- To induce strong immunity and avoid autoimmunity, predicted antigens must not have sequence similarity to host (e.g., human) proteins.

- To filter out antigens which have sequence similarity to proteins of humans, Blast search was performed against the human genome using the BLASTP program.

  For carrying out this step, we gave GI ids of all the protein sequence as input to the program. Proteins which have homolog's in other cariogenic microorganisms but not in humans were selected for the next step.

Since the major immunological response in oral cavity is of humoral type, the B-cell epitopes were predicted in the selected proteins.

4.5 PREDICTION OF B-CELL EPITOPES IN THE SELECTED PROTEINS

- BCPREDS (http://ailab.ist.psu.edu/bcpreds/predict.html) was used to predict the linear B-cell epitopes (20 amino acids long) in *S.mutans* proteins.
- BCPREDS predicts epitopes using two different methods: (i) BCPRED, which makes predictions based on subsequence kernels (Manzalawy *et al*., 2008) and (ii) AAP, which predicts B-cell epitopes based on amino acid pair antigenicity (Chen *et al*, 2007). BCPREDS also assigns an antigenicity score to each epitope.
- The protein sequence of each protein selected in STEP 4.4 was given as input and the following parameters were selected:

(i) Epitope length of 20 amino acids,

(ii) Specificity of 90% and

(iii) Prediction of only non-overlapping epitopes was selected.

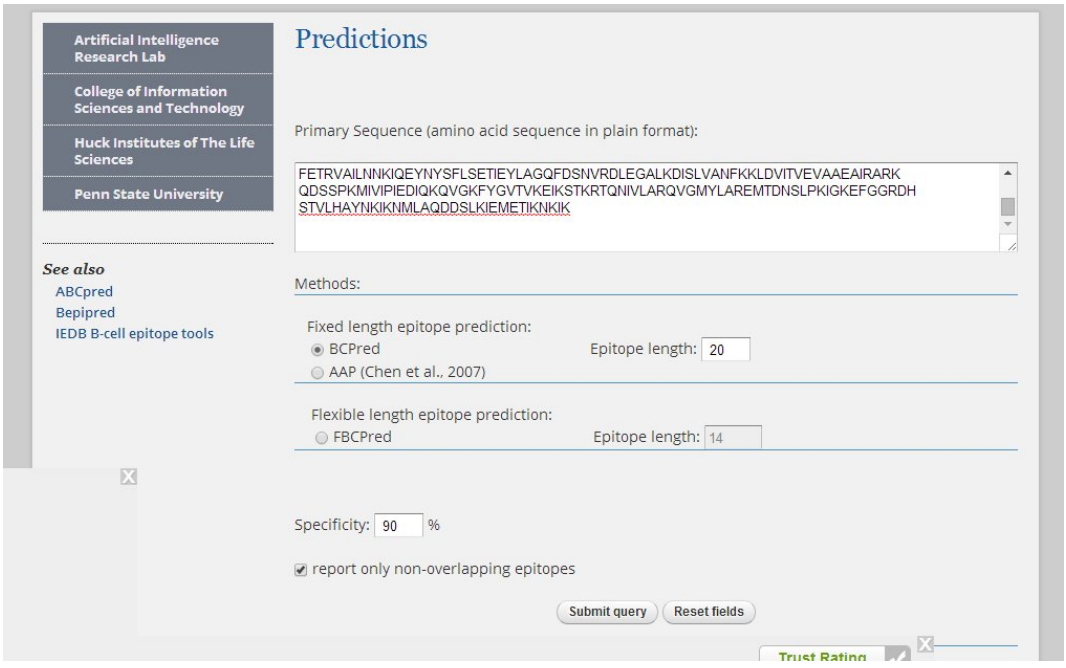


Figure 4.5.1: Showing sequence submission in BCPREDS

Antibodies cannot bind the B-cell epitopes lying in the transmembrane regions. Hence, B-cell epitopes lying in the transmembrane regions cannot be used as vaccine candidates.

4.6 SELECTION OF SURFACE EXPOSED B-CELL EPITOPES

Surface exposed B-cell epitopes were selected on the basis of the following criteria:

(i) The B-cell epitope should not lie in transmembrane region.

- For selecting B-cell epitopes lying in the transmembrane region, TMHMM Server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) was used to predict the trans-membrane regions in the proteins containing the B-cell epitopes.
- TMHMM predicts the residues of a protein lying in the transmembrane region (helices) along with the residues lying inside and outside the cell (Krogh *et al*., 2001).
- The B-cell epitopes lying only in the surface exposed regions (outside the cell) were selected.

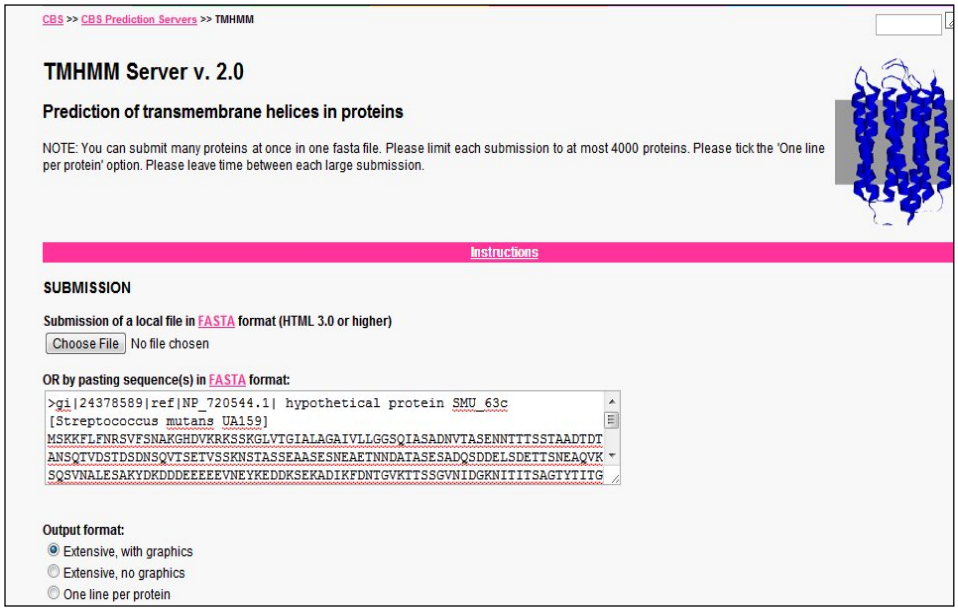


Fig 4.6.1:Showing sequence submission in TMHMM server

(ii) The B-cell epitope should have antigenicity >0.8 (BCPREDS) and >0.4 (VaxiJen)

- The antigenicity of remaining B-cell epitopes was predicted using VaxiJen server (http://www.ddg-pharmfac.net/VaxiJen/VaxiJen/VaxiJen.html).
- VaxiJen predicts the antigenicity of an amino acid sequence solely on the basis of the physicochemical properties of proteins without recourse to sequence alignment. The default threshold value for a bacterial amino acid sequence to be anitgenic is 0.4 (Doytchinova *et al*., 2007).
- Epitope sequence in plain format was given as input in VaxiJen selecting "bacteria" as the target organism.

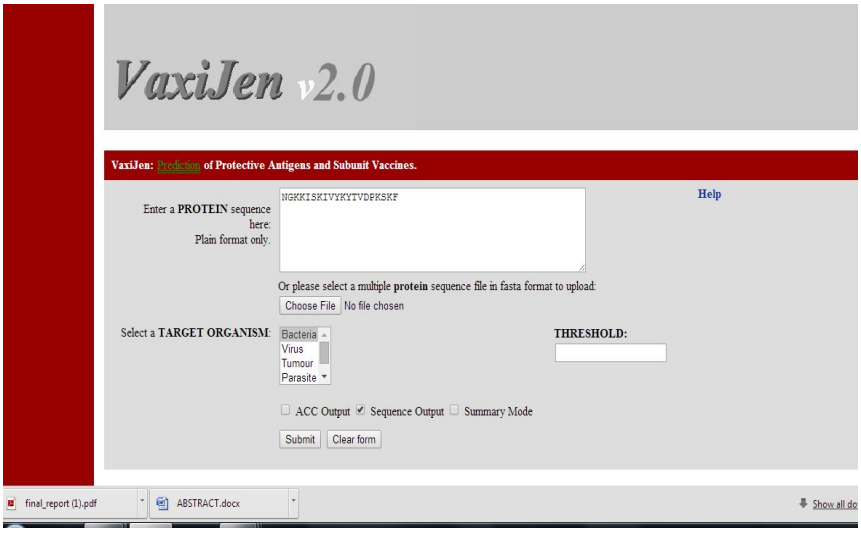


Figure 4.6.2: Showing sequence submission in VaxiJen.

Having all the antigenic B-cell epitopes present in *S.mutans*, now the next step was the selection of epitopes that are conserved in other microorganisms associated with dental caries.

Instead of performing sequence similarity search for the B-cell epitopes, we performed the search for the whole protein sequence. This was done because protein sequence as a whole rather than a small peptide will result in removing false positives from identification of homolog's in other cariogenic microorganisms and thus, the probability of the similar proteins to have the same sub-cellular localization as that of the query protein is quite high.

4.7 SELECTION OF B-CELL EPITOPES THAT ARE CONSERVED IN OTHER CARIOGENIC MICROORGANISMS

- Using BlastP, homologs of *S.mutants* proteins selected in step 4.4 were predicted.
- Sub-cellular localization of all the homologous proteins was predicted using PSORTb, CELLO and Gpos-mPLoc.
- Multiple sequence alignment of each selected protein with all its homologs in other cariogenic microorganism was performed to detect regions of similarity. The program Clustal omega (available at: http://www.ebi.ac.uk/Tools/msa/clustalw2/) was used to perform multiple sequence alignment.
- Clustal omega performs high quality multiple sequence alignments (Sievers *et al*., 2011).
- Out of the B-cell epitopes selected in step 4.6, the epitopes having 9 or more consecutive amino acid residues conserved in other cariogenic microorganisms were selected.

CD4+ T-cells can recognise the antigenic peptides presented by antigen-presenting cells and activate the B-cells, which produce IgG antibody. Hence, CD4+ T-cells activation plays a very important role in eliciting immune response in oral cavity.

Not every administered vaccine is capable of eliciting immune response as it may be degraded by proteases. To increase the chances of eliciting an immune response by a vaccine, B-cell epitopes have been selected in a manner that T-cell epitopes can be derived from these epitopes only, i.e., the B-cell epitopes contain T-cell epitopes also so that even if the B-cell epitope is degraded by proteases, there are chances that T-cell epitopes may elicit immune response.

As 15 amino acid long T-cell epitopes are efficient in stimulating CD4+ T-cells (Darzynkiewicz *et al.*, 2004; Holland *et al.*, 2013), the B-cell epitopes having >14 consecutive amino acid residues conserved in more than 2 cariogens have been used in this step.

4.8 PREDICTION OF T-CELL EPITOPES FROM B-CELL EPITOPES

- 15 amino acids long T-cell epitopes have been predicted from the B-cell epitopes.
- NetMHCIIpan server 2.0 (available at: http://www.cbs.dtu.dk/services/NetMHCIIpan-2.0/) was used to predict the binding of T-cell epitopes to 655 HLA-DR alleles.
- This server takes peptide sequences in fasta format as input and allows the user to select the HLA-DR alleles for which prediction is to be made. A maximum of 15 alleles can be selected per submission (Nielsen *et al.*, 2010).

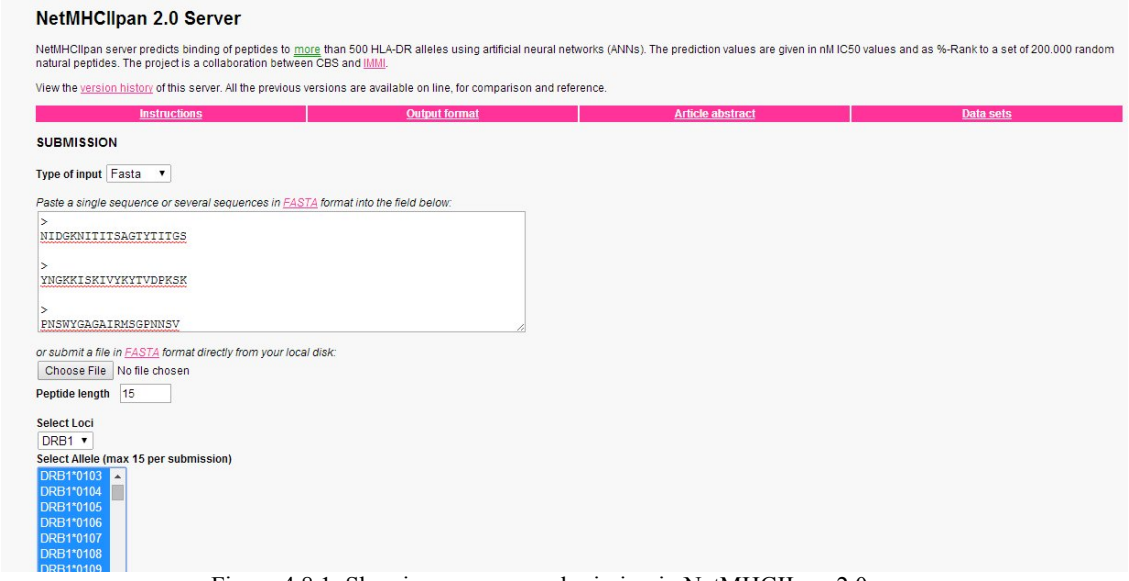


Figure 4.8.1: Showing sequence submission in NetMHCIIpan 2.0 server.

The next step is the selection of antigenic T-cell epitopes which bind to the maximum number of HLA-DR alleles. HLA-DRB1*0101, HLA-DRB1*0301, HLA-DRB1*0401, HLA-DRB1*0701, HLA-DRB1*1101 and HLA-DRB1*1501 are the most frequently occurring alleles in the human population (Panigada *et al.*, 2002). Therefore, T-cell epitopes binding to these most frequently occurring alleles have been selected so as to ensure maximum population coverage.

HLA-DRB1*0101 is the commonest bound allele (Eaper BR, 2009), therefore the epitopes interacting with this allele should produce better antigenic responses.

4.9 SELECTION OF T-CELL EPITOPES

- The antigenicity of each T-cell epitope predicted in step 4.8 was predicted using VaxiJen.
- The T-cell epitopes :

(a) Having VaxiJen antigenic score > 0.4, and
(b) Binding to HLA-DRB1*0101 with IC_{50} value < 100nM, and
(c) Binding to maximum number of alleles among 655 HLA-DR alleles listed on NetMHCIIpan server 2.0 with IC50 value<100nM were selected.

Alignment of the epitopes on the structure of the protein containing the epitopes can be used for confirming that the location of epitopes in the protein.

4.10 CONFIRMATION OF EXO-MEMBRANE TOPOLOGY OF SELECTED EPITOPES

To confirm the exomembrane topology of the selected epitope:

- Sequence based search was performed at RCSB Protein Data Bank (available at: http://www.rcsb.org/pdb/search/advSearch.do) for the protein containing the T-cell epitope selected in step 4.9.
- The PDB structure having maximum similarity to the query protein was selected.
- Using the Pepsurf (Mayrose *et al*., 2006) algorithm available at Pepitope server (available at: http://pepitope.tau.ac.il/), the epitopes were aligned on the selected PDB structure.
- Pepitope server does not accept epitope sequences longer than 14 amino acids. So, the selected B-cell epitopes were entered as seven overlapping 14 amino acid peptides.

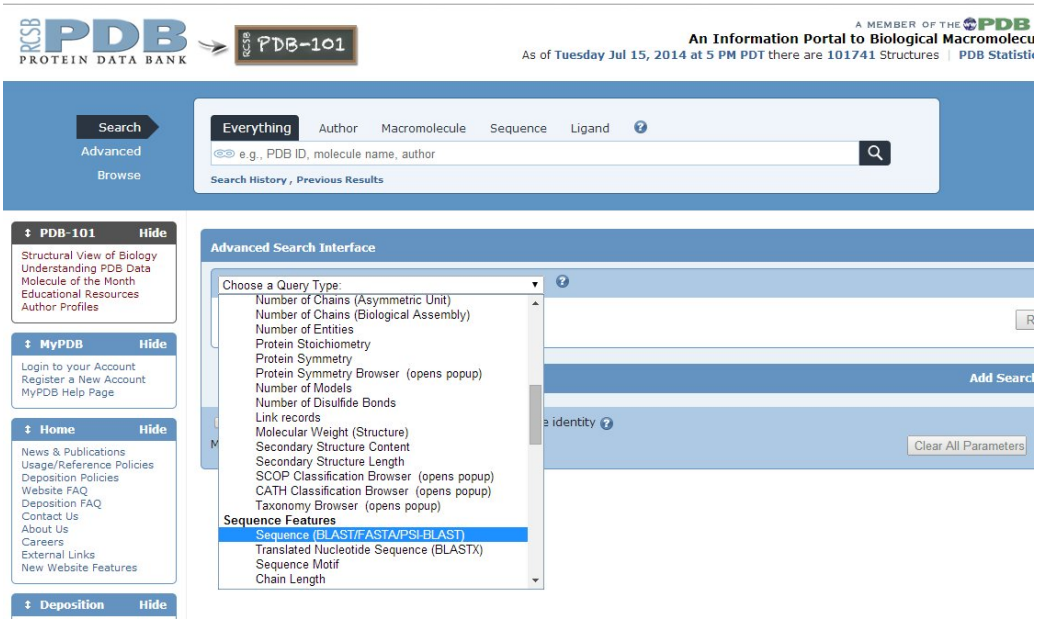


Figure 4.10.1: Showing advanced search option at PDB.

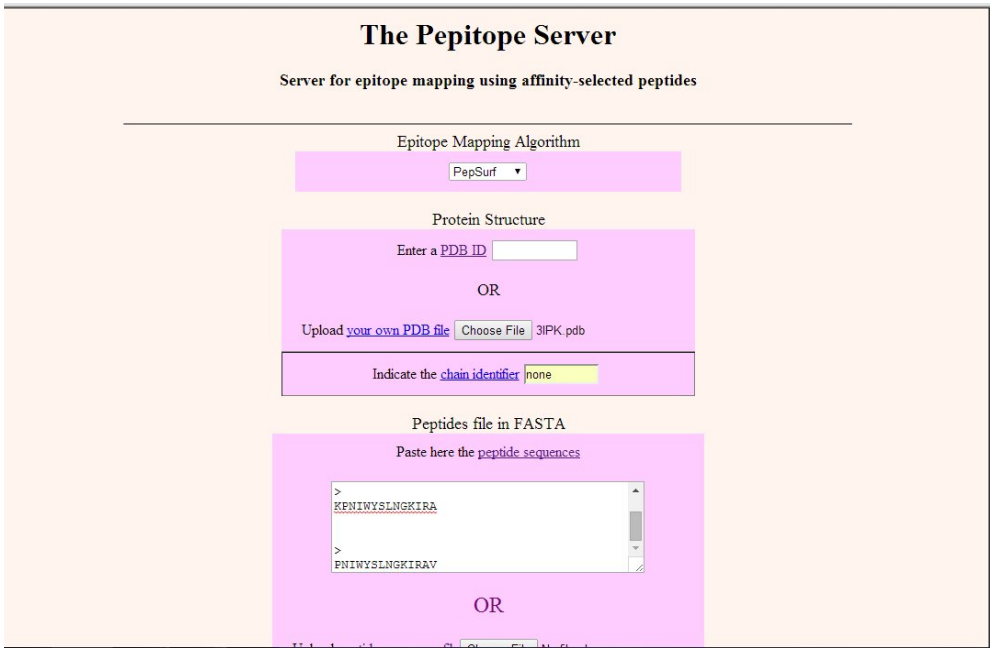


Figure 4.10.2: Showing PDB structure and sequence submission at Pepitope server.

Anti-sera from rabbits immunized with *S.mutans* is known to cross-react with human heart tissue. In order to prevent such cross-reactivity, the predicted vaccine candidates are analysed for having similarity to host proteins (Ferretti *et al.*, 1980).

4.11 CONFIRMATION OF NON CROSS-REACTING NATURE OF PREDICTED VACCINE CANDIDATES TO HOST

- BlastP search was performed against Human genome using the vaccine candidates selected in step 4.9 as query.

5. RESULTS

5.1 IDENTIFICATION OF CARIOGENIC MICROORGANISMS

From literature, 69 microorganisms (including both gram positive and gram negative bacteria) were found to be cariogenic. These microorganisms along with 17 *S.mutans* strains whose genome sequence is available at NCBI have been listed in Appendix I.

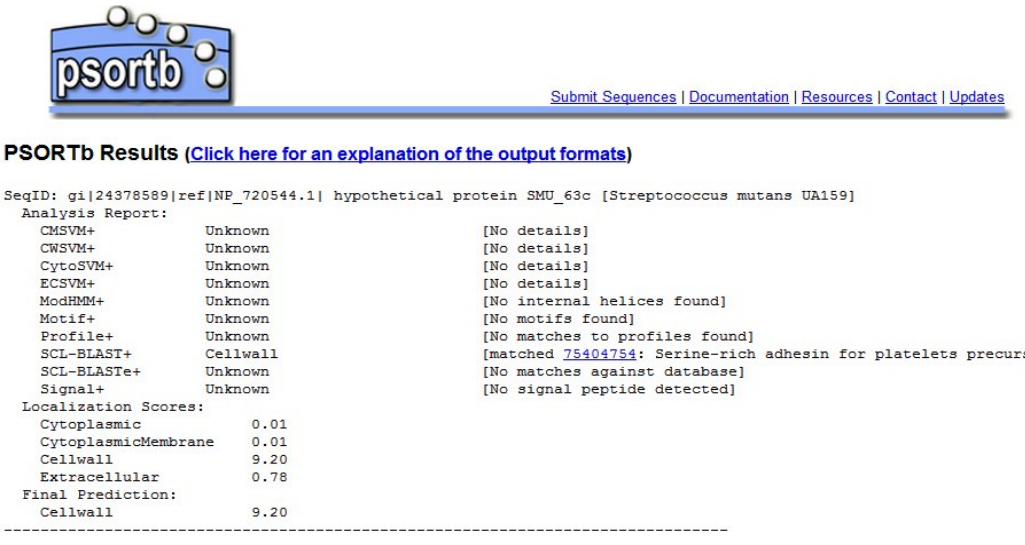
5.2 RETRIEVAL OF PROTEIN SEQUENCES OF *S.mutans* UA159 FROM NCBI

- As *S.mutans* UA159 genome is the first fully sequenced *S.mutans* genome (Song *et al*., 2013), more information is available about this genome compared to the other strains of *S.mutans*.
- 1960 amino acid sequences encoding the entire proteome of *S.mutans* UA159 were retrieved from NCBI.

5.3 PREDICTION OF SUB-CELLULAR LOCALIZATION OF PROTEINS

The whole set of proteins of *S.mutans* was then screened via different protein localization prediction servers/tools so as to mine out the proteins which could behave as antigens, i.e., the proteins/ peptides that are either surface exposed (present on the cell wall) or secreted by the cell.

- Protein localization was predicted using 3 different servers, so as to minimize false positives in the result.
- Out of 1960, 26 proteins were predicted to be either surface exposed or secreted by all the 3 servers. These proteins have been listed in Appendix II.



psortb

Submit Sequences | Documentation | Resources | Contact | Updates

PSORTb Results (Click here for an explanation of the output formats)

```
SeqID: gi|24378589|ref|NP_720544.1| hypothetical protein SMU_63c [Streptococcus mutans UA159]
  Analysis Report:
    CMSVM+            Unknown                    [No details]
    CWSVM+            Unknown                    [No details]
    CytoSVM+          Unknown                    [No details]
    ECSVM+            Unknown                    [No details]
    ModHMM+           Unknown                    [No internal helices found]
    Motif+            Unknown                    [No motifs found]
    Profile+          Unknown                    [No matches to profiles found]
    SCL-BLAST+        Cellwall                   [matched 75404754: Serine-rich adhesin for platelets precursor
    SCL-BLASTe+       Unknown                    [No matches against database]
    Signal+           Unknown                    [No signal peptide detected]
  Localization Scores:
    Cytoplasmic           0.01
    CytoplasmicMembrane   0.01
    Cellwall              9.20
    Extracellular         0.78
  Final Prediction:
    Cellwall              9.20
--------------------------------------------------------------------------------
```

Figure 5.3.1 : Showing subcellular localization prediction by PSORTb.

CELLO RESULTS

SeqID: gi|24378589|ref|NP_720544.1| hypothetical protein SMU_63c [Streptococcus mutans UA159]

Analysis Report:

SVM	LOCALIZATION	RELIABILITY
Amino Acid Comp.	Membrane	0.825
N-peptide Comp.	Extracellular	0.652
Partitioned seq. Comp.	Extracellular	0.645
Physico-chemical Comp.	Extracellular	0.759
Neighboring seq. Comp.	Extracellular	0.617

CELLO Prediction:

Extracellular	2.791 *
Membrane	1.684
CellWall	0.434
Cytoplasmic	0.090

Figure 5.3.2 : Showing subcellular localization prediction by CELLO.

Your input sequence (613aa) is:

```
>
MSKKFLFNRSVFSNAKGHDVKRKSSKGLVTGIALAGAIVLLGGSQIASADNVTASENNTT
TSSTAADTDTANSQTVDSTDSDNSQVTSETVSSKNSTASSEAASESNEAETNNDATASES
ADQSDDELSDETTSNEAQVKSQSVNALESAKYDKDDDEEEEEVNEYKEDDKSEKADIKFD
NTGVKTTSSGVNIDGKNITITSAGTYTITGSASGYSISVADKVTDTVKLKLDAVNLTDST
LYSSRDLDIKVLSDSSISSSLKNTIETGGALYISSKKKSGLKVTSTAGHAIKANSLEADK
VTLELSSTAKDGINATSNVSIKKSNVTISAEDDGIQAEDNTDVNSGDIQIKDSIVKITST
SKGITANDEITVKGSTFITITSGSEGIEGRYVNLKKGQITINAGDDAINATEWTTKDDAD
LSHLKNSKKDIENEVAIVISGANISGIGKDDGVDSNGNLYITDGSLKIQSITDYSSAIDY
DGTGFASGGTTWAIGHMGFAQGFSKGTKQAYIAAIVSGLAGDTITITTDSKGHIVAKTKAD
VDFDHVVFSNKTIKAGKTYTVTTSDGHKAVIKATKDTTTHPSGRHVSKDTVPLLPNGHHP
AFPGNGTPPNDKN
```

---------- Gpos-mPLoc Computation Result ----------

Query protein	Predicted location(s)
	Extracell.

Figure 5.3.3 : Showing subcellular localization prediction by Gpos-mPLoc

5.4 SELECTION OF SURFACE EXPOSED OR SECRETED PROTEINS WHICH HAVE HOMOLOGS IN OTHER CARIOGENIC MICROORGANISMS BUT NOT IN HUMANS

5.4.1 Selection of *S.mutans* proteins having regions conserved in other cariogenic microorganisms.

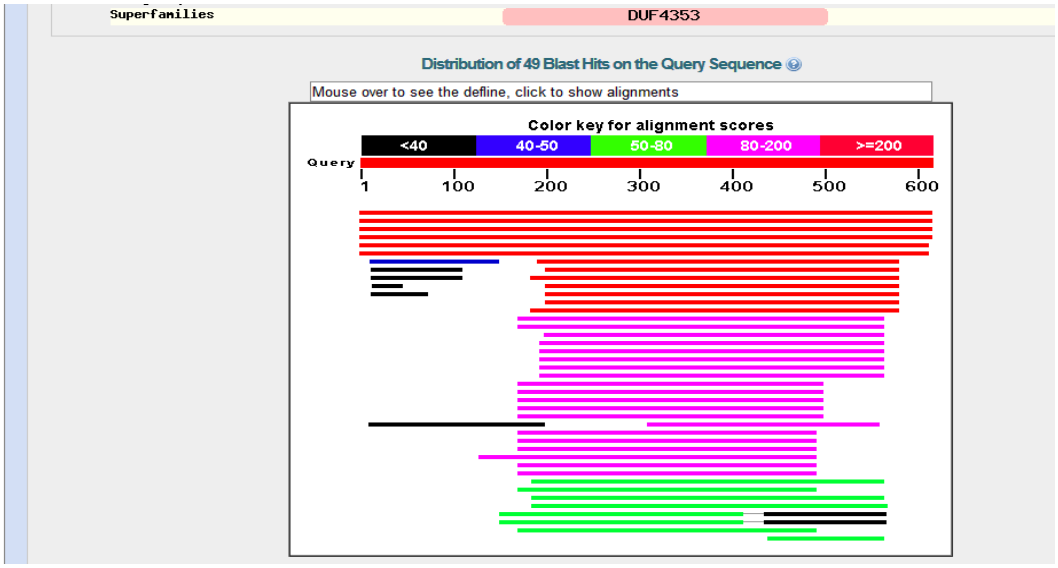


Figure 5.4.1: Showing BLAST hits when *S.mutans* proteins were blasted against 70 cariogens.

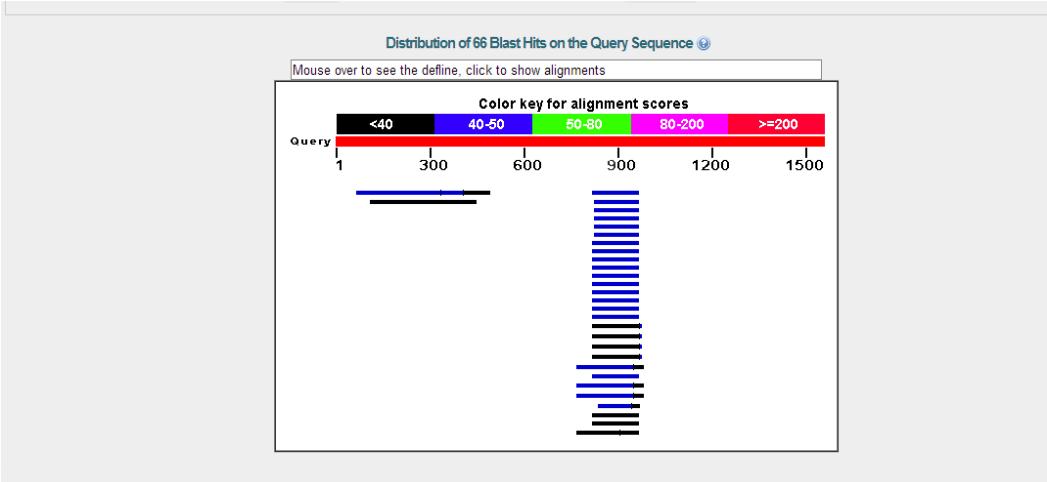


Figure 5.4.2: Showing BLAST hits when *S.mutans* proteins were blasted against human genome (taxid 9606).

Using BlastP, only 16 proteins of *S.mutans* out of 26 proteins were found to have regions conserved in other cariogens also. These 16 proteins have been listed in TABLE 5.4.1

TABLE 5.4.1 showing the proteins having homologs in other cariogens.

S.No.	GI number	Protein name	Localization
1	24378589	hypothetical protein SMU_63c	Cellwall
2	24378602	exo-beta-D-fructosidase; fructanase FruA	Cellwall
3	24378603	exo-beta-D-fructosidase FruB	Cellwall
4	24378708	transfer protein	Cellwall
5	24379087	cell surface antigen SpaP	Cellwall
6	24379801	glucan-binding protein GbpC	Cellwall
7	24379528	cell wall protein, WapE	Cellwall
8	24380381	dextranase	Cellwall
9	24379091	hypothetical protein SMU_616	Extracellular
10	24379231	glucan-binding protein D	Extracellular
11	24379291	hypothetical protein SMU_836	Extracellular
12	24379358	glucosyltransferase-S	Extracellular
13	24379444	glucosyltransferase-I	Extracellular
14	24379445	glucosyltransferase-SI	Extracellular
15	24380370	beta-D-fructosyltransferase	Extracellular
16	24380444	glucan-binding protein GbpA	Extracellular

This step served to remove the *S.mutans* proteins which do not have homologs in other cariogenic microorganisms and thus, are not suitable as vaccine targets for other cariogenic microorganisms, in the initial phase of this study.

5.4.2 Prediction of homologs in humans

Using BlastP, no human homologs were found for the proteins enlisted in TABLE 5.4.1. As a result, all the proteins listed in TABLE 5.4.1 were used in the next step.

This step ensured that none of the selected proteins of *S.mutans* has similarity to the human proteins. Hence, no cross-reactivity should be observed for the vaccine candidates derived from these proteins.

5.5 PREDICTION OF B-CELL EPITOPES IN THE SELECTED PROTEINS

- BCPREDS predicted a total of 349 B-cell epitopes in the 16 proteins listed in TABLE 5.4.1.
- The protein with GI 24379091 did not contain any B-cell epitope so it was not taken up in the next steps.

Submitted sequence: 613 amino acids
Epitope length: 20 amino acids
Classifier Specificty: 90%
Prediction method: bcpred
Use overlap filter: yes

BCPred Predictions

Position	Epitope	Score
154	KDDDEEEEEVNEYKEDDKSE	1
54	ASENNTTTSSTAADTDTANS	1
92	SSKNSTASSEAASESNEAET	1
592	PLLPNGHHPAFPGNGTPPND	1
192	NIDGKNITITSAGTYTITGS	0.991
374	GSTFITITSGSEGIEGRYVN	0.989
344	NSGDIQIKDSIVKITSTSKG	0.975
570	VIKATKDTTTHPSGRHVSKD	0.968
441	GANISGIGKDDGVDSNGNLY	0.967

Figure 5.5.1: Showing B-cell epitopes predicted by BCPREDS

5.6 SELECTION OF SURFACE EXPOSED B-CELL EPITOPES

The surface exposed B-cell epitopes were selected on the basis of :

(i) Transmembrnae topology

- Based on TM-HMM results listed in TABLE 5.6.1, 3 B-cell epitopes were rejected as they were lying in the transmembrane regions.

TABLE 5.6.1 showing the transmembrane regions in the proteins listed in Table 5.4.1.

	TM-HMM PREDICTION RESULTS		
PROTEIN GI	**INSIDE**	**HELIX**	**OUTSIDE**
24378589			1-613
24378602	1-19	20-39	40-1423
24378603			1-519
24378708	1-16	17-39	40-365
24379087	1558 – 1562	1538 -1557	1-1537
24379801			1-583
24379528	1-12; 501-507	13-35; 483-500	36-482
24380381	845-850	827-844	1-826
24379231			1-726
24379291			1-544
24379358			1-1462
24379444			1-1476
24379445			1-1455
24380370	1--12; 795	13-35; 777-794	36-776
24380444	1-16	17-39	40-565

(ii) Antigenicity of B-cell epitopes

- Based on VaxiJen scores, 69 B-cell epitopes have antigenicity score below the set threshold of 0.4.
- So, said 69 epitopes were rejected.

In this step, 72 B-cell epitopes (three lying in transmembrane region and 69 non-antigenic) were rejected. The remaining epitopes were checked for conservancy in the next step.

5.7 SELECTION OF B-CELL EPITOPES THAT HAVE REGIONS CONSERVED IN OTHER CARIOGENIC MICROORGANISMS

- By aligning each protein listed in TABLE 5.4.1 (except 24379091) with its homologs in other cariogenic microorganisms, B-cell epitopes having 9 or more than 9 consecutive residues conserved in 3 or more than 3 cariogenic microorganisms (including *S.mutans*) were selected.

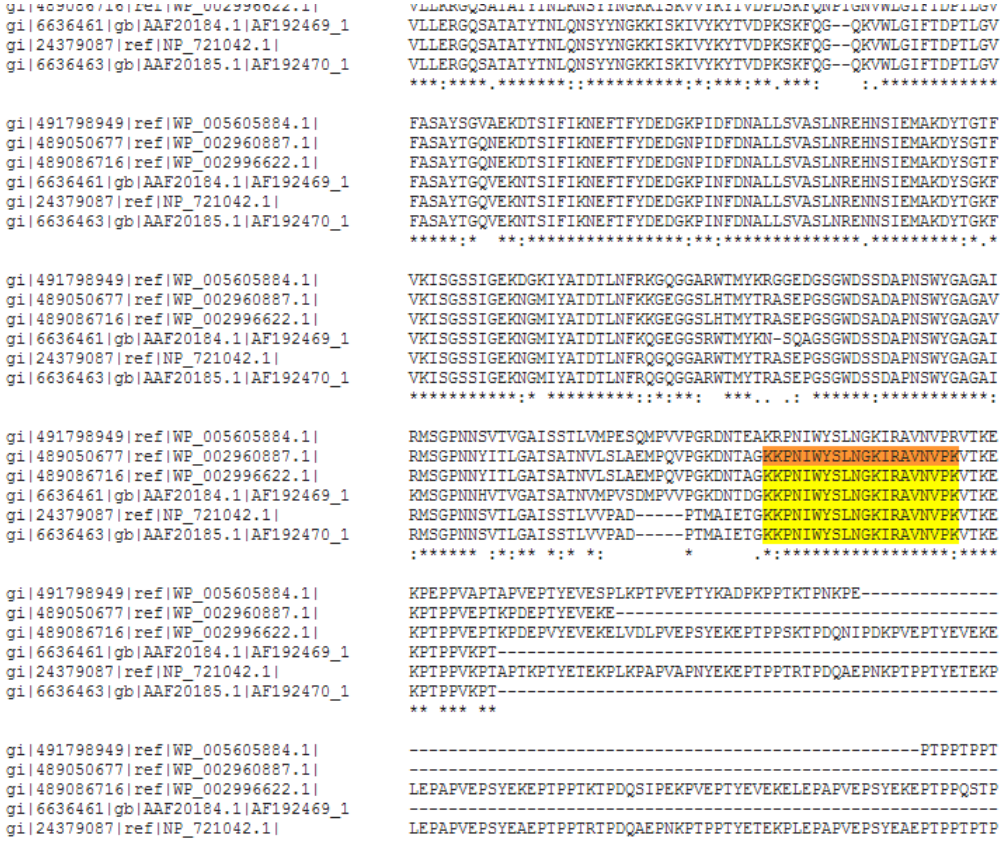


Figure 5.7.1: Showing BlastP results for conserved B-cell epitope in cariogens (out of 70)

Figure 5.7.2: Showing conserved B-cell epiotpe in all the 17 S.mutans strains

- The regions of B-cell epitopes conserved in 3 or more than 3 microorganisms have been listed in TABLE 5.7.1.
- A total of 26 B-cell epitopes having conserved regions (9-20 amino acids long) in other cariogenic microorganisms also were identified.
- Out of 26 B-cell epitopes, 6 had 15 or more amino acids conserved in more than 2 cariogens. These 6 epitopes have been selected for further analysis.

TABLE 5.7.1 showing the B-cell epitopes conserved in 3 or more than 3 cariogens. (*Highlighted in Grey colour-Epitopes have >14 consecutive amino acid residues conserved)

Protein GI	**method of prediction**	**CONSERVED IN**	**conserved region**	**starting position**	**b-cell epitope**	**BCPREDS SCORE**	**VaxiJen_ SCORE**
24378589	BCPRED	*Streptococcus mutans, Streptococcus downei, Streptococcus sobrinus Streptococcus mutans, Streptococcus downei, Streptococcus sobrinus*	NIDGKNITITSAGTYTITGS	192	NIDGKNITITSAGTYTITGS	0.991	1.112
24378589	BCPRED		EGIEGRYVN	374	GSTFITITSGSEGIEGRYVN	0.989	**1.0925**
24378589	AAP		DGVDSNGNL	441	GANISGIGKDDGVDSNGNLY	0.967	**0.9543**
24378589	AAP		EDNTDVNSGDI	338	EDNTDVNSGDIQIKDSIVKI	1	**1.4819**
24378589	AAP		QGSDYYGANF	287	QGSDYYGANFYQESDCVVKS	1	**0.8651**
24378708	AAP	*Streptococcus mutans, Streptococcus agalactiae, Streptococcus intermedius*	VAVSDGQTK	191	VAVSDGQTKADKIESWATTI	1	**0.8006**
24379087	BCPRED	*Streptococcus mutans, Streptococcus sobrinus, Streptococcus constellatus, Streptococcus intermedius, Granulicatella adiacens, Streptococcus downei*	QKETEIKEDYTKQA	126	EEAVQKETEIKEDYTKQAED	0.995	**1.0097**
24379087	BCPRED		TGNPATNLPEAQGS	50	KVVGTQTGNPATNLPEAQGS	0.992	**0.6476**
24379087	BCPRED		GSGWDSSDAPNS	749	MYTRASEPGSGWDSSDAPNS	0.986	**0.5589**
24379087	BCPRED		NGKKISKIVYKYTVD	618	YNGKKISKIVYKYTVDPKSK	0.981	**0.9420**
24379087	AAP		PNSWYGAGAIRMSGPNN	766	PNSWYGAGAIRMSGPNNSVT	1	**0.4100**
24379087	AAP		KKPNIWYSLNGKIRAVNVPK	807	KKPNIWYSLNGKIRAVNVPK	1	**0.5645**
24379087	AAP		INNVPKINPKKDVTLTLDPA	1324	INNVPKINPKKDVTLTLDPA	1	**0.6662**
24379087	AAP		DPTLGVFASAYTG	650	DPTLGVFASAYTGQVEKNTS	1	**0.5268**
24379087	AAP		KEIRNNNDINIDRT	996	VQPQVNKEIRNNNDINIDRT	1	**0.9242**
24379087	AAP		DLTKSV….TIYPTVVGQ	1088	TFNADLTKSVATIYPTVVGQ	1	**0.5209**
24379087	AAP		SAVDDAFSK	528	LKASAVDDAFSKSTSKAKYD	1	**0.8552**

24379358	AAP	*Streptococcus mutans, Streptococcus sobrinus, Streptococcus salivarius, Streptococcus downei*	NFDGVRVDA	450	IVANDPEANFDGVRVDAVDN	1	**0.8954**
24379358	AAP	*Streptococcus mutans, Streptococcus sobrinus, Streptococcus salivarius, Streptococcus downei*	YRLLNRTPT	393	YRLLNRTPTSQTGKPKYFED	1	**0.4422**
24379358	AAP		GGYDFLLANDIDNSNP	414	SSGGYDFLLANDIDNSNPVV	1	**0.6727**
24379444	BCPRED	*Streptococcus mutans, Streptococcus salivarius, Streptococcus sobrinus, Streptococcus downei*	ANFDSIRVDA	439	NDPDANFDSIRVDAVDNVDA	0.995	**0.7825**
24379444	AAP		ANFDSIRVDA	438	ANDPDANFDSIRVDAVDNVD	1	**0.6663**
24380370	BCPRED	*Streptococcus mutans, Lactobacillus acidophilus, Lactobacillus gasseri, Streptococcus salivarius, Lactobacillus johnsonii, Lactobacillus vaginalis, Lactobacillus jensenii*	DWRTATYSYYAVPV	569	DWRTATYSYYAVPVAGSSDT	0.978	**0.6674**
24380370	AAP		DVWDSWPVQD	243	DVWDSWPVQDAKTGEVINWN	1	**0.7483**
24380370	AAP		ANAAIGILKL	463	ANAAIGILKLKGDKKTPEVD	1	**0.6949**
24380370	AAP		DWRTATYSYYAVPV	565	SVPADWRTATYSYYAVPVAG	1	**0.4562**

5.8 PREDICTION OF T-CELL EPITOPES FROM B-CELL EPITOPES

- 15 amino acids long T-cell epitopes were predicted from 6 B-cell epitopes selected in step 5.7.
- The selected 6 B-cell epitopes contain 24 T-cell epitopes as shown in TABLE 5.8.1.
- Out of these 24 T-cell epitopes, 15 epitopes bind to HLA-DRB1*0101 with IC_{50}<100nM.
- Table 5.8.1 also shows the number of HLA-DR alleles (out of 655 alleles) bound by the predicted T-cell epitopes.

TABLE 5.8.1 showing IC_{50} values of T-cell epitopes for HLA-DR alleles (IC_{50} value < 100nM highlighted in pink)

B-CELL EPITOPE	T-CELL EPITOPE	**DRB1*0101**	**DRB1*0301**	**DRB1*0401**	**DRB1*0701**	**DRB1*1101**	**DRB1*1501**	**No. of HLA-DR alleles bound**
NIDGKNITITSAGTYTITGS	NIDGKNITITSAGTY	831.65	4660.1	1636.62	231.65	2220.47	1939.88	14
	IDGKNITITSAGTYT	349.73	3142.26	835.08	135.84	1652.31	1056.63	29
	DGKNITITSAGTYTI	90.2	1562.74	323.69	33.1	1046.96	372.95	100
	GKNITITSAGTYTIT	58.34	1010.34	215.35	28.7	764.43	256.37	143
	KNITITSAGTYTITG	63.21	1404.75	187.21	36.49	907.18	293.82	108
	NITITSAGTYTITGS	92.75	2216.03	222.76	60.27	1413.07	446.99	72
YNGKKISKIVYKYTVDPKSK	NGKKISKIVYKYTVD	537.95	2090.65	2638.51	323.39	70.18	127.15	64
PNSWYGAGAIRMSGPNNSVT	PNSWYGAGAIRMSGP	25.83	7392.04	613.29	405.32	765.48	1250.05	32
	NSWYGAGAIRMSGPN	37.06	7690.25	682.35	642.43	768.5	1720.31	29
	SWYGAGAIRMSGPNN	65.63	7807.38	932.75	947.49	1113	2062.25	23
KKPNIWYSLNGKIRAVNVPK	KKPNIWYSLNGKIRA	12.09	151.5	96.94	11.22	12.94	17.87	268
	KPNIWYSLNGKIRAV	6.98	112.54	81.82	10.72	9.86	17.23	276
	PNIWYSLNGKIRAVN	8.74	168.34	120.57	20.6	11.71	33.58	250
	NIWYSLNGKIRAVNV	10.36	165.43	163.14	33.67	13.26	50.33	243
	IWYSLNGKIRAVNVP	15.98	313.21	292.28	70.36	17.31	101.83	218
	WYSLNGKIRAVNVPK	53.25	604.47	578.18	268.98	38.24	279.95	162
INNVPKINPKKDVTLTLDPA	INNVPKINPKKDVTL	3445.68	6095.14	6613.04	3410.01	1992.72	4118.08	1
	NNVPKINPKKDVTLT	2915.29	5307.34	6445.35	3204.82	2755	4256.87	0
	NVPKINPKKDVTLTL	1367.8	1440.19	3582.27	1332.35	2945.33	2879.74	0
	VPKINPKKDVTLTLD	1253.08	1140.85	3267.74	1551.69	4346.58	3027.25	0
	PKINPKKDVTLTLDP	1706.14	1094.87	3383.17	1921.4	4992.59	3862.62	0
	KINPKKDVTLTLDPA	1610.51	927.5	2231.32	2108.94	4627.02	4191.67	1
SSGGYDFLLANDIDNSNPVV	GGYDFLLANDIDNSN	7.97	2202.96	125.14	284.9	768.99	603.74	59
	GYDFLLANDIDNSNP	11.59	1806.2	142.65	409.97	1047.14	887.19	53

5.9 SELECTION OF T-CELL EPITOPES

- The antigenicity scores of the T-cells predicted in step 5.8 have been listed in TABLE 5.9.1. The antigens derived from PNSWYGAGAIRMSGPNNSVT were not antigenic.

TABLE 5.9.1 showing antigenicity score of conserved T-cell epitopes (Antigenic epitopes- highlighted in pink)

B-CELL EPITOPE	CONSERVED T-CELL EPITOPE	VAXIJEN SCORE
NIDGKNITITSAGTYTITGS	NIDGKNITITSAGTY	1.0560 (Probable ANTIGEN)
	IDGKNITITSAGTYT	0.9348 (Probable ANTIGEN)
	DGKNITITSAGTYTI	0.6725 (Probable ANTIGEN)
	GKNITITSAGTYTIT	0.5620 (Probable ANTIGEN)
	KNITITSAGTYTITG	0.8169 (Probable ANTIGEN)
	NITITSAGTYTITGS	0.7426 (Probable ANTIGEN)
YNGKKISKIVYKYTVDPKSK	NGKKISKIVYKYTVD	0.8264 (Probable ANTIGEN)
PNSWYGAGAIRMSGPNNSVT	PNSWYGAGAIRMSGP	0.0522 (Probable NON-ANTIGEN)
	NSWYGAGAIRMSGPN	0.2096 (Probable NON-ANTIGEN)
	SWYGAGAIRMSGPNN	0.1211 (Probable NON-ANTIGEN)
KKPNIWYSLNGKIRAVNVPK	KKPNIWYSLNGKIRA	0.5439 (Probable ANTIGEN)
	KPNIWYSLNGKIRAV	0.6667 (Probable ANTIGEN)
	PNIWYSLNGKIRAVN	0.6923 (Probable ANTIGEN)
	NIWYSLNGKIRAVNV	0.9957 (Probable ANTIGEN)
	IWYSLNGKIRAVNVP	0.9172 (Probable ANTIGEN)
	WYSLNGKIRAVNVPK	0.9419 (Probable ANTIGEN)
INNVPKINPKKDVTLTLDPA	INNVPKINPKKDVTL	0.4937 (Probable ANTIGEN)
	NNVPKINPKKDVTLT	0.7325 (Probable ANTIGEN)
	NVPKINPKKDVTLTL	0.7112 (Probable ANTIGEN)
	VPKINPKKDVTLTLD	0.9361 (Probable ANTIGEN)
	PKINPKKDVTLTLDP	1.0605 (Probable ANTIGEN)
	KINPKKDVTLTLDPA	0.6205 (Probable ANTIGEN)
SSGGYDFLLANDIDNSNPVV	GGYDFLLANDIDNSN	0.8558 (Probable ANTIGEN)
	GYDFLLANDIDNSNP	8686 robable ANTIGEN)

- Based on the criteria:

(a) VaxiJen antigenic score > 0.4, and

(b) Binding to HLA-DRB1*0101 with IC_{50} value < 100nM, and

(c) Binding to the maximum number of alleles from 655 HLA-DR alleles listed on NetMHCIIpan server 2.0.

The T-cell epitopes KKPNIWYSLNGKIRA, KPNIWYSLNGKIRAV, PNIWYSLNGKIRAVN, NIWYSLNGKIRAVNV and IWYSLNGKIRAVNVP were selected. All these T-cell epitopes were derived from the B-cell epitope KKPNIWYSLNGKIRAVNVPK.

TABLE 5.9.2 lists the IC_{50} values, number of HLA-DR alleles bound and antigenicity of the selected T-cell epitopes (IC_{50} values <100nM highlighted in pink)

T-cell epitope	DRB1*0101	DRB1*0301	DRB1*0401	DRB1*0701	DRB1*1101	DRB1*1301	DRB1*1501	No. of HLA-DR alleles bound	VaxiJen score
KKPNIWYSLNGKIRA	12.09	151.5	96.94	11.22	12.94	12.01	17.87	268	0.5439
KPNIWYSLNGKIRAV	6.98	112.54	81.82	10.72	9.86	10.37	17.23	276	0.6667
PNIWYSLNGKIRAVN	8.74	168.34	120.57	20.6	11.71	19.06	33.58	250	0.6923
NIWYSLNGKIRAVNV	10.36	165.43	163.14	33.67	13.26	33.34	50.33	243	0.9957
IWYSLNGKIRAVNVP	15.98	313.21	292.28	70.36	17.31	82.34	101.83	218	0.9172

5.10 CONFIRMATION OF EXO-MEMBRANE TOPOLOGY OF SELECTED EPITOPES

- Sequence based search at RCSB Protein Data Bank (available at: http://www.rcsb.org/pdb/search/advSearch.do) for the protein containing the T-cell epitopes selected in step 5.9 retrieved the PDB structure 3IPK as having 93.7% sequence similarity to the protein with GI 24379087.
- As the B-cell epitope KKPNIWYSLNGKIRAVNVPK is conserved in 3IPK, this structure was selected to confirm the exo-membrane topology of the selected epitopes.
- Using the Pepitope server (available at: http://pepitope.tau.ac.il/), the entire B-cell epitope containing the selected T-cell epitopes was aligned on the selected PDB structure.
- Pepitope predicted that all the seven 14 amino acids peptides derived from KKPNIWYSLNGKIRAVNVPK belong to a single cluster (Score= 92.564, no. of residues=28). The alignment score and P-value for each alignment are given in TABLE 5.10.1.
- The B-cell epitope KKPNIWYSLNGKIRAVNVPK (pink) aligned on the structure 3IPK has been shown in Figure 5.10.1.

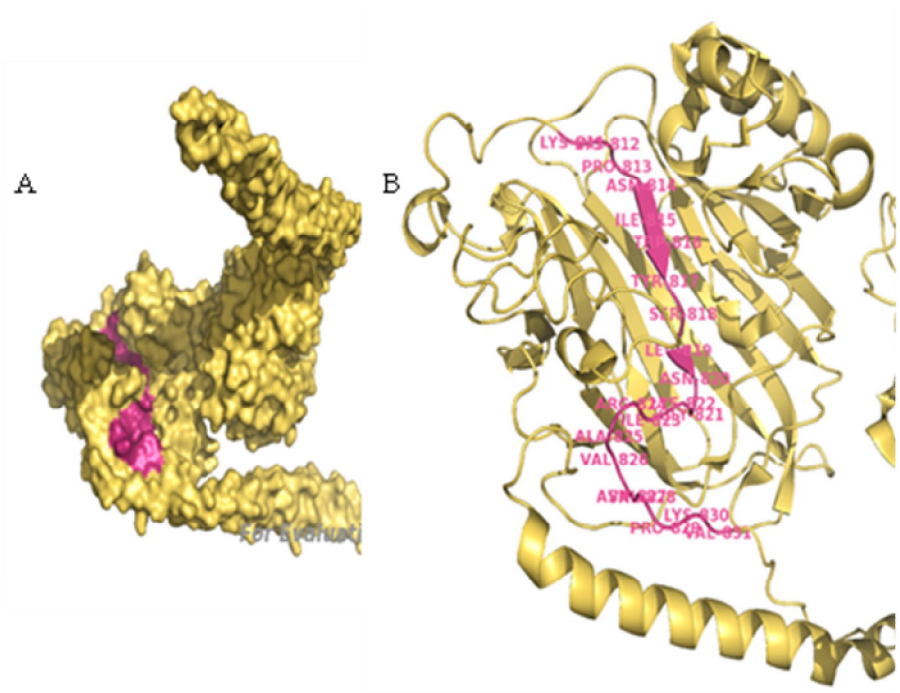


Figure 5.10.1: The alignment of epitope (pink) on 3D-structure of protein 3IPK (A). The residues of the aligned epitope on the structure (B).

TABLE 5.10.1 shows the alignments scores and P-value for all the overlapping peptides derived from KKPNIWYSLNGKIRAVNVPK

PEPTIDE	**SCORE**	**P-value**
KKPNIWYSLNGKIR	27.6583	1.15945e-05
KPNIWYSLNGKIRA	28.3635	2.78835e-06
PNIWYSLNGKIRAV	27.3149	1.82643e-06
NIWYSLNGKIRAVN	29.513	1.29045e-06
IWYSLNGKIRAVNV	28.276	7.7406e-07
WYSLNGKIRAVNVP	28.141	8.9188e-07
YSLNGKIRAVNVPK	29.6136	1.19896e-06

5.11 CONFIRMATION OF NON CROSS-REACTING NATURE OF PREDICTED VACCINE CANDIDATES TO HOST

- The selected vaccine candidate KKPNIWYSLNGKIRAVNVPK was used as a query and BLASTP was performed against human genome.
- The best hit had a score of 24.7 with E-value of 1.9 (insignificant)

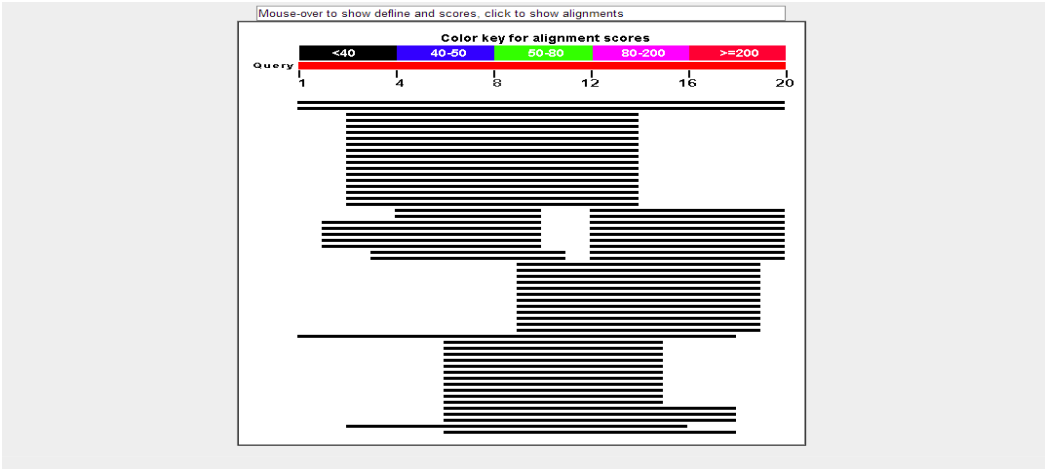
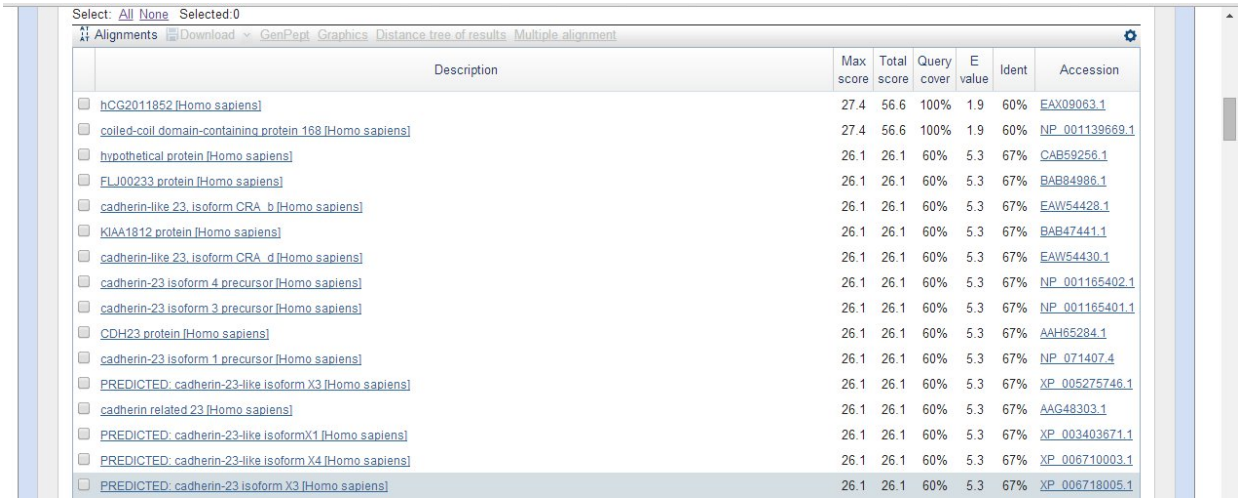


Figure 5.11.1: Showing blast hits retrieved when the selected B-cell epitope was blasted with human genome.


Select: All None Selected:0

Alignments Download GenPept Graphics Distance tree of results Multiple alignment

Description	Max score	Total score	Query cover	E value	Ident	Accession
hCG2011852 [Homo sapiens]	27.4	56.6	100%	1.9	60%	EAX09063.1
coiled-coil domain-containing protein 168 [Homo sapiens]	27.4	56.6	100%	1.9	60%	NP_001139669.1
hypothetical protein [Homo sapiens]	26.1	26.1	60%	5.3	67%	CAB59256.1
FLJ00233 protein [Homo sapiens]	26.1	26.1	60%	5.3	67%	BAB84986.1
cadherin-like 23, isoform CRA_b [Homo sapiens]	26.1	26.1	60%	5.3	67%	EAW54428.1
KIAA1812 protein [Homo sapiens]	26.1	26.1	60%	5.3	67%	BAB47441.1
cadherin-like 23, isoform CRA_d [Homo sapiens]	26.1	26.1	60%	5.3	67%	EAW54430.1
cadherin-23 isoform 4 precursor [Homo sapiens]	26.1	26.1	60%	5.3	67%	NP_001165402.1
cadherin-23 isoform 3 precursor [Homo sapiens]	26.1	26.1	60%	5.3	67%	NP_001165401.1
CDH23 protein [Homo sapiens]	26.1	26.1	60%	5.3	67%	AAH65284.1
cadherin-23 isoform 1 precursor [Homo sapiens]	26.1	26.1	60%	5.3	67%	NP_071407.4
PREDICTED: cadherin-23-like isoform X3 [Homo sapiens]	26.1	26.1	60%	5.3	67%	XP_005275746.1
cadherin related 23 [Homo sapiens]	26.1	26.1	60%	5.3	67%	AAG48303.1
PREDICTED: cadherin-23-like isoformX1 [Homo sapiens]	26.1	26.1	60%	5.3	67%	XP_003403671.1
PREDICTED: cadherin-23-like isoform X4 [Homo sapiens]	26.1	26.1	60%	5.3	67%	XP_006710003.1
PREDICTED: cadherin-23 isoform X3 [Homo sapiens]	26.1	26.1	60%	5.3	67%	XP_006718005.1

Figure 5.11.2: Showing the scores and E-values of the retrieved hits.

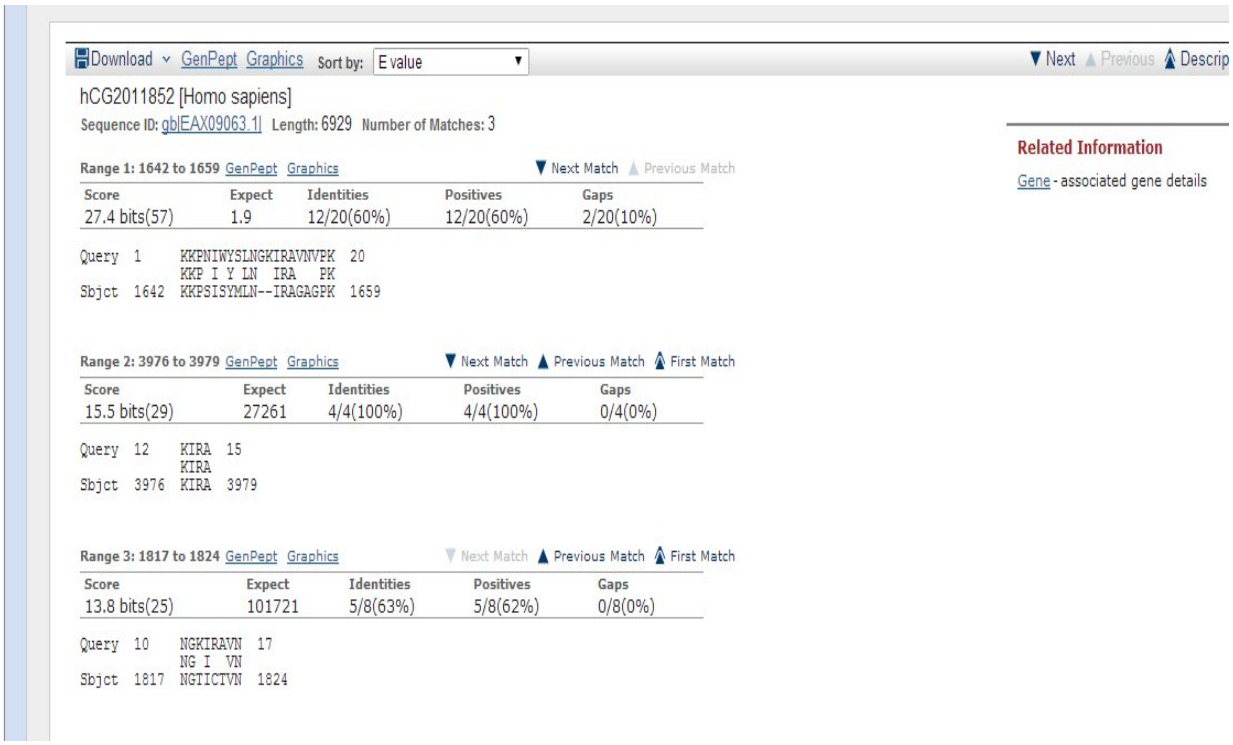


Figure 5.11.3: Showing alignment of the B-cell epitope with the best hit

6. CONCLUSION

This study, directed towards the identification of vaccine candidates for dental caries has resulted in the prediction of probable epitopes that can be used to elicit immune response against a number of microorganisms growing in a biofilm. KKPNIWYSLNGKIRAVNVPK has been recognized as an antigen that can be used as a vaccine against *Streptococcus mutans* (all 17 strains), *Streptococcus sobrinus, Streptococcus constellatus, Streptococcus intermedius, Granulicatella adiacens* and *Streptococcus downei*. This entire epitope is conserved in 5 cariogens, *Streptococcus mutans, Streptococcus sobrinus, Streptococcus constellatus, Streptococcus intermedius* and *Streptococcus downei* as shown in Figures 5.7.1 and 5.7.2.

As shown in TABLE 5.9.2, this B-cell epitope contains six overlapping 15 amino acid long MHC class-II binding peptides. All the six peptides bind to HLA-DRB1*0101, but only KPNIWYSLNGKIRAV and KKPNIWYSLNGKIRA bind to DRB1*0401 also. Though NIWYSLNGKIRAVNV (0.9957) is the most antigenic one, it does not bind to DRB1*0401 and compared to KPNIWYSLNGKIRAV and KKPNIWYSLNGKIRA binds to a lesser number of HLA-DR alleles out of the 655 alleles for which NetMHCIIpan predicts binding. Further, the IC_{50} values for peptide-MHC class II allele binding are lower for KPNIWYSLNGKIRAV compared to KKPNIWYSLNGKIRA and KPNIWYSLNGKIRAV binds to the maximum number of HLA alleles considered in this study.

In the selected vaccine candidates, the best three T-cell epitopes based on the criteria used to select T-cell epitopes are KPNIWYSLNGKIRAV followed by KKPNIWYSLNGKIRA and PNIWYSLNGKIRAVN.

Further, this B-cell epitope does not have any significant similarity to the human proteome. This observation is based on the Blast hits depicted in Figures 5.11.1- 5.11.3. The best hit had a score of 27.4 with E-value of 1.9, which represents insignificant similarity.

This particular B-cell epitope is also found to be surface exposed based on the sequence structure alignment results as shown in Figures 5.10.1-5.10.2.

So, the best B-cell epitope is KKPNIWYSLNGKIRAVNVPK and the best T-cell epitope derived from the said B-cell epitope is KPNIWYSLNGKIRAV as it binds to maximum number of HLA-DR alleles with lowest IC50 values and is capable of eliciting immune response against 5 cariogens. Further as the 17 amino acid stretch PNIWYSLNGKIRAVNVP (0.7512- VaxiJen score) of the selected vaccine candidate KKPNIWYSLNGKIRAVNVPK is also conserved in *Granulicatella adiacens* as shown in Figure 6.1, the selected vaccine candidate can potentially elicit immune response against *Granulicatella adiacens* also.

In conclusion, the best vaccine candidate for targeting a number of cariogens is KKPNIWYSLNGKIRAVNVPK, which contains overlapping T-cell epitopes capable of eliciting immune response against 6 cariogens, *Streptococcus mutans* (all 17 strains), *Streptococcus sobrinus, Streptococcus constellatus, Streptococcus intermedius, Streptococcus downei* and *Granulicatella adiacens.*

Figure 6.1: Showing the conservancy of selected B-cell epitope KKPNIWYSLNGKIRAVNVPK in 6 cariogens.

7. DISCUSSION AND FUTURE PERSPECTIVE

S.mutans has been identified as the main etiological agent in dental caries (Loesche WJ, 1986). However, the oral cavity is inhabited by a large number of microorganisms which produce acids by utilizing sugars and decrease the pH of oral cavity. This decrease in pH for a prolonged time results in creating an imbalance between the demineralization and re-mineralization of tooth surface (Fehr VD, 1965). As the sugar acids produced by bacteria are responsible for cavitations of tooth enamel and dentine, the best strategy for preventing and controlling dental caries would be to inhibit all the acid producing bacteria.

Most of the studies aimed at preventing dental caries target *S.mutans* but various studies have shown that bacteria other than *S.mutans* can also cause dental caries as they produce acids and thus, promote demineralization of tooth surface. Further, once a biofilm has been established it is not necessary that only *S.mutans* will be responsible for demineralization of tooth surface. In view of this, a strategy targeting most of the acid producing bacteria will be more apt for preventing a disease like dental caries which is caused by a number of bacteria (Kleinberg I., 2002).

This study adopts a strategy that targets multiple microorganisms associated with the causation of dental caries. Till date, vaccines for preventing dental caries target only *S.mutans* and not other microorganisms which either initiate or help in the progression of dental caries. In this study, an attempt has been made to target caries initiating as well as microorganisms associated with progression of caries using a single vaccine. Advances in sequencing techniques and the tools available for analyzing the sequence data have enabled prediction of probable vaccine candidates from the large number of proteins synthesized by an organism. For targeting multiple microorganisms along with the major etiological agent *S.mutans*, vaccine candidates that are conserved in a number of microorganisms have been predicted.

As a plethora of microorganisms have been associated with dental caries, the first and foremost step for targeting multiple microorganisms associated with the causation of dental caries was to identify the microorganisms that have to be targeted. Using literature search, 70 microorganisms were selected as probable caries associated microorganisms. These may be responsible for initiation (*S.mutans, S.sobrinus*) or progression (*Lactobacilli* and others) of dental caries.

An antigenic peptide conserved in a number of microorganisms growing in a biofilm can be used as a vaccine to targ*et al*l those microorganisms. For a protein to be used as a vaccine it should either be secreted or should be present on the outer side of the cell, i.e, it should be

surface-exposed. So, for assessing the vaccine like character of a protein, the first and foremost step is the elucidation of its sub-cellular localization. Though a number of bioinformatics tools are available for localization prediction but most of them perform well for gram negative bacteria. To validate the sub-cellular localization of proteins instead of one, three localization prediction tools have been used, namely PSORTb, CELLO and Gpos-mPLoc. The proteins predicted as extracellular/secreted or cell wall components by all the three servers have been considered in this study. This decreases the rate of false positives in the result. Another important consideration for a protein to act as a vaccine candidate in humans is that it should not have homolog's in humans, otherwise the vaccine will either be a poor immunogen or if an immune response is triggered, it will result in auto immunity (Wilson *et al*., 2000; Weber *et al*., 2009). Out of the selected *S.mutans* proteins, no protein showed a significant level of homology when blasted with the human genome.

As the role of CD8+ cell mediated immunity is not prominent in oral cavity (Setia *et al*., 2012) and the major immunological response is due to salivary IgA and the IgG antibody produced by the B-cells activated by CD4+ T cell, epitope prediction has been restricted to B-cell and MHC class II branch of antigen presentation only. Also, such linear B-cell epitopes have been predicted from which MHC-class II restricted T-cell epitopes can be derived. This type of epitopes have higher chances of eliciting immune responses because even if the B-cell epitope is not recognized by the immune system, there are chances that T-cell epitope will elicit immune response in the host.

To detect B-cell epitopes that are conserved in a number of cariogenic bacteria, protein similarity was searched in 70 cariogenic microbes (including *S.mutans*) on the assumption that there may be some B-cell epitopes that may have stretches of 9 or more than 9 amino acid residues that may be conserved in all or a number of cariogenic microorganisms. Blast was performed for each of the selected *S.mutans* protein. In addition to *S.mutans* UA159, all other strains (16 strains) of *S.mutans* whose genome has been sequenced were also included in this study. Each *S.mutans* protein and its homolog's in other cariogens were aligned to detect the conserved regions. As expected, small antigenic regions were found to be conserved in diverse proteins in diverse organisms, e.g. Protein with GI number 24380370 has antigenic regions conserved in 7 species. The objective of this study is to predict vaccine candidates which can elicit immune response against a number of microorganisms and there may be proteins which are not conserved or which do not have homologs in other cariogenic microorganisms. This step eliminated the proteins which were specific for *S.mutans* and thus, ruled out the possibility of using such proteins in the subsequent steps of this study. BLASTp search was performed for the entire proteins and not for the B-cell epitopes because the epitopes are just 20 amino acids long and there are very high chances of such small sequences to be conserved in non-homologous proteins also.

B-cell epitopes of length 20 amino acids were predicted so that CD4+ T-cell epitopes can be derived from these B-cell epitopes only. For B-cell epitopes prediction, both BCPRED

((Manzalawy *et al*., 2008) and AAP (Chen *et al*., 2007) methods have been used. BCPRED predicts B-cell epitope based on subsequence kernel and AAP uses the amino acid pair (AAP) scale which is based on the finding that certain amino acid pairs are favoured in B-cell epitopes. Though BCPREDS gives antigenic score to each predicted epitope, antigenicity of all epitopes has also been predicted using VaxiJen. Epitopes predicted to be antigenic based on the scores of both BCPREDS and VaxiJen have been selected.

As T-cell epitopes of length 15 amino acids are known to stimulate both the CD4+ and CD8+ cells and they stimulate CD4+ cells more efficiently compared to 9 amino acids long peptides (Darzynkiewicz *et al*., 2004; Holland *et al*., 2013), 15 amino acids long MHC class II restricted T-cell epitopes have been predicted.

Most of the peptide - MHC class II allele binding prediction servers predict binding to HLA-DR locus only as very limited data is available for other class II HLA alleles. In this study, T-cell epitope prediction has been done using NetMHCIIPan 2.0, which performs better compared to other methods of MHC-class II restricted T-cell epitope prediction (Nielsen *et al*., 2010). NetMHCIIPan 2.0 predicts peptide-HLA binding even for the HLA-DR alleles for which experimental binding data is not available. In this study, 655 HLA-DR alleles have been considered.

The predicted T-cell epitopes have been selected based on the criteria: (a) antigenicity of the epitope should be >0.4 (VaxiJen), (b) must bind to HLA-DRB1*0101 with IC_{50} value < 100nM , (c) bind to maximum number of HLA-DR alleles with IC_{50} < 100nM. This criterion helps in the selection of T-cell epitopes which will cover the maximum HLA-DR allele diversity while binding to the most frequently occurring alleles (Panigada *et al*, 2002).

Of all the predicted T-cell epitopes, only the T-cell epitopes derived from the B-cell epitope KKPNIWYSLNGKIRAVNVPK, which is conserved in 5 cariogens, namely *Streptococcus mutans* (all 17 strains), *Streptococcus sobrinus, Streptococcus constellatus, Streptococcus intermedius* and *Streptococcus downei* showed high affinity binding to the most frequently occurring HLA-alleles and bind to the maximum number of HLA-DR alleles. None of the epitopes derived from KKPNIWYSLNGKIRAVNVPK was predicted to bind to HLA-DRB1*0301.

Out of KKPNIWYSLNGKIRAVNVPK derived T-cell epitopes, the epitope KPNIWYSLNGKIRAV binds with lowest IC_{50} values with the most frequently occurring alleles and is conserved in 5 cariogens, namely *Streptococcus sobrinus, Streptococcus constellatus, Streptococcus intermedius, Streptococcus downei*, and the 17 *S.mutans* strains. So, KPNIWYSLNGKIRAV is the best T-cell epitope and KKPNIWYSLNGKIRAVNVPK is the best B-cell epitope of all the epitopes considered in this study.

The finally selected vaccine candidate KKPNIWYSLNGKIRAVNVPK for dental caries belongs to Cell surface antigen SpaP (GI : 24379087) and is expected to act as an vaccine

for *Streptococcus mutans, Streptococcus sobrinus, Streptococcus constellatus, Streptococcus intermedius* and *Streptococcus downei*, thereby sensitizing B-cells and CD4+ T-cells against 5 cariogens. The region PNIWYSLNGKIRAVNVP of the selected vaccine candidate is also conserved in *Granulicatella adiacens*. As this region is antigenic based on VaxiJen score and long enough to sensitize both B-cells and CD4+ T-cells and contains the T-cell epitopes PNIWYSLNGKIRAVN, NIWYSLNGKIRAVNV and IWYSLNGKIRAVNVP, the selected vaccine candidate can potentially be used as a vaccine against *Granulicatella adiacens* also.

Ferretti *et al.* observed that antisera from rabbits immunized with *S.mutans* cross reacted with human heart tissue **(**Ferretti *et al*., 1980**)**. In order to confirm that the selected vaccine candidate will not induce auto-immunity, Blast search against human genome was performed using the selected vaccine candidate as query. No significant sequence similarity was observed for the selected vaccine candidate in the human genome as shown in Figures 5.11.1-5.11.3.

Future Perspective:

Until the emergence of field of Bioinformatics, wet lab experiments were the sole source for identifying targets that can be used as vaccines. Bioinformatics gave rise to the field called Immunoinformatics, which allows for the selection of probable vaccines *in-silico*, thereby saving time and money required to perform wet lab experiments. Immunoinformatics considerably decreases the number of putative targets to be explored for their effectiveness as vaccines in animal models. This study has applied the various tools available for identifying vaccine targets to predict probable vaccine candidates against dental caries.

As dental caries is a poly-microbial disease, a deeper understanding of the mechanisms involved in adhesion and signaling and the interactions between the microbes will aid in the determination of the role of known virulence determinants and the factors which are of utmost importance. For combating polymicrobial diseases, factors like host-microbe pathogenic mechanisms, microbe-microbe interactions, host-immunity mediated antimicrobial defenses and environmental factors need to be considered. So, the future studies should be focused on the study of biofilms in vivo in order to gain deeper insights into complex dynamics within the microbial populations and their interactions with the host.

The composition of microbial consortia in oral cavity differs from person to person and different microbes have been shown to initiate and develop dental caries. In view of this observation, a vaccine targeting a number of cariogenic microbes present in the oral biofilms can cover a much larger percentage of human population. This study predicted a vaccine candidate that can potentially be used as a vaccine against 6 cariogens. For targeting more caries associated microbes, these microbes can be divided into several groups based on

the level of homology between their genomes and then studies similar to the present one can be conducted to identify vaccine targets for each group. Vaccine targets for each group may then be formulated into a single vaccine by protein engineering. The engineered protein will present multiple epitopes on its surface, thereby eliciting immune response to multiple microbes. This approach can further be extrapolated to prevent multiple diseases also, e.g., a multi-epitope vaccine for dental caries may also include epitopes specific for the microbes responsible for other oral diseases like periodontal diseases and halitosis. Further, other diseases like urinary tract infections, endocarditis, and infections in cystic fibrosis, associated with the formation of biofilms can also be targeted using the same strategy.

The immunogenic potential of vaccine candidates predicted *in-silico* needs to be validated in animal models and humans. These studies are helpful in deciding the right time for immunization, the route of administration, the adjuvant to be used and other factors affecting the efficacy of a vaccine.

8. REFERENCES

1. Alexandra K Marr, William J Gooderham, Robert EW Hancock, (2006). Antibacterial peptides for therapeutic use: obstacles and realistic outlook. Current Opinion in Pharmacology; 6(5):468–72.
2. Anders Krogh, Björn Larsson, Gunnar von Heijne, Erik L.L Sonnhammer, (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. Journal of Molecular Biology; 305(3):567-80.
3. Barh D, Misra AN, Kumar A, Vasco A, (2010). A novel strategy of epitope design in *Neisseria gonorrhoeae*. Bioinformation; 5(2):77-85.
4. Bell Raj Eapen, (2009). Delayed-Type hypersensitivity to latex: Computational prediction of MHC class II epitopes on latex allergens. Nature Precedings; hdl.handle.net/10101/npre.2009.2931.1
5. Binod Kumar Patro, B Ravi Kumar, Anil Goswami, Vijay Prakash Mathur, Baridalyne Nongkynrih, (2008). Prevalence of dental caries among adults and elderly in an urban resettlement colony of New Delhi. Indian Journal of Dental Research; 19(2): 95-98.
6. Brailsford SR, Tregaskis RB, Leftwich HS, Beighton D, (1999).The predominant Actinomyces spp. isolated from infected dentin of active root caries lesions. Journal of Dental Research; 78(9):1525-34.
7. Brittany Bower, Anton Dormer, (2014). A novel strategy of *In-silico* peptide vaccine design for *Streptococcus mutans*. Online Journal of Bioinformatics; 15 (1): 1-34
8. C. A. Weber, P. J. Mehta, M. Ardito, L. Moise, B. Martin, and A. S. De Groot, (2009). T cell epitope: friend or foe? Immunogenicity of biologics in context. Advanced Drug Delivery Reviews; 61(11): 965–976.
9. Chen J, Liu H, Yang J, Chou K, (2007). Prediction of linear B-cell epitopes using amino acid pair antigenicity scale. Amino Acids; 33: 423-428.
10. Christopher G Daly, (2009). Prescribing good oral hygiene for adult. Australian Prescriber; 32:72–5.
11. Christopher J. Holland, David K. Cole, Andrew Godkin, (2013). Re-directing CD4+ T cell responses with the flanking residues of MHC class II-bound peptides: the core is not enough. Frontiers in Immunology: 4:172.
12. C. Wilson, H. Tiwana, A. Ebringer, (2000). Molecular mimicry between HLA-DR alleles associated with rheumatoid arthritis and Proteus mirabilis as the aetiological basis for autoimmunity. Microbes and Infection; 2(12): 1489–1496.
13. da Conceicao Tavares Gomes MD, das Neves Ferreira da Silva PM, Rua Vilanova MJ, (2009). Vaccine against dental caries based on virulence-associated immunomodulatory extracellular proteins produced by the cariogenic bacteria *Streptococcus sobrinus* and *Streptococcus mutans*. US7541041.

14. D.J. Smith, (2002).Dental caries vaccines: Prospects and concerns. Critical Reviews in Oral Biology & Medicine; 13(4): 325-339.
15. DJ Smith, MA Taubman, (1990). Effect of local deposition of antigen on salivary immune responses and reaccumulation of *mutans streptococci*. Journal of Clinical Immunology; 10: 273–281.
16. DJ Smith, MA Taubman, (1987). Oral immunization of humans with *Streptococcus sobrinus* glucosyltransferase. Infection and Immunity; 55: 2562–2569. 27.
17. DM Walker (2004). Oral mucosal immunology: An overview. Annals of the Academy of Medicine, Singapore; 33(4):27-30.
18. D Preza, I Olsen, JA Aas,T Willumsen, B Grinde ,BJ Paster, (2008). Bacterial profiles of root caries in elderly patients. Journal of Clinical Microbiology; 46(6):2015-21
19. EC Reynolds, (1998). Anticariogenic complexes of amorphous calcium phosphate stabilized by casein phosphopeptides: a review. Special Care in Dentistry; 18(1):8–16.
20. EL-Manzalawy Y, Dobbs D, Honavar V, (2008). Predicting linear B-cell epitopes using string kernels. Journal of Molecular Recognition; 21: 243-255.
21. Erin L. Gross, Clifford J. Beall, Stacey R. Kutsch, Noah D. Firestone, Eugene J. Leys and Ann L. Griffen (2012). Beyond *Streptococcus mutans*: Dental Caries Onset Linked to Multiple Species by 16S rRNA Community Analysis. PLoS ONE; 7(10): e47722.
22. Fabian Sievers, Andreas Wilm, David Dineen, Toby J Gibson, Kevin Karplus, Weizhong Li, Rodrigo Lopez, Hamish McWilliam, Michael Remmert, Johannes Söding, Julie D Thompson, Desmond G Higgins, (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology: 7: 539.
23. Franklin García-Godoy and M. John Hicks, (2008). Maintaining the integrity of the enamel surface: the role of dental biofilm, saliva and preventive agents in enamel demineralization and remineralization. The Journal of the American Dental Association; 139(Suppl): 25S–34S.
24. Fu Chen, Dong Wang, (2010).Novel technologies for the prevention and treatment of dental caries: a patent survey. Expert Opinion on Therapeutic Patents. 20(5): 681–694.
25. H.H.K. Xu, J.L. Moreau, L. Sun, L.C. Chow, (2010). Novel CaF2 Nanocomposite with High Strength and Fluoride Ion Release. Journal of Dental Research; 89(7): 739–745.
26. Hong-Bin Shen and Kuo-Chen Chou, (2009). Gpos-mPLoc: A Top-Down Approach to Improve the Quality of Predicting Subcellular Localization of Gram-Positive Bacterial Proteins. Protein and Peptide Letters; 16: 1478-1484.
27. I. Kleinberg, (2002). A mixed-bacteria ecological approach to understanding the role of the oral bacteria in dental caries causation: an alternative to *Streptococcus mutans* and the specific-plaque hypothesis. Critical Reviews in Oral Biology & Medicine; 13(2):108-25.
28. Irini A Doytchinova, Darren R Flower, (2007). VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics; 8:4.

29. Itay Mayrose, Tomer Shlomi, Nimrod D. Rubinstein, Jonathan M. Gershoni, Eytan Ruppin, Roded Sharan, Tal Pupko, (2006). Epitope mapping using combinatorial phage-display libraries: a graph-based algorithm. Nucleic Acid Research; 35(1):69-78.
30. JD Featherstone, (2008). Dental caries: a dynamic disease process. Australian Dental Joural; 53(3):286–91.
31. JD Featherstone, (2004). The continuum of dental caries--evidence for a dynamic disease process. Journal of Dental Research; 83: 39-42.
32. JD Featherstone, (2000).The science and practice of caries prevention. Journal of the American Dental Association; 131(7):887–899.
33. JD Featherstone, (1999). Prevention and reversal of dental caries: role of low level fluoride. Community Dentistry and Oral Epidemiology; 27(1):31–40.
34. JD Hillman, TA Brooks, SM Michalek, (2000). Construction and characterization of an effector strain of *Streptococcus mutans* for replacement therapy of dental caries. Infection and Immunity; 68(2):543–9.
35. J.H. Guo, R. Jia, M.W. Fan, Z. Bian, Z. Chen, B. Peng, (2004).Construction and Immunogenic Characterization of a Fusion Anti-caries DNA Vaccine against PAc and Glucosyltransferase I of *Streptococcus mutans*. Journal of Dental Research; 83(3): 266-270.
36. J J Ferretti, C Shea, and M W Humphrey, (1980). Cross-reactivity of *Streptococcus mutans* antigens and human heart tissue. Infection and Immunity; 30(1): 69–73.
37. Joyson Moses, B N Rangeeth, Deepa Gurunathan , (2011). Prevalence Of Dental Caries, Socio-Economic Status And Treatment Needs Among 5 To 15 Year Old School Going Children Of Chidambaram. Journal of Clinical and Diagnostic Research; 5(1):146-151.
38. JP Pessan, NS Al-Ibrahim, MA Buzalaf, KJ Toumba, (2008). Slow-release fluoride devices: a literature review. Journal of Applied Oral Science; 16(4):238–46.
39. KA Brogden, (2005). Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? Nature Reviews. Microbiology. ; 3(3):238–50.
40. KJ Cross, NL Huq, EC Reynolds, (2007). Casein phosphopeptides in oral health--chemistry and clinical applications. Current Pharmaceutical Design; 13(8):793–800.
41. KM Shivkumar, SK Vidya, GN Chandu, (2009). Dental caries vaccine. Indian Journal of Dental Research; 20:99-106.
42. Krithika AC, Kandaswamy D, Gopikrishna V,(2004). Caries Vaccine-I Today's Myth. The Journal of Indian Association of Public Health Dentistry; 4:21-5.
43. Kruger C, Pearson SK, Kodama Y, Vacca Smith A, Bowen WH, Hammarstrom L, (2004). The effects of egg-derived antibodies to glucosyltransferase on dental caries in rats. Caries Research; 38:9-14.
44. Li F, Michalek SM, Dasanayake AP, Li Y, Kirk K, *et al*. (2003) Intranasal immunization of humans with *Streptococcus mutans* antigens: Oral Microbiology AND Immunology; 18: 271–277.

45. Lifu Song, Wei Wang, Georg Conrads, Anke Rheinberg, Helena Sztajer, Michael Reck, Irene Wagner-Döbler, An-Ping Zeng, (2013). Genetic variability of mutans streptococci revealed by wide whole-genome sequencing. BMC Genomics; **14**:430.
46. L Torlakovic, V Klepac-Ceraj,B Ogaard, SL Cotton, BJ Paster, I Olsen , (2012). Microbial community succession on developing lesions on human enamel. Journal of Oral Microbiology; 4.
47. Maddalena Panigada, Tiziana Sturniolo, Giorgio Besozzi, Maria Giovanna Boccieri, Francesco Sinigaglia, Giuliana Gialdroni Grassi, Fabio Grassi, (2002). Identification of a Promiscuous T-Cell Epitope in *Mycobacterium tuberculosis* Mce Proteins. Infection And Immunity; 70(1): 79-85.
48. Marsh P, Martin MV, (1999). Oral Microbiology. 4th edn. Wright; Oxford.
49. Mitzi R. Becker, Bruce J. Paster, Eugene J. Leys, Melvin L. Moeschberger, Sarah G. Kenyon, Jamie L. Galvin, Susan K. Boches, Floyd E. Dewhirst, and Ann L. Griffen, (2002). Molecular Analysis of Bacterial Species Associated with Childhood Caries. Journal Of Clinical Microbiology; 40(3):1001–1009.
50. Morten Nielsen, Sune Justesen, Ole Lund, Claus Lundegaard, Søren Buus, (2010). NetMHCIIpan-2.0 - Improved pan-specific HLA-DR predictions using a novel concurrent alignment and weight optimization training procedure. Immunome Research; **6**:9.
51. Nancy Y. Yu, James R. Wagner, Matthew R. Laird, Gabor Melli, Sébastien Rey, Raymond Lo, Phuong Dao, S. Cenk Sahinalp, Martin Ester, Leonard J. Foster, Fiona S. L. Brinkman, (2010). PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics; 26(13):1608-1615.
52. N Nabi, A Gaffar, (1990). Antibacterial antiplaque oral composition. US4894220.
53. Niu Y, Sun J, Fan M, Xu QA, Guo J, (2009). Construction of a New Fusion Anti-caries DNA Vaccine. Journal of Dental Research; 88: 455-460.
54. NK Childers, SS Zhang, SM Michalek (1994). Oral immunization of humans with dehydrated liposomes containing *Streptococcus mutans* glucosyltransferase induces salivary immunoglobulin A2 antibody responses. Oral Microbiology AND Immunology; 9: 146–153.
55. N Takahashi , B Nyvad , (2008). Caries ecology revisited: microbial dynamics and the caries process. Caries Research; 42:409–418.
56. Okada M, Soda Y, Hayashi F, Doi T, Suzuki J, Miura K and Kozai K, (2005). Longitudinal study of dental caries incidence associated with *Streptococcus mutans* and *Streptococcus sobrinus* in pre-school children. Journal of Medical Microbiology; 54(7):661-5.
57. Pedro Belda-Ferre, Luis David Alcaraz1, Raúl Cabrera-Rubio, Héctor Romero, Aurea Simón-Soro, Miguel Pignatelli, Alex Mira, (2012). The oral metagenome in health and disease. The ISME Journal; 6, 46–56.
58. PE Petersen, D Bourgeois, H Ogawa, S Estupinan-Day and C Ndiaye, (2005). The global burden of oral diseases and risks to oral health. Bulletin of the World Health Organization; 83:661–669.

59. Ramandeep Singh Gambhir, Simarpreet Singh, Gurminder Singh, Rina Singh, Tarun Nanda and Heena Kakar, (2012). Vaccine against Dental Caries- An Urgent Need. Vaccines & Vaccination; 3:2.
60. RG Rozier, (2001). Effectiveness of Methods used by Dental Professionals for the Primary Prevention of Dental Caries: A Review of the Evidence. Journal of Dental Education; 65(10):1063-72.
61. Robert H Selwitz, Amid I Ismail and Nigel B Pitts, (2007). Dental Caries. The Lancet; 369(9555): 51-59.
62. RO Mattos-Graner, DJ Smith,(2004). The vaccination approach to control infections leading to dental caries. Braz J Oral Sci; 3:595-608.
63. Russel MW, Childers NK, Michalek SM, Smith DJ, Taubman MA, (2004) A caries vaccine? The state of science of immunization against dental caries. Caries Research; 38:230-5.
64. Saniya Setia, Ramandeep Singh Gambhir and Vinod Kapoor, (2012). Immunology in prevention of dental caries. Universal Research Journal of Dentistry; 2(2): 58-63.
65. Sakeenabi Basha, Hiremath Shivalinga Swamy (2012). Dental caries experience, tooth surface distribution and associated factors in 6- and 13- year- old school children from Davangere, India. Journal of Clinical and Experimental Dentistry; 4(4):e210-6.
66. T Hoshino, Y Kondo, K Saito, Y Terao, N Okahashi, S Kawabata, T Fujiwara, (2011). Novel epitopic region of glucosyltransferase B from *Streptococcus mutans*. Clinical and Vaccine Immunology; 18(9): 1552-61.
67. T Lehner, (1992). Immunology of dental caries. Immunology of oral diseases. 3 rd ed. Blackwell scientific publications.
68. Weatherell JA, (1975). Composition of dental enamel. Br. Med. Bull; 31:115–119.
69. WH Bowen, (2002). Do we need to be concerned about dental caries in the coming millennium? Critical Reviews in Oral Biology And Medicine; 13:126-31.
70. WH Bowen, (1996). Vaccine against dental caries: A personal view. Journal of Dental Research; 75:1530-9.
71. WJ Loesche, (1986). Role of *Streptococcus mutans* in human dental decay. Microbiological Reviews; 50:353–380.
72. Yu CS, Chen YC, Lu CH, Hwang JK, (2006). Prediction of protein subcellular localization. Proteins: Structure, Function and Bioinformatics; 64:643-651.
73. Zbigniew Darzynkiewicz, Mario Roederer, and Hans Tanke, (2004). Cytometry, 4th Edition: New Developments ;Volume 75, Pages 1-876
74. Z Luo, DJ Smith, Taubman MA, WF King, (1988). Cross sectional analysis of serum antibody to oral streptococcal antigens in children. Journal of Dental Research; 67:554-60.

APPENDIX I

LIST OF MICROORGANISMS ASSOCIATED WITH THE CAUSATION OF DENTAL CARIES

1. *Actinomyces viscosus (taxid:1656)*
2. *Acinetobacter baumannii (taxid:470)*
3. *Actinobaculum sp. oral taxon 183 (taxid:712888)*
4. *Actinomyces georgiae (taxid:52768)*
5. *Actinomyces gerencseriae (taxid:52769)*
6. *Actinomyces israelii (taxid:1659)*
7. *Actinomyces johnsonii (taxid:544581)*
8. *Actinomyces naeslundii (taxid:1655)*
9. *Actinomyces odontolyticus (taxid:1660)*
10. *Atopobium parvulum (taxid:1382)*
11. *Bacteroidetes bacterium oral taxon 272 (taxid:651591)*
12. *Bacteroidetes bacterium oral taxon 274 (taxid:712899)*
13. *Bifidobacterium dentium (taxid:1689)*
14. *Bifidobacterium longum (taxid:216816)*
15. *Campylobacter gracilis (taxid:824)*
16. *Campylobacter showae (taxid:204)*
17. *Capnocytophaga ochracea (taxid:1018)*
18. *Corynebacterium diphtheriae (taxid:1717)*
19. *Dialister invisus (taxid:218538)*
20. *Enterococcus faecalis (taxid:1351)*
21. *Eubacterium alactolyticum (taxid:113287)*
22. *Filifactor alocis (taxid:143361)*
23. *Gemella sanguinis (taxid:84135)*
24. *Granulicatella adiacens (taxid:46124)*
25. *Granulicatella elegans (taxid:137732)*
26. *Haemophilus influenzae (taxid:727)*
27. *Haemophilus parainfluenzae (taxid:729)*
28. *Helicobacter pylori (taxid:210)*
29. *Klebsiella pneumoniae (taxid:573)*
30. *Lactobacillus acidophilus (taxid:1579)*
31. *Lactobacillus brevis (taxid:1580)*
32. *Lactobacillus casei (taxid:1582)*
33. *Lactobacillus fermentum (taxid:1613)*
34. *Lactobacillus gasseri (taxid:1596)*

35. *Lactobacillus jensenii (taxid:109790)*
36. *Lactobacillus johnsonii (taxid:33959)*
37. *Lactobacillus paracasei subsp. paracasei (taxid:47714)*
38. *Lactobacillus plantarum (taxid:1590)*
39. *Lactobacillus rhamnosus (taxid:47715)*
40. *Lactobacillus salivarius (taxid:1624)*
41. *Lactobacillus vaginalis (taxid:1633)*
42. *Mitsuokella multacida (taxid:52226)*
43. *Mobiluncus mulieris (taxid:2052)*
44. *Neisseria mucosa (taxid:488)*
45. *Olsenella profusa (taxid:138595)*
46. *Parascardovia denticolens (taxid:78258)*
47. *Parvimonas micra (taxid:33033)*
48. *Prevotella multisaccharivorax (taxid:310514)*
49. *Prevotella intermedia / Prevotella nigrescens-like organism (PINLO) (taxid:60133)*
50. *Propionibacterium acnes (taxid:1747)*
51. *Propionibacterium avidum (taxid:33010)*
52. *Rothia dentocariosa (taxid:2047)*
53. *Scardovia inopinata (taxid:78259)*
54. *Scardovia wiggsiae (taxid:230143)*
55. *Selenomonas sputigena (taxid:69823)*
56. *Slackia exigua (taxid:84109)*
57. *Staphylococcus aureus (taxid:1280)*
58. *Streptococcus agalactiae (taxid:1311)*
59. *Streptococcus constellatus (taxid:76860)*
60. *Streptococcus downei (taxid:1317)*
61. *Streptococcus intermedius (taxid:1338)*
62. *Streptococcus parasanguinis (taxid:1318)*
63. *Streptococcus pyogenes (taxid:1314)*
64. *Streptococcus salivarius (taxid:1304)*
65. *Streptococcus sobrinus (taxid:1310)*
66. *Streptococcus vestibularis (taxid:1343)*
67. *Veillonella atypica (taxid:39777)*
68. *Veillonella dispar (taxid:39778)*
69. *Veillonella parvula (taxid:29466)*
70. *Streptococcus mutans (taxid:1309)*
71. *Streptococcus mutans GS-5 (taxid:1198676)*
72. *Streptococcus mutans LJ23 (taxid:1155071)*
73. *Streptococcus mutans NN2025 (taxid:511691)*
74. *Streptococcus mutans SF1 (taxid:857121)*

75. Streptococcus mutans U2A (taxid:857116)
76. Streptococcus mutans 3SN1 (taxid:857149)
77. Streptococcus mutans SM4 (taxid:857109)
78. Streptococcus mutans U138 (taxid:857135)
79. Streptococcus mutans NFSM1 (taxid:857130)
80. Streptococcus mutans S1B (taxid:857105)
81. Streptococcus mutans N66 (taxid:857124)
82. Streptococcus mutans NV1996 (taxid:857123)
83. Streptococcus mutans OMZ175 (taxid:857099)
84. Streptococcus mutans ST1 (taxid:857118)
85. Streptococcus mutans 24 (taxid:857107)
86. Streptococcus mutans NLML1 (taxid:857114)

APPENDIX II

Proteins predicted to be either extracellular or localized in cell wall by all the three servers.

SNo.	GI number	Protein name	Localization
1	24378589	hypothetical protein SMU_63c	Cellwall
2	24378602	exo-beta-D-fructosidase; fructanase FruA	Cellwall
3	24378603	exo-beta-D-fructosidase FruB	Cellwall
4	24378708	transfer protein	Cellwall
5	24379087	cell surface antigen SpaP	Cellwall
6	24379801	glucan-binding protein GbpC	Cellwall
7	24379424	hypothetical protein SMU_984	Cellwall
8	24379427	cell wall-associated protein WapA	Cellwall
9	24379528	cell wall protein, WapE	Cellwall
10	24379599	thioredoxin family protein	Cellwall
11	24380381	dextranase	Cellwall
12	24380475	hypothetical protein SMU_2147c	Cellwall
13	24378802	bacteriocin peptide	Extracellular
14	24379091	hypothetical protein SMU_616	Extracellular
15	24379175	autolysin; amidase	Extracellular
16	24379231	glucan-binding protein D	Extracellular
17	24379291	hypothetical protein SMU_836	Extracellular
18	24379358	glucosyltransferase-S	Extracellular

19	24379444	glucosyltransferase-I	Extracellular
20	24379445	glucosyltransferase-SI	Extracellular
21	24380122	hypothetical protein SMU_1752c	Extracellular
22	24380237	hypothetical protein SMU_1882c	Extracellular
23	24380370	beta-D-fructosyltransferase	Extracellular
24	24380386	hypothetical protein SMU_2048	Extracellular
25	24380444	glucan-binding protein GbpA	Extracellular
26	24380474	hypothetical protein SMU_2146c	Extracellular