UNIVERSITY OF CATANIA

INTERNATIONAL PhD IN BASIC AND APPLIED BIOCHEMICAL SCIENCES

XXXIII CYCLE

Dr. Giovanni Sposito

DESIGN, SYNTHESIS AND EXPRESSION OF RECOMBINANT PROTEINS FOR THE DEVELOPMENT OF HEPATITIS E-MALARIA VACCINE

DOCTORAL THESIS

Coordinator:	Tutor:	Co-supervisor
Prof.ssa Stefania Stefani	*Prof.ssa Agata Campisi*	*Dott.ssa Agata Fazzio*

Academic Years 2018-2021

CHAPTER I

Malaria Etiopathogenesis

Nowadays malaria represents the most important disease transmitted by a carrier in tropical and subtropical areas. World Health Organisation (WHO) has recently reported that there are circa 216 million cases and 445 thousand deaths in 91 countries with high malaria endemic level

(http://www.salute.gov.it/portale/temi/p2_6.jsplingua=italiano&id=1983&area=Malattie%20infettive&menu=viaggiatori).

Malaria is a parasitic disease caused by *Plasmodium* species. In particular, *Plasmodium falciparum,* responsible of a severe form of malaria, presents a life cycle which is characterised by several phases, where each is capable of inducing direct immune responses against specific-stage antigens (Fig. 1).

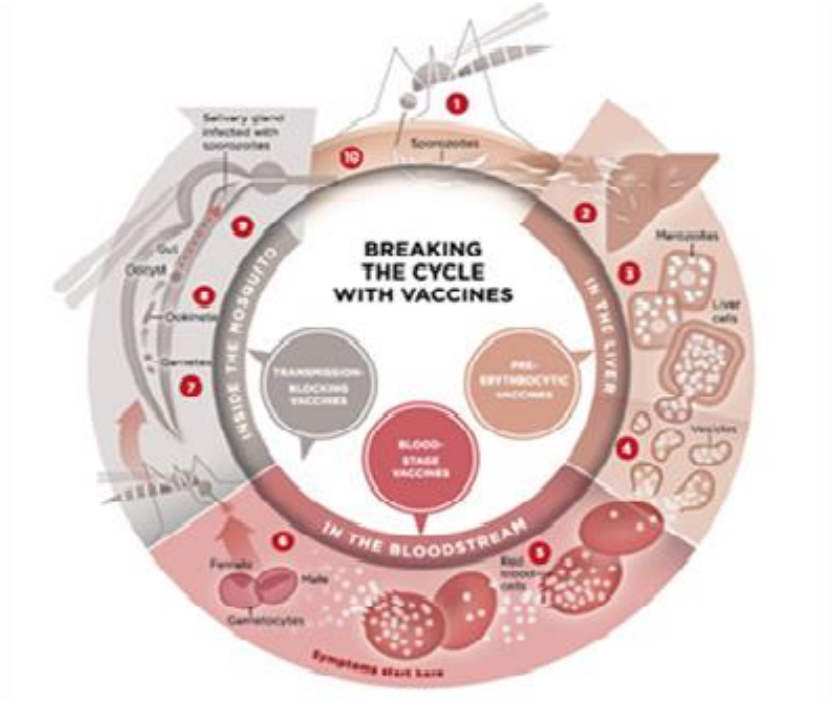


Fig. 1: *Plasmodium falciparum* reproductive cycle

(https://www.malariavaccine.org/malaria-and-vaccines/vaccine-development/life-cycle-malaria-parasite)

Several female *Anopheles* mosquitoes transmit malaria parasites to humans; while septic mosquitoes inoculate a parasite “sporozoite” into the blood stream of mammalians.

Before invading hepatocytes, this sporozoite, which situates in an extracellular environment, is exposed to attack of antibodies directed against “circumsporozoite protein” (CSP). CSP is indeed expressed during a sporozoite phase and also at the beginning of parasite hepatic phase (Coppi A et al., 2011).

Once this parasite locates in the liver, it replicates itself and develops a “schizont” shape in a ratio of 20.000 per infected cell. After approximately a week of liver infection, thousands

of “merozoites” are released into the blood stream. This form of parasite expresses in its surface proteins, such as *Merozoite Surface Protein* MSP-1, MSP-2, MSP-3. They are essential components of haematic-asexual phase of the merozoite, which is responsible of erythrocytes invasion.

Once MSP-1, -2 and -3 are in the blood cells, they become a target of humoral immune response (which is mediated by antibodies) and cellular one (which is correlated to cytokines production). Furthermore, during erythrocytes invasion, the merozoite produces a protein, such as an antigen 175 capable of binding to those erythrocytes (EBA-175), which combines with glycophorin A.

These merozoites undergo various stages of replication, generating “trophozoites”, “schizonts”, and other forms which could infect new cells. Nevertheless, a limited quantity of trophozoites can evolve in “gametocytes”, both male and female.

Those fecundate and they generate a zygote that, subsequently, matures in ookinete, then in oocyst, and finally in sporozoite, which migrates on the level of the vector’s salivary glands. This event would cause a new infective cycle (De Carneri I, 2012).

1.2 Symptoms

This disease symptoms are associated with various infective stages of the parasite. They include severe anaemia, fever, thrombocytopenia, shivers, cephalalgy, emesis, muscular pain, anorexia, sternness, diarrhoea, abdominal discomfort, cough, seizure, breathing difficulty, hypoglycaemia, metabolic acidosis, overproduction of production of lactic acidic, coma associated with increase of intracranial pressure (cerebral malaria), retinopathy,

pregnancy complications, including preterm childbirth and low weight at birth, due to the restrictions of foetal growth/development (FGR) (Mawson AR, 2013).

These symptoms are associated with an infected erythrocytes rupture and with a release of alleged malaria toxins, which activate mononuclear blood peripheral cells and stimulate cytokines' release. It is believed that the balance between pro-inflammatory cytokines and anti-inflammatory chemokines, which are growth factors and effector molecules, is correlated with the degree of the disease's severity (Langhorne J et al., 2008).

1.3 Diagnosis

A clinical diagnosis of *malaria* is still difficult due to no specificity of its symptom's nature; thus, it leads to overlap *malaria* with other common and potentially life-threatening diseases, such as viral or bacterial infections and other febrile disorders.

Therefore, the overlay of *malaria* symptoms with another tropical disease compromises diagnostic specificity; hence, it can threaten the care quality for patients with no malarial fever in endemic areas (Mwangi TW et al., 2005; McMorrow ML et al., 2008).

Laboratory diagnostics expects use of several techniques, such as conventional microscopic *via* coloration Giemsa of peripheric blood smears (Ngasala B et al., 2008); other methods provide the use of concentration techniques, such as *quantitative buffy coat* method (QBC) (Bhandari PL et al., 2008). QBC was designed to improve microscopic detection of parasites and to simplify malaria diagnosis (Clendennen TE et al., 1995). This method includes coloration of parasite's deoxyribonucleic acid (DNA) in microhematocrit test tubes with fluorescent colourants, such as acridine orange, and its subsequent detection through epi-fluorescent microscopy.

Since WHO recognised the urgent need of new, simple, rapid, accurate, economical diagnostic tests to determine malarial parasites presence, to overcome deficiency of optical microscopy, several new malaria diagnostic techniques (OMS), including rapid diagnostic tests (RDT) were developed. They are fast and easy to execute and they do not require electricity or specific equipment (Bell D et al., 2006).

Unlike conventional microscopic diagnosis via coloration of peripheric blood smears and QBC technique, all RDT centre upon the same principle and reveal malaria antigen in blood which flows through a membrane containing specific antimalarial antibodies. Most of RDT aim to a specific protein of *Plasmodium falciparum*, for example protein II, which is rich of histidine, (HRP-II) or to lactate dehydrogenase (LDH).

Polymerasic chain reaction (PCR) continues to be wildly used to confirm malarial infection, to practice a therapeutic response and to identify drug resistance (Chotivanich K et al., 2006).

This technique shows a greater sensibility and specificity among microscopic diagnosis of coloured peripheric blood smears and now it seems the best method for a diagnosis of this disease (Morassin B et al., 2002). Albeit PCR appears to have exceeded both its main problems of sensibility and specificity to malarial diagnosis, many complicated methodologies, such as high expenses and necessity of specialized technicians, limit its utility.

1.4 Treatment

Once diagnosed, malaria has to be treated as a medical emergency because *Plasmodium falciparum* infections can progress rapidly and become fatal. Other species of *Plasmodium* rarely cause severe display. A pharmacological choice relies on the infectious *Plasmodium*

species; it depends on clinical condition of patients (for example, anaemia, renal failure, pulmonary edema, shock, acidosis, jaundice and convulsions) and also on geographical area where infection occurs to be contracted. As a matter of fact, some areas host resistant parasites to certain medicaments. Amidst antimalarial medicines, intravenous quinine is the most used therapy in malaria *Plasmodium falciparum* treatment resistant to other antimalarials. Quinine provides a rapid onset of action with a brief half-life elimination and it has an addictive effect whether combined with antibiotics such as clindamycin, tetracycline or doxycycline.

In the US, quinidine gluconate, an optical diastereomer of quinine, is the only available antimalarial intravenous medicine and it can be used in lieu of quinine. Nevertheless, it detains many serious adverse effects, including cardiotoxicity, which require an electrocardiographic monitoring. Therefore, most doctors choose to use quinine.

Quinine and quinidine, alkaloids isolated from *Cinchona officinalis* L. cortex, are capable of killing a parasite before entering inside red blood cells or during a phase where it is silent within its host.

Chloroquine phosphate, a synthetic form of quinine, was introduced after the Second World War and it is still drug of choice for *Plasmodium falciparum* infections that are not resistant to chloroquine. Thus, chloroquine is also drug of choice for *Plasmodium vivax* and *Plasmodium ovale*, because few chloroquine-resistant strains were reported in *Plasmodium malariae*. Unfortunately, *Plasmodium falciparum* is highly resistant to chloroquine in most areas of the world, particularly in Africa, therefore it was replaced by sulphonamide-pyrimethamine in many African regions (https://www.uspharmacist.com/article/malaria-treatment-and-prevention).

1.5 Vaccination

RTS, S vaccine was created in 1987 as part of a collaboration established in 1984 between GlaxoSmithKline (GSK) and Walter Reed Army Institute of Research (WRAIR). During this time, both research groups have been trying to develop a vaccine based on the proof that sporozoites, which were attenuated by radiations, protect from malarial infection.

This vaccine is composed by a repeated region of circumsporozoite protein (CSP), which is added to superficial antigen of hepatitis B virus (HBsAg) and to adjuvant AS01. It causes induction of a high level of human immunity (Coelho CH et al., 2017) and it belongs to a group of vaccines with sporozoite subunits (pre-erythrocytes).

An experimentation known as RTS, S/AS01 started in 2009 was subsequently completed (Chan M, 2016). Several studies behind III phase led to the permission of its commercialisation and involved 15.460 children in 7 sub-Saharan countries.

Each child received three doses of immunisation at intervals of a month and they were divided in two different age groups: 6-12-weeks old neonates and 5-17-months old infants. They received a fourth dose after 18 to 20 months. During over 48 follow-up months, RTS, S/AS01 efficacy was estimated to be 36,3% amongst the oldest group after four doses of immunisation and to be 28,3% after three doses.

An observation period for neonates was shorter: during over 38 follow-up months, protection against clinic malaria was estimated to be 25,9% after four doses and 18,3% after three doses (Olotu A et al., 2016). Therefore, efficacy is moderated amongst the group of older children, but it is not sufficient for neonates, showing that further studies are necessary.

On one hand RTS, S/AS01 is generally well tolerated, with typical side effects similar to other consolidated infancy vaccines; on the other an increased risk of febrile seizures has

been identified among the older children group, although it presents no serious consequences afterwards.

It was reported that, amongst only the group of older children, 16 cases of meningitis caused death and 8 cases of cerebral malaria. Nevertheless, it is not evident that there is a clear connection between meningitis or cerebral malaria and administration of RTS, S/AS01 and it is still under evaluation through a pilot study conducted in Africa.

That level of protection depends on the anti-corporal title against sporozoite surface and decreases over time (Olotu A et al., 2013; Olotu A et al., 2016).

Main limitations to RTS, S/AS01 vaccine include: moderate protection, amount of doses to maintain efficacy, administration system in African countries, a vaccine cost, probable interference between acquired maternal antibodies and *Plasmodium*, side effects and a safety concern.

CHAPTER II

Liver stage (Hepatocytes) and Blood stage (Erythrocyte) proteins

2.1 CSP

CSP is present in every *Plasmodium* species and it is situated on the sporozoite surface. CPS coding genes are instead located in chromosome 3 of *Plasmodium falciparum* genome. This protein is synthesized as precursor going from 50 to 70 kDa; it is subsequently processed in a protein with a mature surface from 40 to 60 kDa (Coppi A et al., 2005). CSP is an important multifunctional molecule for a parasite, which performs several roles, according to its vital development stage.

It has been reported that CPS participates in invading salivary glands of mosquitoes (Sidjanski SP et al., 1997), in the relation between sporozoites and liver cells (Cerami C et al., 1992; Rathore D et al., 2003), and in the inactivation of a host cell's protein synthesis machinery. CPS generally presents in various *Plasmodium* species common structural features, with a variable central region composed by repeated amino acid sequences in tandem and two highly conserved portions (regions I and II).

Region I, situated in a distal amino terminal part, contains KLKQP motif, which is preserved within several *Plasmodium* species: this region proves to be involved with the invasion of mosquito's salivary glands (McCutchan TF et al., 1996). Fragments of the protein and peptides containing this motif can inhibit an invasion of salivary glands sporozoites.

Furthermore, on the level of its carboxy terminal region CPS contains thrombospondin (TSR), a known cell adhesion protein which has high affinity for hepatocytes (Sidjanski SP, 1997; Myung JM et al., 2004; Coppi A et al., 2005) (Fig. 2).

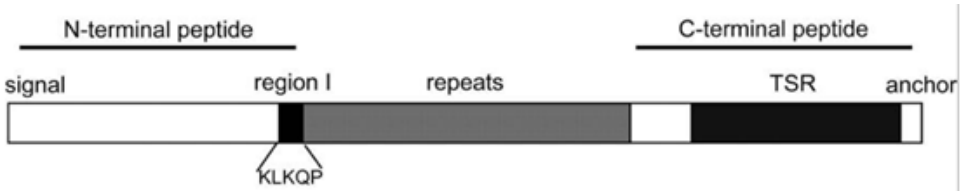


Fig. 2: *Plasmodium* CPS detains several conserved features: a repeated central region (gray box) and two conserved regions (black stripes i.e., region I and TSR) (Coppi A et al., 2005)

2.2 PfF2 protein

Every pathological and clinical manifestation of the disease is caused by an asexual erythrocyte stage. Therefore, the invasion is a logical objective to develop useful interventions for the control of malaria. During this phenomenon, *Plasmodium falciparum* protein EBA-175 binds with glycophorin receptor A (GpA) which is present on human erythrocytes (Sim BK et al., 1990; Adams JH et al., 1992; Klotz FW et al., 1992; Orlandi PA et al., 1992). Some antibodies, going against EBA-175, inhibit its bound with GpA and block merozoite invasion *in vitro* (Sim BK et al., 1990). The erythrocyte binding domain of EBA-175 is region II (RII). RII is a 616 amino acid fragment that contains two cysteine-rich regions F1 and F2.

Both F1 and F2 are homologous to *Duffy* binding protein of *Plasmodium vivax* and are, therefore, called Duffy binding-like domains (DBL). Presence of one or two DBL domains,

a cysteine-rich C-terminal region (VI region) and a type I transmembrane domain define EBA-175 as a member of EBL superfamily (Erythrocyte Binding-like) (Adams JH et al., 1992) (Fig. 3). DBL domains, which represent a characteristic of the EBA-175 family, are important for GpA receptor (Mayor A et al., 2005); these domains appear also to be unique to *Plasmodium*, making themselves a great pharmacological target (Aravind L et al., 2003).

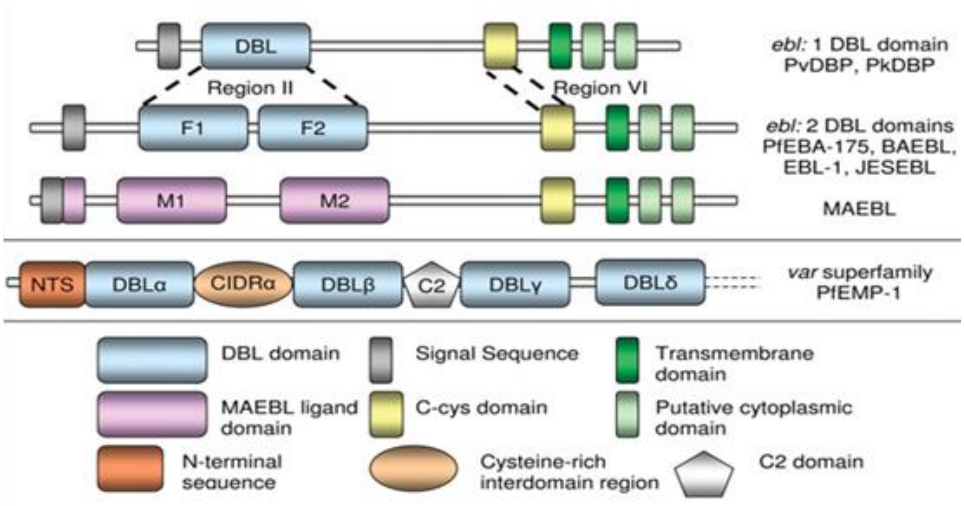


Fig. 3: DBL Containing proteins from *Plasmodium*
(Tolia NH et al., 2005)

It has been previously shown/demonstrated that F2 alone can bind with erythrocytes, but that F1 is insufficient. Furthermore, there are dimeric interactions exclusively between F2 monomers, but F1 depends on F2 for dimerization. Therefore, it has been hypothesized/proposed that F2 by itself may form a weak dimer or may be induced to dimerize by the receptor. Moreover, F2 also contributes with 75% of the glycan contacts observed in the complex with sialyllactose. Thus, it appears to provide most of the interaction sites for its receptor.

2.3 PfMSPFu24 Protein

The approach to build chimeric fusion proteins were used to develop vaccines in haematic phase. Several studies have reported that the chimeric proteins, based on merozoite surface antigens, show significantly increased immune responses compared to each single antigen (Shi Q et al., 2007; Tamborrini M et al., 2011). A chimeric gene known as MSPFu24 corresponds to immunodominant and highly conserved regions of MSP1 and MSP3 protein. In detail, 19 kDa carboxy-terminal region of MSP1 (PfMSP-119) is linked to the highly conserved 24 kDa region of MSP3 (PfMSP-324) (Mazumdar S et al., 2010; Deepak G et al., 2013). PfMSP-1 and PfMSP-3 are two candidate proteins to a vaccine development for its haematic phase.

The surface protein of merozoite 1 (MSP-1) is synthesized as a precursor protein of 195 kDa, which is proteolytically processed to form a multi-sub-unit complex, expressed on a merozoite surface. A 42 kDa glycosylphosphatidylinositol-anchored component, PfMSP-142, is further cleaved upon invasion, leaving only a 19 kDa C-terminal domain (PfMSP-119) attached to the merozoite surface (Holder AA, 2009).

CHAPTER III

Hepatitis E

Hepatitis E (HEV) is a greater international public health issue and it is estimated that globally 2,3 billion people have been infected. HEV is the main cause of acute viral hepatitis in the world, especially in developing countries. A first epidemy of hepatitis E happened between 1955-1956 in New Delhi, India, and it provoked more than 29,000 cases of infected people with symptomatic jaundice (Viswanathan R, 1957).

Since then, many major outbreaks have occurred in Asia, Africa and Mexico (Velázquez O et al., 1990; Khuroo MS et al., 1994). Furthermore, sporadic hepatitis E outbreaks usually occur in developing Asian and African countries, as well as in industrialised countries (Coursaget P et al., 1998; Das K et al., 2000).

HEV is a circular, non-enveloped, single-stranded, positive-sense RNA virus with a diameter between 32 nm and 34 nm (Lu L et al., 2000). HEV genome organisation is substantially different from other viruses and it has its own family, known as hepevyridae, *hepevirus* genus, HEV species (Schlauder GG et Mushahwar IK, 2001). HEV genome organises itself in three overlying *Open Reading Frame* (ORF). These three codifying frames are used to express different proteins.

ORF1 encodes a polyprotein composed by about 1690 amino acids, which undergoes post-translational cleavage into multiple non-structural proteins necessary for virus replication,

including a methyltransferase, a presumed papain-like cysteine protease, an RNA helicase, and an RNA-dependent RNA polymerase. ORF2 does not overlap with ORF1; it is located at the 3 'end of this genome and encodes its main structural protein, 660 amino acid capsid protein. ORF3 begins with the last nucleotide of ORF1; overlaps with ORF2 and encodes a small immunogenic phosphoprotein of 123 amino acids that binds to a cytoskeleton; this suggests its potential role in assembly of viral particles. ORF organisation differs slightly according to each genotype, nonetheless, its function remains the same (Bradley DW, 1995) (Fig. 4).

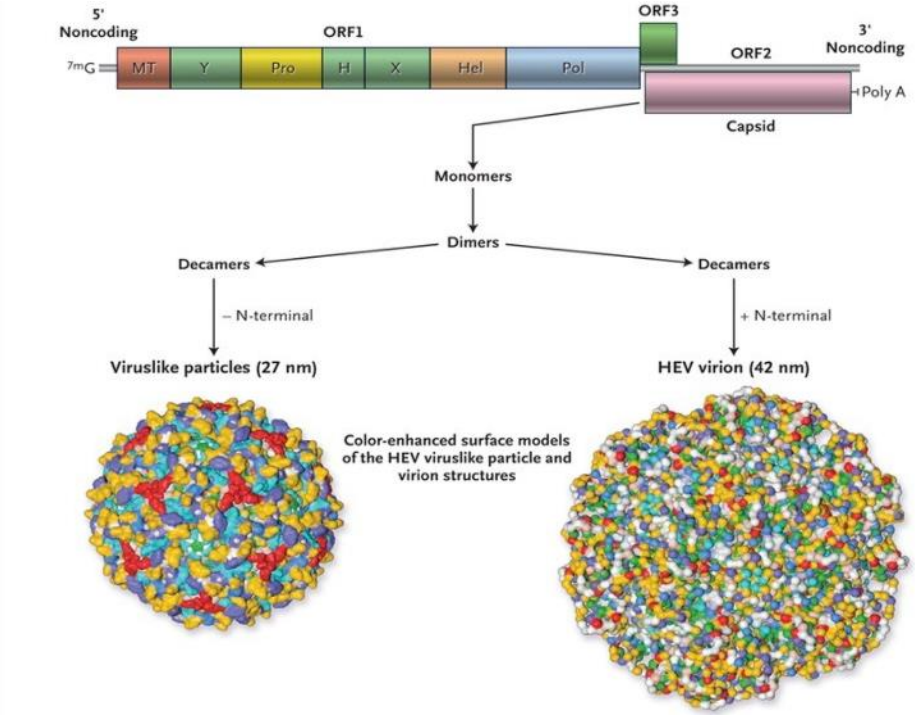


Fig. 4: Hepatitis E (HEV) and its genome
(Hoofnagle JH et al., 2012)

HEV is classified in at least four different main genotypes, from 1 to 4, and in 24 subtypes (Lu L et al., 2009). Nonetheless, HEV has only a serotype. Genotype 1 is the most frequent cause of sporadic Hepatitis E epidemic in developing countries; while genotype 2 was identified for the first time during the 1986 outbreak in Mexico and later in Chad and in Nigeria (Velázquez O et al., 1990). Meng et al., (1997) first described an isolated case of HEV, which was genetically divergent from Burmese and Mexican strains and was widespread in porcine population.

This strain is eventually isolated from a US case which occurred in a person whereby no clear risk factor of infection was identified (Kwo PY et al., 1997). HEV genotype 3 is globally predominant in the swine population and is now increasingly detected in human cases in the developing world. Genotype 4 was first described in Taiwan and afterwards in China, Japan and India. Genotypes 3 and 4 were isolated also from pigs living in United States, Africa, Japan and India. There are clear differences in epidemical potential within various genotypes; epidemics occur exclusively in developing countries, where the predominant circulating human strain is HEV 1 genotype (Teshale EH et al., 2010).

During epidemical conditions, HEV is principally transmissible by drinking water contaminated from faeces. In Southeast Asia, pandemics usually take place during rainy seasons, when floods can contaminate drinking water supplies (Viswanathan R, 1957). Nevertheless, several outbreaks occurred also during dry seasons or in conditions where there were neither evident sign of flooding nor of contamination from drinking water supplies.

Recent results suggest that there is a significant person-to-person transmission in epidemic situations, even though it is unclear whether this way/mode of transmission is comparable

to a case of person-to-person transmission of *Hepatitis A* infection (Teshale EH et al., 2010). There have been reports of transfusion-related transmission and nosocomial transmission of isolated cases of HEV infection (Robson SC et al., 1992; Arankalle VA et Chobe LP, 1999). Vertical transmission of HEV from a pregnant woman to her unborn foetus is very well documented. Khuroo et al., (2009) studied foetal results of HEV infection in pregnant women and detected transmission in utero with foetal outcomes ranging from intrauterine foetal death to symptomatic and asymptomatic neonatal infection. There has currently been no evidence for sexual transmission of HEV.

The incubation period of HEV infection ranges between 15 days to 60 days. HEV causes a number of clinical manifestations, including asymptomatic infection and jaundice hepatitis. Hepatitis E is an acute disease with sudden onset of non-specific symptoms, followed by right upper quadrant pain, jaundice, anorexia, malaise, nausea, vomiting and, in most severe conditions, can lead to death due to fulminant liver failure.

Some HEV outbreak investigations have shown that HEV is not just a deadly disease for pregnant women and very young children. There are actually growing evidences which demonstrate that HEV can cause a fatal disease in healthy young adults and HEV superinfection can also increase mortality in stable individuals with chronic liver disease (Teshale EH et Hu DJ, 2011).

3.1 Diagnosis and treatment

HEV infection causes an increase in both immunoglobulin M (IgM) and anti-HEV IgG antibodies. Anti-HEV IgM response is rapid, it takes place approximately one month after

infection and peaks at the onset of biochemical abnormalities and/or symptoms. RNA of HEV can be detected in both blood and stool at the peak of an acute serological response.

There are many commercial enzyme immunoassays to detect anti-HEV IgM and IgG in serum; although there is considerable variability in their sensitivity and specificity, therefore it makes to obtain a diagnosis of HEV infection difficult. Reverse transcription polymerase chain reaction can be used to identify RNA of HEV in serum and stool, but is usually not available in commercial laboratories. Interpretation of test results can be difficult and a definitive diagnosis generally requires: the use of multiple tests, consideration of risk factors, and exclusion of other causes of acute hepatitis, in regions of low endemicity.

There is no specific therapy for hepatitis E and treatment is currently supportive. The disease typically resolves within 4-6 weeks from the beginning of symptoms, usually with no long-term consequences (De Cock KM et al., 1987). Patients with severe hepatitis E complications require hospitalization and it is generally considered that this disease may be more severe than hepatitis A (Chau TN et al., 2006). Vulnerable populations, such as pregnant women and people with pre-existing chronic liver disease should be identified and receive necessary supportive treatment. Immunoglobulins have not been yet demonstrated to be effective in preventing hepatitis E in people affected by HEV infection, although there are some evidences to suggest that a previous infection can be protective from this disease (Khuroo MS et Dar MY, 1992).

3.2 Vaccination

In Nepal, a phase II recombinant HEV vaccine study carried out expressing the ORF2 protein in baculovirus demonstrated that its vaccine is safe and highly effective (Shrestha MP et al.,

2007). Nonetheless, this study included men only and therefore, it did not provide information on the vaccine's safety and efficacy in women and children. Zhu et al., (2010) reported a phase III study on another recombinant vaccine expressed in *E. coli*. The study included nearly 110,000 randomized people to receive the vaccine or placebo (hepatitis B vaccine). This study found the vaccine to be 99% effective, although it did not enrol children and pregnant women. There are other vaccines in different stages of development.

Two vaccine candidates, which underwent phase II and phase III studies respectively, are truncated ORF2 recombinant proteins of SAR55 strain. Since HEV has only one serotype, these vaccines are expected to be equally protective against infection from any of the four HEV genotypes.

Considering Asian recurrent outbreaks in its population, population-wide prevalence of immunity markers against HEV is not as high as that one against hepatitis A; it seems plausible that this vaccine may prove to be useful for prevention against HEV transmission during epidemics in these countries.

Nowadays, the vaccine used against HEV is Hecolin®, which is a recombinant vaccine utilised to prevent against this virus, developed and manufactured by Xiamen Innovax Biotech Co., Ltd. in Xiamen, China. In China, the vaccine is approved to use in people aged between 16 and over and is recommended for individuals at high risk of HEV infection. Hecolin® is well tolerated and has been shown to be safe to administer in adults.

Main complications associated with its use are mostly local reactions at its injection site. Current evidence demonstrates that this vaccine is highly immunogenic, having nearly all recipients which seroconvert after three doses administered in a 0.1- and 6-month schedule.

Hecolin® is prepared using a genetically modified strain of *E. coli* and is based on a 239 amino acid long recombinant HEV peptide, called "HEV 239", corresponding to amino acids 368-606 of ORF2 which encodes the HEV capsid protein. This amino acid sequence is derived from a Chinese genotype 1 HEV strain. While the results of recent vaccine studies are promising, many questions remain to be answered before the vaccine can be implemented for field use.

The most important matter is the safety and efficacy of a vaccine among the most vulnerable populations, including pregnant women, young children and those with pre-existing chronic liver disease. Other issues include the effectiveness of a vaccine when used in an immediate post-exposure situation to prevent and control transmission during an outbreak, cost of the vaccine, duration of vaccine-induced immunity, and when to vaccinate (Teshale EH et Hu DJ, 2011).

CHAPTER IV

Aim of the thesis

To develop a multivalent recombinant vaccine candidate against malaria and against hepatitis E in *Pichia pastoris.* The research was conducted via a collaboration between the Department of Pharmaceutical and Health Sciences, Section of Biochemistry, of the University of Catania and Etna Biotech S.r.l, Catania, Italy, for Phd research project *(PON RI 2014-2020) Plasmodium falciparum* gene predominant in pre-erythrocyte phase of each parasite coding for CSP was selected as a target (Fig. 5), together with two additional proteins: PfF2 (Fig. 6) and MSPFu24 (Fig. 7) in erythrocyte phase.

MMRKLAILSVSSFLFVEALFQEYQCYGSSSNTRVLNELNYDNAGTNLYNE
LEMNYYGKQENWYSLKKNSRSLGENDDGNNNNGDNGREGKDEDKRD
GNNEDNEKLRKPKHKKLKQPADGNPDPNANPNVDPNANPNVDPNAN
PNVDPNANPNANPNANPNANPNANPNANPNANPNANPNANPNANP
NANPNVDPNANPNANPNANPNANPNANPNANPNANPNANPNANP
NANPNANPNANPNANPNANPNANPNANPNANKNNQGNGQGHNM

Fig. 5: Amino acid sequence of CSP

MEKREHIDLDDFSKFGCDKNSVDTNTKVWECKNPYILSTKDVCVPPRRQEL
CLGNIDRIYDKNLLMIKEHILAIAIYESRILKRKYKNKDDKEVCKIINKTFADI
RDIIGGTDYWNDLSNRKLVGKINTNSKYVHRNKKNDKLFRDEWWKVIKK
DVWNVIWVFKDKTVCKEDDIENIPQFFRWFSEWGDDYCQKMIETLKVECK
EKPCEDDNCKSKCNSYKEWISKKKEEYNKQAKQYQEYQKGNNYKMYSEF
KSIKPEVYLKKYSEKCSNLNFEDEFKEELHSDYKNKCTMCPEVKDVPISIIRN
NEQTSQEAVPEENTEIAHRTETPSISEGPKGNEQKERDDD

Fig. 6: Amino acid sequence of PfF2

MAKNAYEKAKNAYQKANQAVLKAKEASSYDYILGWEFGGGVPEHKK
EENMLSHLYVSSKDKENISKENDAMGNISQHQCVKKQCPQNSGCFR
HLDEREECKCLLNYKQEGDKCVENPNPTCNENNGGCDADAKCTEED
SGSNGKKITCECTKPDSYPLFDGIFCSS

Fig. 7: Amino acid sequence of MSPFu24

For hepatitis E (HEV), P239 (Fig. 8), which represents a highly immunogenic element of the virus capsid protein, was selected and subsequently added to 5’ end of the malarial gene.

Fig. 8: Amino acid sequence of HEV P239

These fusion genes were then expressed as recombinant proteins in *Pichia pastoris* to develop a "combined Malaria-hepatitis E" vaccine with "additive" or "synergistic" effects. Starting from the procedure of design and cloning recombinant HEV-CSP, HEV-PfF2 and HEV-MSPFu24 genes, this research activity evolved in developing their expression in *Pichia pastoris* yeast cells until their extraction and partial purification.

CHAPTER V

Materials and Methods

5.1 Design and synthesis of the recombinant HEV-CSP synthetic gene in intermediate plasmid pMS-RQ

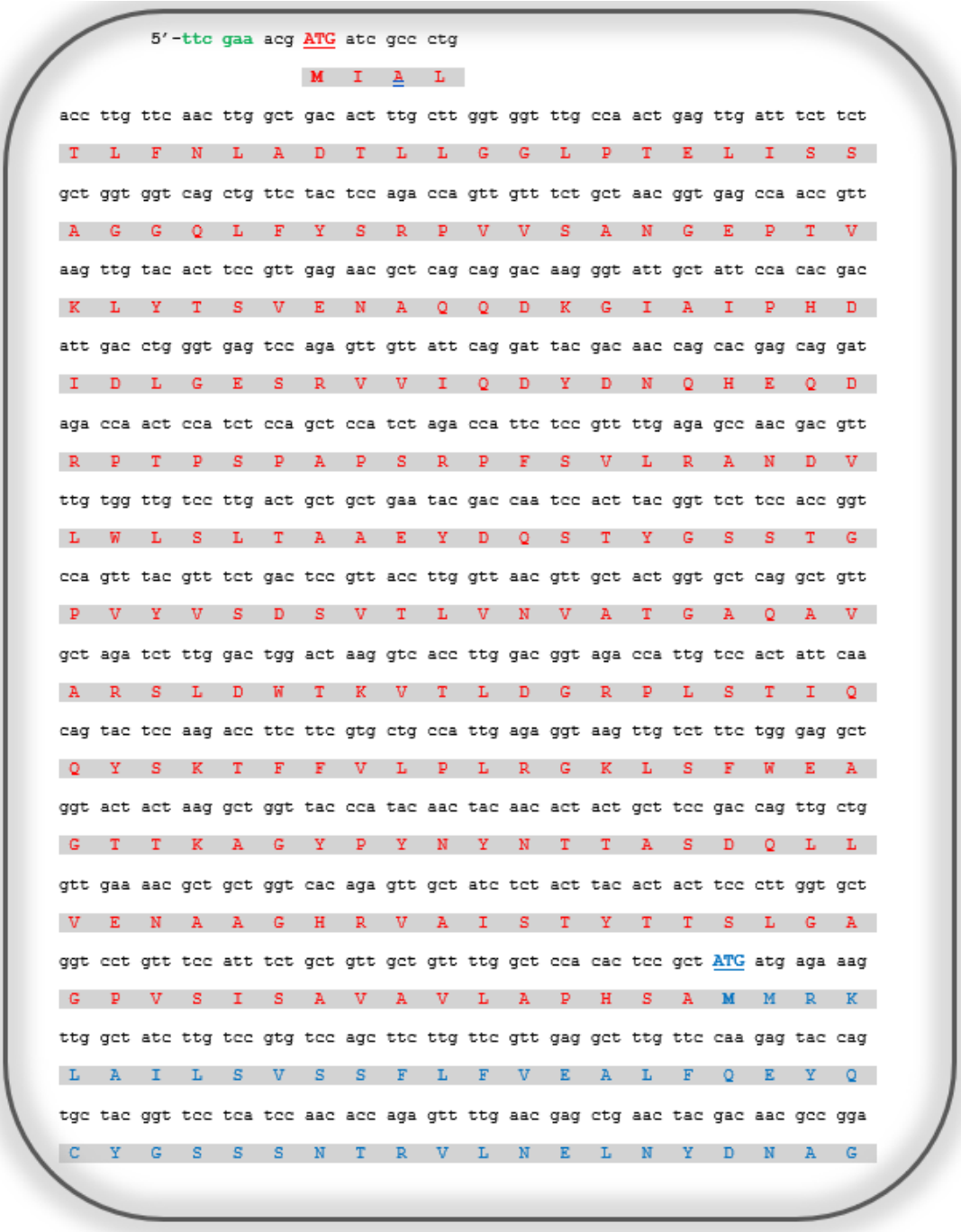
The circumsporozoite (CS) surface protein gene is based on EMBL DNA sequence CQ830509 and EMBL protein sequence CAH04007, in which the C-terminal 14 amino acids are truncated. N-terminal sequence is a consensus assembled by alignment of various sequences present in GenBank, while C-terminal sequence is based on the 3D7 P falciparum sequence. For hepatitis E virus capsid protein gene (GenBank: BAG09239.1) a part from amino acid 368 to 606 was selected in ORF2, corresponding to the smallest region of the protein, still capable of forming capsid structure.

Figure 9 shows the sequence of HEV-CSP fusion gene with considerable modifications for the further cloning and subsequent expression in *Pichia pastoris*.

Unique restriction sites *BstBI* and *NotI* have been added respectively upstream and downstream of the “Open Reading Frame” (ORF) of the fusion gene.

- *BstBI* (TTCGAA) site in position 5 'of the gene
- *NotI* (GCGGCCGC) site in position 3 'of the gene
- A STOP codon TAA has been inserted at the 3’ end of the fusion gene in order to terminate the translation of the protein at this specific point.

- Codon usage was adapted to the codon bias of Pichia pastoris genes. In addition, regions of very high (> 80 %) or very low (< 30 %) GC content have been avoided where possible.



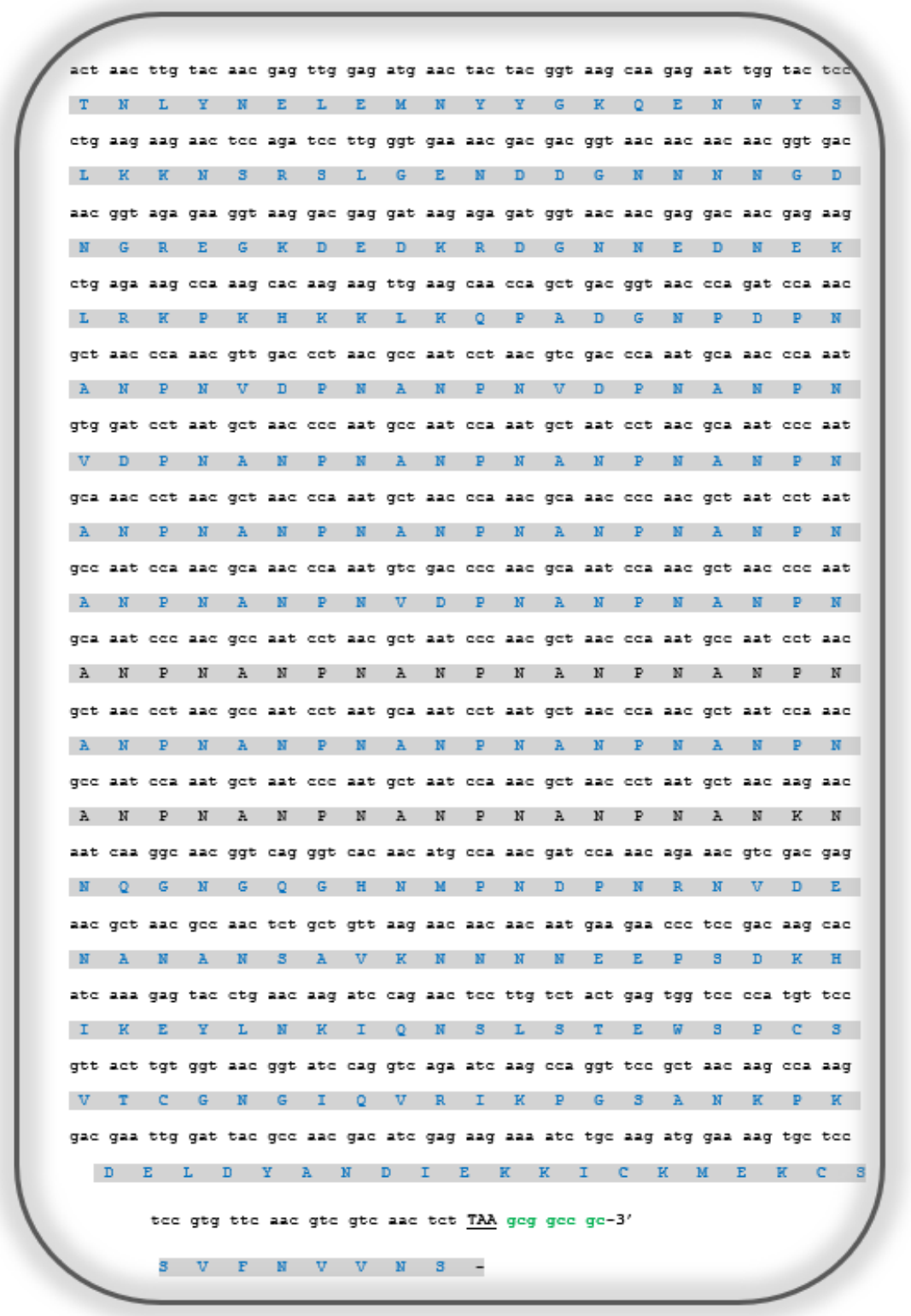


Fig. 9: HEV-CSP fusion gene and relative protein sequence optimized for Pichia pastoris expression. In green are the restriction enzymes sites, red highlighted is HEV ORF and blue highlighted is the CSP ORF

During optimization negative, cis-acting sites (such as splice sites, TATA-boxes, etc.) which may negatively influence expression, were eliminated wherever possible. GC content was adjusted to prolong mRNA half-life. The parameter CAI (codon adaptation index) describes how well codons match the codon usage preference of the target organism. Thus, a CAI of 1.0 would be perfect. However, a CAI of > 0.9 is considered very good (i.e. allowing high expression). Codon usage was adapted to the bias of Pichia pastoris resulting in a CAI* value of 0.87.

The optimized gene should therefore allow high and stable expression rates in *Pichia pastoris*. A gene synthesis was carried out by Thermo Fisher Scientific (Milan, Italy) and fusion HEV-CSP gene was cloned into the intermediate plasmid 17ADBLEC_HEV-CSP_pMS-RQ (Fig. 10).

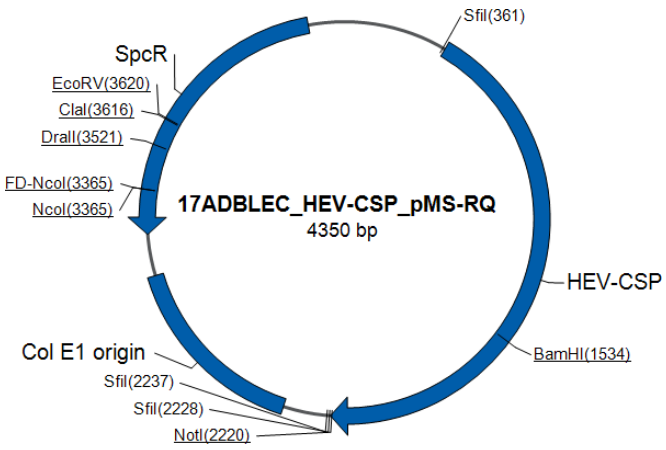
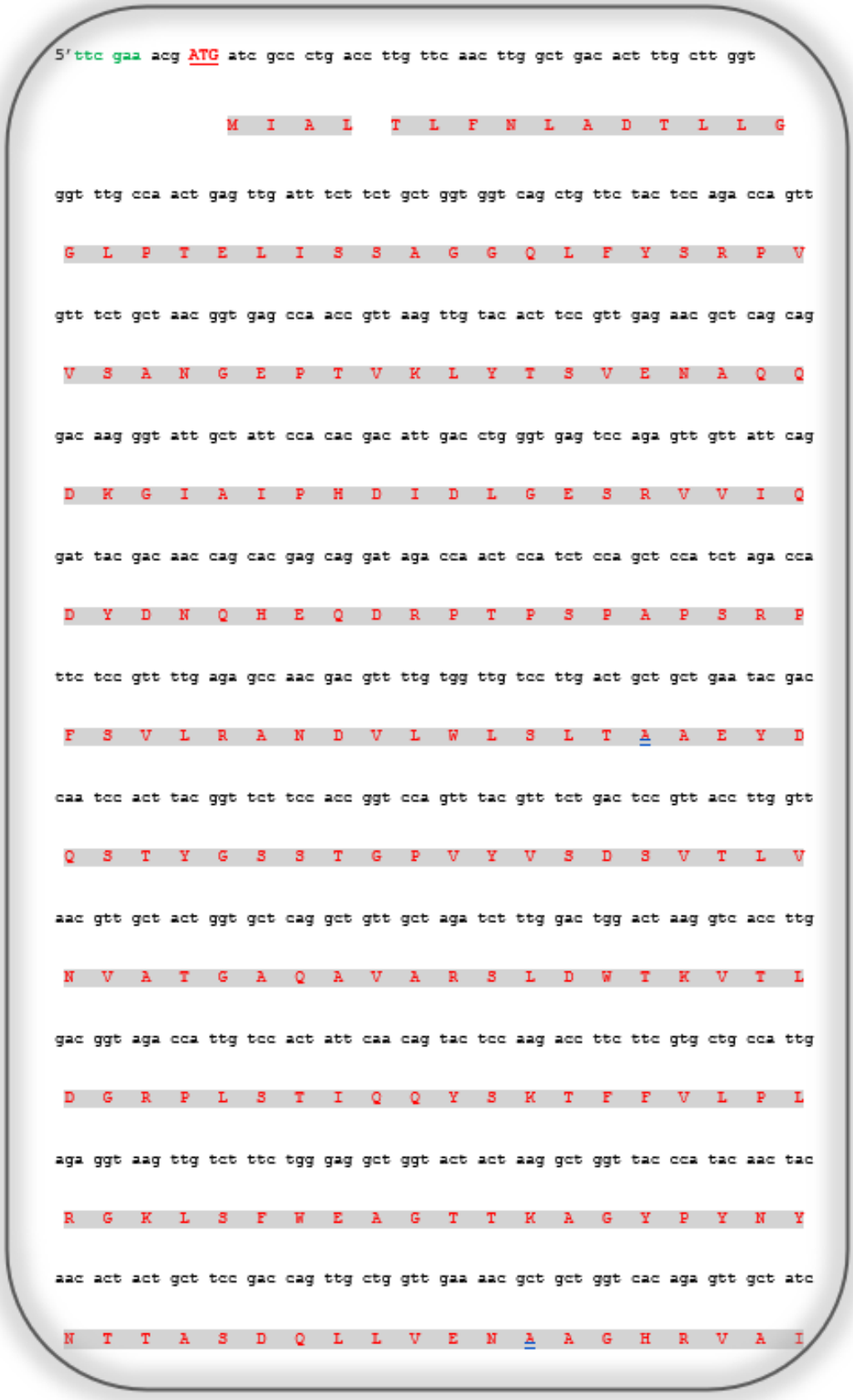


Fig 10: Plasmid map: synthetic gene HEV-CSP was inserted into pMS-RQ. Plasmid DNA was purified from transformed bacteria and concentration determined by UV spectroscopy. The final construct was verified by sequencing. The sequence congruence within the insertion sites was 100%

5.2 Design and synthesis of the recombinant HEV-PfF2 synthetic gene in intermediate plasmid pMA-T

EBA-175 is a *Plasmodium falciparum* protein that binds human erythrocyte receptor glycophorin A (GpA) during invasion. EBA is formed by several domains and the one taken in examination is DBL, which is Duffy binding like protein formed by 61 aa. The DBL is formed by two fragments F1 (residues 8-282) and F2, rich in cysteines (residues 297-603); they have a very similar structure. The F2 domain alone can bind erythrocytes.

Figure 11 shows the design of recombinant HEV-PfF2 protein, in which PfF2 gene has been fused upstream with p239 of HEV capsid protein in a single gene sequence called HEV-PfF2 (Fig. 11). Furthermore, in position 5 ' a single restriction site for BstBI was inserted at the level of the start codon (ATG), while in position 3' another single restriction site for NotI was inserted at the level of stop codon (TAA); the latter is necessary for the subsequent cloning steps in pPICZαA expression vector.


5'ttc gaa acg ATG atc gcc ctg acc ttg ttc aac ttg gct gac act ttg ctt ggt
M I A L T L F N L A D T L L G
ggt ttg cca act gag ttg att tct tct gct ggt ggt cag ctg ttc tac tcc aga cca gtt
G L P T E L I S S A G G Q L F Y S R P V
gtt tct gct aac ggt gag cca acc gtt aag ttg tac act tcc gtt gag aac gct cag cag
V S A N G E P T V K L Y T S V E N A Q Q
gac aag ggt att gct att cca cac gac att gac ctg ggt gag tcc aga gtt gtt att cag
D K G I A I P H D I D L G E S R V V I Q
gat tac gac aac cag cac gag cag gat aga cca act cca tct cca gct cca tct aga cca
D Y D N Q H E Q D R P T P S P A P S R P
ttc tcc gtt ttg aga gcc aac gac gtt ttg tgg ttg tcc ttg act gct gct gaa tac gac
F S V L R A N D V L W L S L T A A E Y D
caa tcc act tac ggt tct tcc acc ggt cca gtt tac gtt tct gac tcc gtt acc ttg gtt
Q S T Y G S S T G P V Y V S D S V T L V
aac gtt gct act ggt gct cag gct gtt gct aga tct ttg gac tgg act aag gtc acc ttg
N V A T G A Q A V A R S L D W T K V T L
gac ggt aga cca ttg tcc act att caa cag tac tcc aag acc ttc ttc gtg ctg cca ttg
D G R P L S T I Q Q Y S K T F F V L P L
aga ggt aag ttg tct ttc tgg gag gct ggt act act aag gct ggt tac cca tac aac tac
R G K L S F W E A G T T K A G Y P Y N Y
aac act act gct tcc gac cag ttg ctg gtt gaa aac gct gct ggt cac aga gtt gct atc
N T T A S D Q L L V E N A A G H R V A I


tct act tac act act tcc ctt ggt gct ggt cct gtt tcc att tct gct gtt gct gtt ttg
S T Y T T S L G A G P V S I S A V A V L
gct cca cac tcc gct ATG gaa aag aga gaa cac att gac ttg gac gac ttc tcc aag ttc
A P H S A M E K R E H I D L D D F S K F
ggt tgc gac aag aac tcc gtt gac acc aac act aag gtt tgg gag tgc aag aac cct tac
G C D K N S V D T N T K V W E C K N P Y
atc ttg tcc acc aag gac gtt tgt gtt cca cca cgt aga caa gag ttg tgc ttg ggt aac
I L S T K D V C V P P R R Q E L C L G N
atc gac aga atc tac gat aag aac ctg ctg atg atc aaa gag cac atc ctg gct atc gct
I D R I Y D K N L L M I K E H I L A I A
atc tac gag tcc aga atc ctg aag aga aag tac aag aac aag gac gac aaa gag gtc tgc
I Y E S R I L K R K Y K N K D D K E V C
aag atc atc aac aag act ttc gcc gac atc aga gac atc atc ggt ggt act gat tac tgg
K I I N K T F A D I R D I I G G T D Y W
aac gac ctg tcc aac aga aag ctg gtc ggt aag atc aac act aac tcc aag tac gtc cac
N D L S N R K L V G K I N T N S K Y V H
aga aac aag aag aac gac aag ctg ttc aga gat gag tgg tgg aag gtc atc aag aag gac
R N K K N D K L F R D E W W K V I K K D
gtc tgg aac gtt atc tcc tgg gtc ttt aag gac aag acc gtc tgt aaa gag gac gac atc
V W N V I S W V F K D K T V C K E D D I
gag aac atc cca cag ttc ttc aga tgg ttc tct gaa tgg ggt gac gac tac tgt caa gac
E N I P Q F F R W F S E W G D D Y C Q D

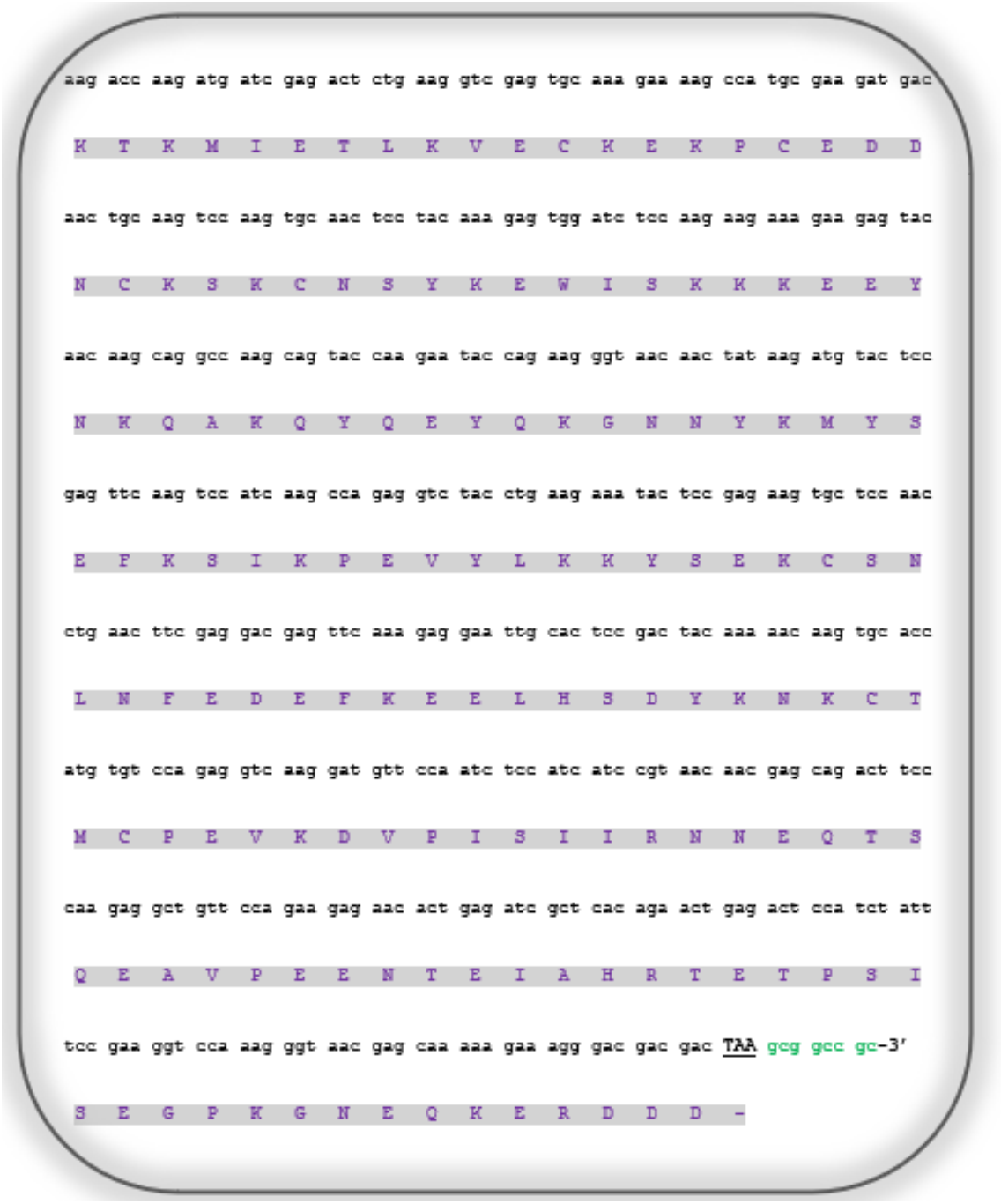


Fig. 11: HEV-PfF2 fusion gene and relative protein sequence optimized for *Pichia pastoris* expression. In green are the restriction enzymes sites, red highlighted is HEV ORF and purple highlighted is the PfF2 ORF

During optimization negative, cis-acting sites (such as splice sites, TATA-boxes, etc.) which may negatively influence expression, were eliminated wherever possible. GC content was adjusted to prolong mRNA half-life (Fig.12).

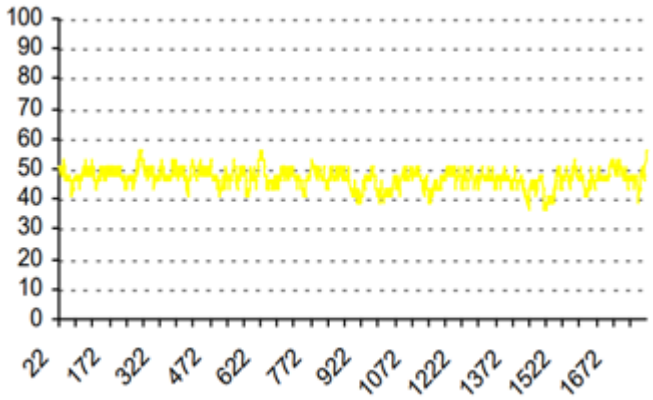


Fig. 12: Codon Quality Distribution of GC

Codon usage was adapted to the bias of *Pichia pastoris* resulting in a CAI* value of 0.85. The optimized gene should therefore allow high and stable expression rates in *Pichia pastoris*. A gene synthesis was carried out by Thermo Fisher Scientific (Milan, Italy) and fusion HEV-PfF2 gene was cloned into the intermediate plasmid 17AEBIBC_HEV-PfF2_ pMA-T of 4206 base pairs (Fig. 13).

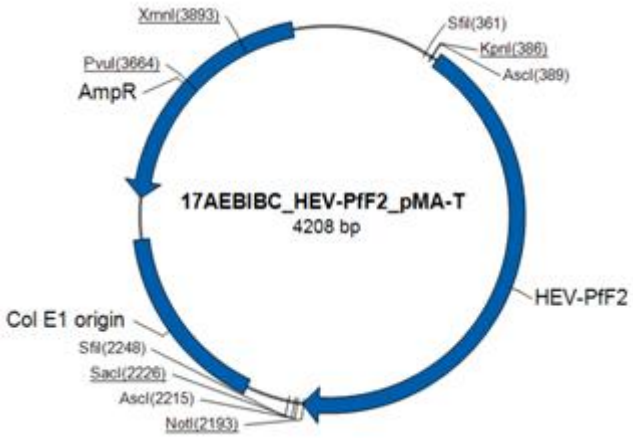


Fig. 13: Plasmid map: The synthetic gene HEV-PfF2 was inserted into pMA-T. Plasmid DNA was purified from transformed bacteria and concentration determined by UV spectroscopy. Final construct was verified by sequencing. Sequence congruence within the insertion sites was 100%

5.3 Design and synthesis of the recombinant HEV-MSPFu24 synthetic gene in intermediate plasmid pMA-T

The chimeric *Plasmodium falciparum* gene, PfMSPFU24, was constructed by coupling genetically immunodominant and highly conserved regions of two merozoite surface proteins present in the erythrocyte phase of a parasite cycle: 19-kDa C-terminal region of merozoite 1 (PfMSP1-19) and highly conserved 24-kDa region of merozoite surface protein 3 (PfMSP3-24). They were used to increase immunogenic potential of this malaria vaccined candidate, MSPFu24 was fused at the 5 'end with p239 of HEV in a single gene sequence called HEV-MSPFu24 (Fig. 14) to which *BstBI* restriction site was added in 5' before ATG start codon, and in 3' a stop codon (TAA) and site of restriction NotI, for subsequent cloning into pPICZαA expression vector.

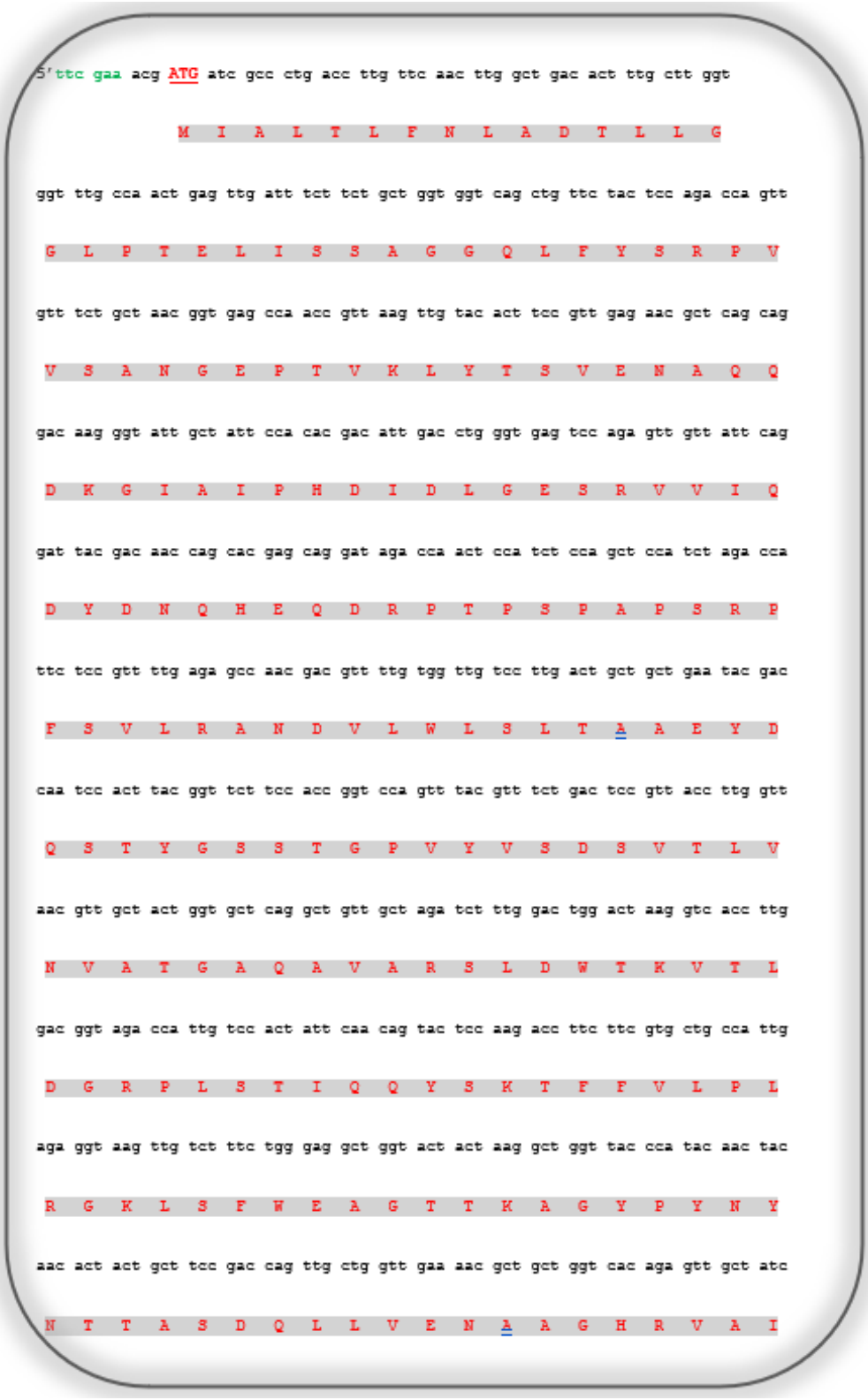

5'ttc gaa acg ATG atc gcc ctg acc ttg ttc aac ttg gct gac act ttg ctt ggt
M I A L T L F N L A D T L L G
ggt ttg cca act gag ttg att tct tct gct ggt ggt cag ctg ttc tac tcc aga cca gtt
G L P T E L I S S A G G Q L F Y S R P V
gtt tct gct aac ggt gag cca acc gtt aag ttg tac act tcc gtt gag aac gct cag cag
V S A N G E P T V K L Y T S V E N A Q Q
gac aag ggt att gct att cca cac gac att gac ctg ggt gag tcc aga gtt gtt att cag
D K G I A I P H D I D L G E S R V V I Q
gat tac gac aac cag cac gag cag gat aga cca act cca tct cca gct cca tct aga cca
D Y D N Q H E Q D R P T P S P A P S R P
ttc tcc gtt ttg aga gcc aac gac gtt ttg tgg ttg tcc ttg act gct gct gaa tac gac
F S V L R A N D V L W L S L T A A E Y D
caa tcc act tac ggt tct tcc acc ggt cca gtt tac gtt tct gac tcc gtt acc ttg gtt
Q S T Y G S S T G P V Y V S D S V T L V
aac gtt gct act ggt gct cag gct gtt gct aga tct ttg gac tgg act aag gtc acc ttg
N V A T G A Q A V A R S L D W T K V T L
gac ggt aga cca ttg tcc act att caa cag tac tcc aag acc ttc ttc gtg ctg cca ttg
D G R P L S T I Q Q Y S K T F F V L P L
aga ggt aag ttg tct ttc tgg gag gct ggt act act aag gct ggt tac cca tac aac tac
R G K L S F W E A G T T K A G Y P Y N Y
aac act act gct tcc gac cag ttg ctg gtt gaa aac gct gct ggt cac aga gtt gct atc
N T T A S D Q L L V E N A A G H R V A I

Fig. 14: HEV-MSPFu24 fusion gene and relative protein sequence optimized for *Pichia pastoris* expression. In green are restriction enzymes sites, red highlighted is HEV ORF and orange highlighted is MSPFu24 ORF

During optimization negative cis-acting sites (such as splice sites, TATA-boxes, etc.) which may negatively influence expression, were eliminated wherever possible. GC content was adjusted to prolong mRNA half-life. Parameter CAI (codon adaptation index) describes how well codons match codon usage preference of the target organism. Thus, a CAI of 1.0 would be perfect. However, a CAI of > 0.9 is considered as very good (in other words, allowing high expression). Codon usage was adapted to the bias of Pichia pastoris resulting in a CAI* value of 0.87 (Fig. 15).

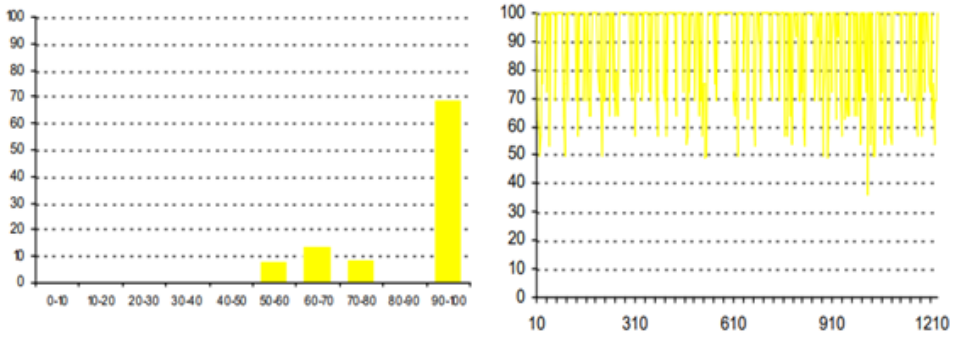


Fig. 15: Histograms show the percentage of sequence codons which fall into a certain quality class. The quality value of the most frequently used codon for a given amino acid in the desired expression system is set to 100, remaining codons are scaled accordingly

The optimized gene should therefore allow high and stable expression rates in. A gene synthesis was carried out by Thermo Fisher Scientific (Milan, Italy) and the fusion HEV-MSPFu24 gene was cloned into intermediate plasmid 17AEBH7C_HEV-MSPFU24_pMA-T di 3656 base pairs (Fig. 16).

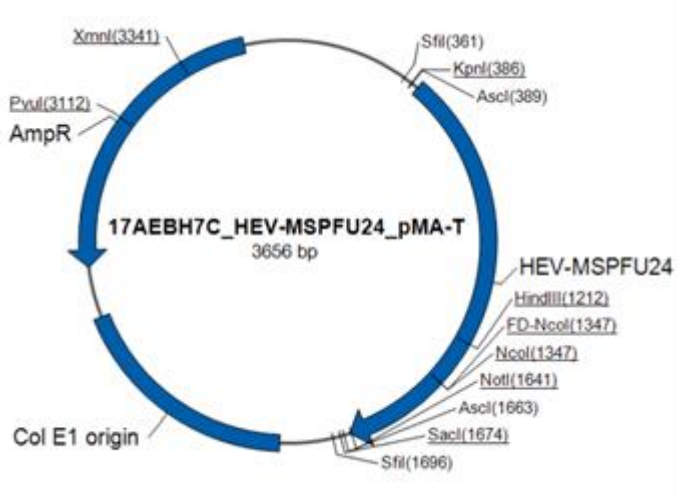


Fig. 16: Plasmid map: Synthetic gene HEV-MSPFu24 was inserted into pMA-T. Plasmid DNA was purified from transformed bacteria and concentration determined by UV spectroscopy. The final construct was verified by sequencing. The sequence congruence within insertion sites was 100%

5.4 *Pichia pastoris* KM71

Pichia pastoris KM71 was chosen as the expression vector. It represents a methylotrophic yeast that has been reassigned to the genus Komagataella, following phylogenetic analyses and called *K. phaffii* or *K. Pichia pastoris* (Kurtzman CP, 2009). This yeast is widely used for the biotechnological production of human proteins due to some advantages it offers, including: a rapid growth rate which allows a high cell density to be obtained in a short time; presence of a strong and inducible promoter; expression of heterologous proteins with an almost always correct tertiary structure (Lu L et al., 2006).

The yeast uses methanol as a carbon source which, by the action of peroxisomal alcohol oxidase (aox) enzyme, converts methanol into formaldehyde and hydrogen peroxide. The latter is detoxified within peroxisome, while formaldehyde is transported outside organelle and used as an energy source. aox enzyme, a homo-octamer that uses flavin adenine

dinucleotide (FAD) as a coenzyme, has low affinity for molecular oxygen and its synthesis is induced at the gene level by methanol.

Within *Pichia pastoris* genome there are two genes that code for two isoforms: aox1, inducible isoform, and aox2, constitutive isoform. However, most of enzymatic aox activities are due to the aox1 isoform. To allow cloning of a heterologous gene in this system it is necessary that the gene is stably integrated into the yeast genome through a stable homologous recombination; this process is possible only with a correct design of the expression vector. In particular, the cloning of transgenes must take place in such a way that its sequence is flanked upstream by a sequence homologous to the one present at 5' of the aox1 gene in the host genome; while downstream by a sequence homologous to the one belonging to a terminator of the same genes, to allow crossing-over and therefore the integration of the transgenes (Lu L et al., 2006).

The use of alternative *Pichia pastoris* strains compared to wild-type strains, such as KM71, which possesses an aox1 gene which is widely disrupted therefore not functioning, allows to considerably improve the expression profile of heterologous proteins. In this case, the mutant strain assumes MutS phenotype, in other words, with a low rate of use of methanol; whose metabolism is entrusted only to the product of the aox2 gene, therefore growing much more slowly. At the same time, methanol being poorly metabolized, it will go more to induce the aox1 promoter, stimulating only the expression of transgenes.

5.5 Bacterial and yeast strains, plasmids and culture conditions

The cloning and expression vector used was pPICZαA (Fig. 17) of 3.6 Kb (Invitrogen TM, Milan, Italy), which allows selection at different concentrations of Zeocin® (Sigma-Aldrich, Milan, Italy) of bacterial transformants, in phase of gene cloning, and of yeast, during the expression phase of heterologous protein.

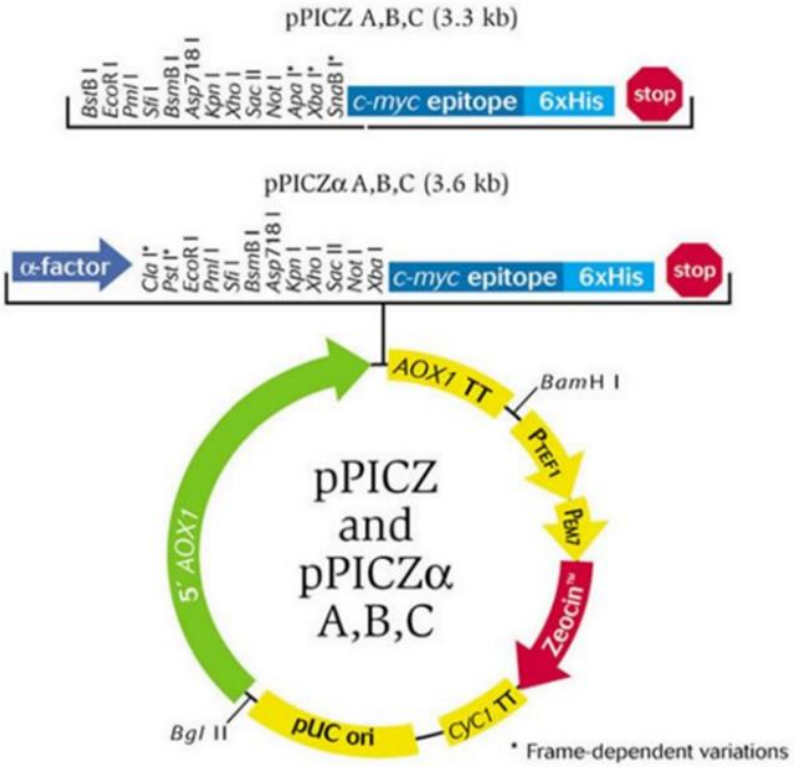


Fig. 17: pPICZαA cloning vector
(Invitrogen, Life Technologies, Milan, Italy, pPICZαA A, B and C, 2010)

The aforementioned vector contains a 5 'aox1 portion, homologous to the promoter of deleted aox1 gene present in *Pichia pastoris* genome, necessary for integration of the vector into a host genome, resulting in stable transformation. Furthermore, the cloning of in-frame transgene with secretion factor α-factor sequence, deriving from *S. cerevisiae*, allows the secretion of cloned protein, in the form of fusion protein. The ShBle gene, deriving from *Streptoalloteichus hindustanus*, encodes the determinant of resistance to Zeocin® (Sigma-Aldrich, Milan, Italy).

Competent cells used for bacterial transformation were XL-10 Gold Ultracompetent (Stratagene, USA), grown in LB broth (Sigma-Aldrich, Milan, Italy) and selected in Zeocina® (Sigma-Aldrich, Milan, Italy) at a concentration of 25 µg/ml.

MOLECULAR BIOLOGY AND GENE CLONING TECHNIQUES

5.6 Plasmid DNA extraction

Plasmid DNA extraction was performed using *QIAprep Spin Miniprep Kit* (QIAGEN, Milan, Italy) for the mini-preparations and *QIAGEN Plasmid Maxi Purification Kit* (QIAGEN, Milan, Italy) for the maxi-preparations, in accordance with the supplier company protocols. Bacterial colonies, selected on plates with their respective antibiotic, were inoculated in LB broth and incubated at 37 °C under agitation for 12-18 hours.

The broth culture was centrifuged at 6000 g for 5 minutes to recover a bacterial pellet and subsequently resuspended in Buffer P1 (50mM Tris-pH 8.0, 10mM EDTA, 100µg/ml Rnase A). Subsequently, cell lysis was carried out with Buffer P2 (200mM NaOH, 1% SDS) at room temperature for 5 minutes and Buffer P3 was added to improve the precipitation of

DNA (3.0 M potassium acetate, pH 5.5) incubated in ice for 5 minutes. Then the DNA precipitated with 100% isopropanol and washed with 70% ethanol, was recovered by centrifuging at 13,000 x g for 10 minutes and resuspended in sterile H_2O.

5.7 Digestion of DNA with restriction enzymes

Restriction and endonuclease digestion reactions were performed with the following restriction enzymes: *BstBI*, maintained at 65 °C for 15 min, *NotI* and *SacI*, maintained at 37 °C for 15 min, in a final volume of 50 µl. The fragments obtained from enzymatic digestions were separated by electrophoretic analysis on 1% concentrated agarose gel in Buffer TAE (40 mM Tris-acetate, 1 mM EDTA) 1X. These restriction enzymes used were purchased from New England Biolabs Ltd. (Milan, Italy) and used according to a protocol provided by this Company.

5.8 Agarose gel electrophoresis

In various stages of the process, DNA analysis was carried out by electrophoresis on 1% agarose gel in 40 mM Tris-acetate and 1 mM EDTA (TAE), in a Sub Cell GT chamber electrophoretic (Bio-Rad, Milan, Italy). The marker used was 1Kb DNA ladder, while the intercalator was SYBR Green (10 mg/ml) (Thermo Fisher, Milan, Italy). Electrophoretic run was performed at 105 Volts. DNA was visualized by UV transilluminator (KODAK 1D Image Analysis, Thermo Fisher, Milan, Italy).

5.9 DNA purification

The DNA, amplified or digested and subjected to agarose gel electrophoresis, was extracted and purified using QIAGEN kit (Milan, Italy), based on the size and concentrations of the fragments to be purified (MinElute Purification kit, QIAEX II Gel Extraction kit, QIAquick Spin, Milan, Italy). Recovery of DNA fragments of interest from agarose gel was performed using the Gel Extraction Extraction kit (QIAGEN, Milan, Italy). Part of the agarose gel containing the fragment of interest was recovered and placed in a 1.5 ml Eppendorf tube, in which 3 volumes of Buffer QG were added for each volume of gel (for each 100 mg of gel they must be added 300 µl of QG Buffer). This tube was later incubated for 10 min at 50 ° C to allow gel solubilization, shaking every 2-3 min. Subsequently, for each volume of gel, 1 volume of isopropanol was later added; the sample of interest, to allow DNA to bind, was then loaded onto an affinity column, which was centrifuged for 30 sec at 13,000 x g. After eliminating the eluted Buffer, 500 µl of Buffer QG were added to the column and the sample was centrifuged for 30 sec to remove any remaining traces of agarose that could interfere with subsequent enzymatic reactions or DNA sequencing. Once this wash Buffer was removed, 750 µl of ethanol-containing Buffer (PE Buffer) was added to the resin. Traces of ethanol were removed by a further 30 sec centrifugations, after eliminating PE Buffer by centrifugation at 13,000 x g for 30 sec. DNA was then eluted after incubating the column for 1 min with 50 µl of sterile H_2O by placing and centrifuging for 30 sec in a 1.5 ml Eppendorf tube and subsequently stored at -20 °C for subsequent use.

5.10 Ligation and bacterial transformation

DNA fragments obtained by enzymatic restriction were ligated to the vector using T4 DNA Ligase (New England Biolabs, Ltd. Milan, Italy). In detail, to facilitate a meeting between

plasmid and insert, a DNA concentration of the insert 2-3 times higher than that of the vector is used, but it is necessary to consider for the calculation of the exact concentrations the molarity of the vector and of the insert. In our case, 30 ng of vector and about 24 ng of insert were used. The samples were placed overnight at 16 °C in order to promote ligation.

The bacterial transformation was carried out with the ligation products at 100 µl of *E. coli* XL Gold Ultracompetent bacterial cells (Stratagene, USA), then incubated on ice for 30 minutes. Thereafter these bacteria were subjected to heat shock by exposure at 42 °C for 2 min, cooled on ice for 5 min and incubated at 37 °C for 1 h in 1 ml of LB, allowing the growth of cells containing the recombinant plasmid, which presents the gene sequence for resistance to the selection antibiotic (Zeocin®, Sigma-Aldrich, Milan, Italy). Finally, all the suspension was seeded in Petri dishes containing LB-Agar added with 50 µg/ml of Zeocin® and incubated over-night at 37 °C, in order to select the transformed bacteria.

5.11 DNA quantification

DNA quantification was performed using Agilent Technologies (Thermo Fisher, Milan, Italy) spectrophotometer with double reading at λ=280 and λ=260 nm, to also determine the DNA purity. To obtain a good quality DNA, the ratio of absorption values at λ=260 and λ=280 nm must be 1.8. If this ratio is lower, it means that there might be a possible contamination of proteins or phenolic compounds. In these cases, DNA will have to be further purified because such contaminations can interfere with subsequent enzymatic reactions. A 5 µl of sample and 995 µl of sterile H_20 were briefly loaded into quartz cuvettes and then placed in spectrophotometer for a reading at λ=260 and λ=280 nm. DNA concentration was calculated considering: its 50 mg sample, read at λ=260 nm and the using

of a cuvette with an optical path of 1 cm; therefore, we arrived to the conclusion that it has an absorbance value of 1.00. Furthermore, we obtained the concentration of this sample in mg/ml I, through an adequate proportion, based on optical density of our DNA and on relative dilution of the sample carried out.

5.12 DNA sequencing and computerized analysis

During the various phases of the study, the sequences of the protein genes (CSP, MSPFu24, PfF2, HEV and recombinant HEV-CSP, HEV-MSPFu24, HEV-PfF2 genes) were verified by automatic sequencing at MWG Biotech Ltd (Ebersberg, Germany). The computerized analysis of the sequences was carried out using the DNA Strider 1.3 software (Macintosh).

5.13 Induction to competence and transfection in *Pichia pastoris* KM71

The process of induction to *Pichia pastoris* competence immediately precedes transfection, which is necessary to introduce exogenous genetic material into the eukaryotic host. It is a particularly delicate process, since it significantly alters permeability of the membrane and cell wall. The procedure was carried out at a controlled temperature of 4 °C.

A single colony of each construct grown and isolated on YPD medium (1% yeast extract, 2% peptone, 2% dextrose, 2% agar) was inoculated in broth and incubated at 28-30 °C for 15-18 hours under agitation. at 200 x g. The next day, having reached an optical density of 0.8-1.5 at λ=300 nm (OD), broth culture was centrifuged at 5,000 x g for 5' and the cell pellet was resuspended in 20 ml of YPD broth with the addition of 400 µl of HEPES Buffer (acid 4-2-hydroxyethyl-1-piperazinyl-ethanesulfonic) 1M at pH 8.0, and 500 µl of DDT

(dithiothreitol) 1M. These cells were then incubated with shaking at 70-80 x g for 15' at 28-30 °C; then sterile water was added at 4 °C bringing the volume to 100 ml.

The cell suspension was centrifuged at 5,000 x g for 10' at 4 °C, supernatant removed and cell pellet was washed a first time with 100 ml of sterile water at 4 °C and a second time with 50 ml of sterile water at 4 °C. A third wash was then carried out with 10 ml of 1M sorbitol and finally the cell pellet, recovered by centrifugation at 5,000 x g for 5' and resuspended in 1 ml of 1M sorbitol at 4 °C.

Thus, these obtained cells, made competent for transformation, were immediately transfected by electroporation. In particular, 80 µl of competent cells were incubated for 10' on ice in the presence of 5-10 µg of DNA, previously linearized with the *SacI* enzyme to facilitate its entry through the membrane and to promote subsequent stable integration in the host genome. Furthermore, strain KM71 transformed with empty plasmid pPICZαA was always used as a negative control in each electroporation experiment.

The device used for electroporation was Gene Pulser XCell (Bio-Rad, Milan, Italy) with the following parameters: 1500 V of voltage, 25 µF of capacity and 200 Ω of resistance.

Immediately after electroporation, the cells were added with 1 ml of sorbitol and incubated at 28-30 °C without shaking for 1-2 hours.

Finally, electroporated cells were plated on YPD agar with Zeocin® (100 µg/ml) to select the transformants. It should be emphasized that during these stages following the selection of transformants in Zeocin®, it is no longer necessary to make the culture medium selective with an antibiotic, as the stable integration of one or more copies of plasmid into the host genome is an irreversible process.

5.14 Replica plating

Screening of transformants by means of the "replica plating" method is necessary in order to monitor the presence of multiple integrations of recombinant vectors within the host. Via this technique, a number of integrated copies of the plasmid is empirically evaluated by determining the level of resistance to Zeocin® of the transformants, taking into account that the more copies of the *ShBle* gene have been integrated into the genome, the greater the resistance conferred.

Colonies of the selected transformants through using Zeocin®, at the maintenance concentration (100 µg/ml), were collected and inoculated in 24-deep wells plates in antibiotic-coated YPD broth, and incubated at 28-30 °C at 200 x g for 1-2 days (Fig. 18).

Fig. 18: Transforming colonies of HEV-CSP, HEV-MSPFu24 and HEV-PfF2 selected in Zeocin®

Subsequently the clones were plated from broth culture in replicate plates of YPD agar added with increasing concentrations of Zeocin® (1,000 µg/ml, 2,000 µg/ml, 2,500 µg/ml) and incubated at 28-30 °C for 2-3 days, to select the High Resistant (HR) clones to be used in the next phase of induction with methanol (Fig. 19).

Fig. 19: "Replica plating" of HEV-CSP, HEV-MSPFu24 and HEV-PfF2 constructs at different concentrations of Zeocin® (1,000 µg/ml; 2,000 µg/ml; 2,500 µg/ml)

5.15 Methanol induction

After selecting the cells that grew best in presence of high concentrations of antibiotic (Zeocin®), it was necessary to induce transcription of the inserted gene, resulting in expression of the desired protein.

Before inducing transcription with methanol, it was necessary to gradually adapt the recombinant clones, previously selected for *replication plating*, to use alternative carbon sources to dextrose. In particular, they were first inoculated into "*baffled flask*" containing 50-100 ml of BGYG broth (1% yeast extract, 2% peptone, 1.34% Yeast Nitrogen Base, 2% glycerol, 100 mM phosphate Buffer pH 6, 0, Biotin 0.04 mg/mL and 0.004% Histidine), with glycerol as a carbon source, and incubated at 28-30 °C at 200 x g for 48-72 hours, in order to increase cell density.

The cells were then recovered by centrifugation at 5,000 x g for 5'; they were then furtherly washed with BGY broth to eliminate any trace of residual glycerol, and finally they were resuspended in a volume of BGYM broth (1% yeast extract, 2% peptone, 1.34% Yeast Nitrogen Base, 0.5-1% Methanol, 100 mM phosphate Buffer pH 6.0, Biotin 0.04 mg/mL and 0.004% Histidine), equal to 1/5 or 1/10 of initial volume of broth culture. The above procedure was necessary for the KM71 MutS strain, in which the induction of aox1 promoter was not established as long as a trace, albeit minimal, of glycerol remains in the culture medium.

In BGYM broth, methanol acts as an agent capable of inducing any exogenous gene. The induction proceeds for 4-5 days, in incubation at 28-30 °C at 200 x g, adding methanol daily at a desired concentration.

5.16 Extraction of intracellular heterologous proteins

Every 24 hours, 1 ml of broth culture was taken in the methanol induction phase, which were centrifuged at 5,000 x g for 5'. A cell pellet, containing the exogenous protein of interest to be recovered, was resuspended in Lysis Buffer (50 mM Tris-HCl, 150 mM NaCl, 0.5%

Triton, 10 mM EDTA, distilled H_2O, pH 7.8). The sample lysis was then conducted with two different methods, depending on the quantity of pellets obtained.

When there were small quantities (1 ml) this process required the use of glass beads (glass beads size 425-600 µm) (Sigma-Aldrich, Milan, Italy). In detail, the pellets, placed in 1.5 ml Eppendorf, were suspended in 100 µl of Lysis Buffer and subsequently the same volume of glass beads was added. We proceeded with 10 cycles of vertexing at 1,500 x g for 30 sec followed by 30 sec of permanence on ice.

Finally, these samples were centrifuged at 13,000 x g to collect the supernatant, which was used for subsequent Western Blot analysis. Instead, for higher quantities (100 ml) we proceeded with homogenization of the samples using the LM 20 Microfluidizer processor (Fig. 20). In particular, 100 ml of each sample in Lysis Buffer (50 mM tris-HCl, 150 mM NaCl, 0.5% Triton, 10 mM EDTA, pH 7.8) and corresponding to an OD (λ=600 nm) of about λ=300 nm was subjected to three cycles of homogenization at high pressure (28,000 Psi) obtaining a uniform reduction in particle size and a halved optical density.

The cell lysates of each construct were recovered by centrifugation at 12,000 x g for 30 'and subsequently analysed separately by Western blot analysis using a specific antibody to determine the presence of the desired protein relating to each of the constructs.

Fig. 20: Homogenisator LM20 Microfluidizer High Pressure

5.17 SDS-PAGE and Western blot Analysis

Protein identification was performed by SDS-PAGE and subsequent Western blot analysis. SDS-PAGE was performed by electrophoresis on pre-casting polyacrylamide gel at 4-12% NuPAGE Novex Bis-Tris Mini (Life Technologies, Milan, Italy), in native or reducing conditions; thus, this procedure followed instructions provided by the manufacturer, with the aid of the 'Novex MiniCell electrophoretic apparatus (Life Technologies, Milan, Italy) at 200 V constant, in MES Buffer. After this electrophoretic run, to highlight a protein pattern, the gel was stained with Simply Blue Safe Stain Pet (Comassie Blue, Life Technologies, Milan, Italy) for about 60 minutes in an oscillating shaker and subsequently decoloured with various washes in distilled H_2O. These proteins were then highlighted by reading the gel in the ChemiDoc (Bio-Rad, Milan, Italy) on a white plate.

The transfer of proteins, separated by electrophoresis from its gel, to a nitrocellulose membrane (Novex Nitrocellulose Membrane 0.2 µm pore size, Life Technologies, Milan,

Italy) was carried out using XCell II Blot Module (Life Technologies, Milan, Italy) apparatus in Transfer Buffer (Life Technologies, Milan, Italy) at constant 30 V for 1 hour.

A mouse monoclonal anti-HEV antibody (LS-C67675 / 46058, LifeSpan BioSciences, Inc. Milan, Italy) against the aa 434-457 residues of pORF2 of the capsid protein was used to perform the Western blot analysis of all the constructs, at concentration of 1 mg/ml and diluted 1:4,000. Specifically, the membrane was initially treated with Blocking Buffer (20 ml PBS, 1gr milk powder, 0.1% TWEEN 20) (Sigma-Aldrich, Milan, Italy) shaken at 60 x g for 1 hour, at room temperature (RT), to saturate the nonspecific antibody binding sites.

Hybridization with a specific antibody was conducted at 4 ° C overnight to allow binding to each target protein. Three washes were then performed with Washing Buffer (400 ml PBS, 0.5% TWEEN 20) (Sigma-Aldrich, Milan, Italy) to eliminate the excess antibody. This membrane was then incubated for 1 hour RT with a specific secondary anti-mouse antibody conjugated with horseradish peroxidase (HRP, Thermo Fisher, Milan, Italy).

After three washes with Washing Buffer, proteins were detected with ECL (Thermo Fisher, Milan, Italy) standard chemiluminescent substrate for detection of HRP (Thermo Fisher, Milan, Italy) enzyme activity, using the ChemiDoc (Bio-Rad, Milan, Italy) instrument.

5.18 Filtration and Ultrafiltration

The samples and all Buffers used in the purification procedures were filtered with 0.22/0.45 μm filters before their use. The sample of interest was rapidly ultrafiltered with Amicon® Ultra-15 (Merk, Milan, Italy) centrifugal filtration devices (Fig. 21).

These filters are classified according to their nominal molecular weight limit (NMWL) or their molecular weight limit (MWCO). For example, a membrane rated at 30,000 will rule out a test protein with a molecular weight of 30 kDa. Ninety percent of that test protein will be retained on the upstream side and 10% will pass into the filtrate, determining the protein concentration. Also, the type of rotor, with fixed or oscillating angle, and the centrifugation time, were suitably determined on the basis of molecular weights of the proteins to be purified (Fig. 21).

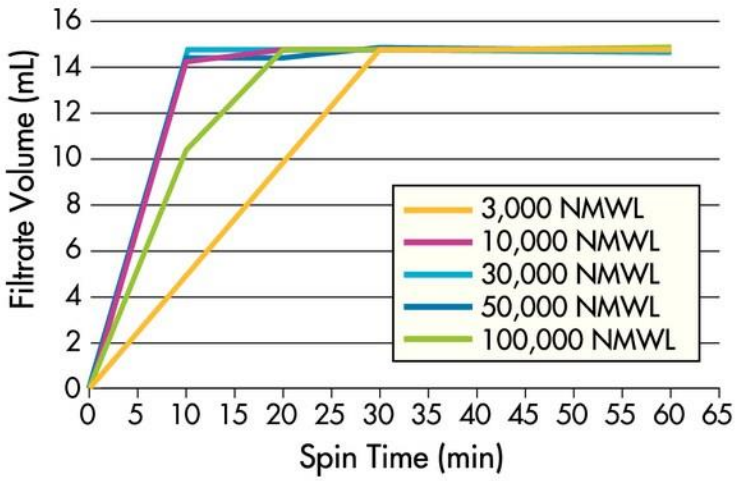


Fig. 21: Specific centrifugation times by volume of filtrate. Initial volume of 15 ml in the swinging rotor at 4,000 x g at 25 °C. Spin times: 3K (40 min); 10K and 100K (20 minutes); 30K (10 minutes); 50K (15 minutes)

In detail, the samples were placed in Amicon ® Ultra-15 (Merk, Milan, Italy) tubes and centrifuged at 4,000 x g for 1 h (Fig. 22). This final concentrate, which was used for the subsequent purification steps, was subsequently recovered.

Fig. 22: Ultrafiltration and concentration system

5.19 Chromatography

The term “chromatography” indicates a set of techniques that aim to separate a mixture into its components to allow qualitative and quantitative recognition. This technique is based on the differential distribution of various components between two phases: one called “the fixed or stationary phase” and the other called “the mobile or eluent phase”, which flows continuously through a fixed phase.

Based on the interactions (hydrogen bonds, dipole-dipole interactions, Van der Waals forces, formation of interaction compounds, steric interactions) that occur between to-be-separated substances and these two phases, where these separation mechanisms are categorised in:

adsorption, partitioning, ion exchange, exclusion and affinity. Therefore, began a purification of our protein by focusing on ion exchange chromatography.

5.20 Ion exchange chromatography (IEC)

This technique is used for separation of ionic or ionizable substances. In short, the stationary phase is constituted by an inert polymer containing ionized or ionizable active sites, whose counter-ions can be exchanged with other ions having a charge of the same sign.

The separation mechanism is based on the competition for the exchange sites between the ions present in the mobile phase and those present in the sample. As a chromatographic technique, we worked with packed column chromatography, where the stationary phase is totally contained within a cylindrical column.

The final resolution of an ion exchange separation is determined by the selectivity and efficiency of the column. These parameters are in turn influenced by various factors such as: grain size, porosity and packing of the column. The separation is affected by how the net surface charge of each protein in the sample varies with pH, the ionic strength of the Buffers, and the elution conditions.

We used 5 ml strong anion exchange (HiTrap Q FF GE Healthcare, Naples, Italy) and cationic (HiTrap SP FF GE Healthcare, Naples, Italy) columns connected to an integrated Akta Pure (*Pure Protein Purification System, GE Healthcare*) high pressure chromatography system with UNICORN ™ hardware and software (Bio-Rad, Milan, Italy) for the integration and processing of chromatographic data and results (Fig. 23).

Fig. 23: AKTA purification system and chromatography column used HiTrap Q FF (GE)

The chromatographic process develops in 4 phases:

Balancing: the column is washed and equilibrated with a Buffer chosen based on the sample characteristics (Start Buffer).

Sample application: Used to bind target molecules. This sample, suitably diluted in the start Buffer and filtered, is applied at a constant rate of 5 ml/minute. Eluate is collected and stored for subsequent testing.

Washing: operation that serves to eliminate everything that is not bound to the resin. It is washed with Start Buffer and the eluate is collected to be subsequently tested.

Elution: bound biomolecules are gradually eluted in an appropriate elution Buffer and the eluate collected in 30 tubes, previously placed in a collector of fractions; they will be subsequently tested to determine how many fractions are situated in the protein of interest.

The elution can be conducted in isocratic conditions, in other words keeping unaltered composition of the mobile phase, or in gradient if one or more parameters of the mobile phase are modified. Gradient may be needed if isocratic conditions do not ensure efficient separation and can be linear or stepped to improve peak resolution (Fig. 24).

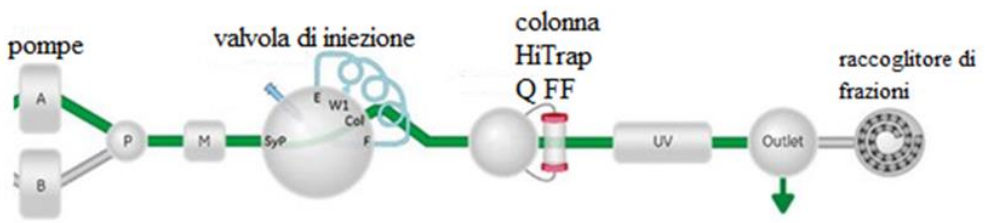


Fig. 24: Chromatographic flow through the Akta Pure Protein Purification System control system

The following steps are necessary for the resin regeneration used and a column appropriate preservation.

Regeneration: once elution is completed, column is regenerated by washing it with a Start Buffer with 1 M NaCl and again with the Start Buffer. The column is now ready for purification of a new sample.

Storage: if no longer used, to preserve the resin, this column is washed with pure water and 20% ethanol to prevent growth of microbes.

CHAPTER VI

Results

Plasmid DNA of each construct and that obtained from pPICZαA expression vector were subjected to ligation process, after suitable digestion with restriction enzymes (*BstBI* and *NotI*), separation on agarose gel and extraction of the bands of each single construct, to obtain recombinant plasmids pPICZαAHEV-CSP (5.1 Kb), pPICZαAHEV-PFf2 (5 Kb) and pPICZαAHEV-MSPFu24 (4.4 Kb). The obtained recombinants were then used to transform *E. coli* XL 10 GOLD competent cells.

6.1 Cloning of recombinant HEV-CSP gene into pPICZαA expression vector

To clone recombinant HEV-CSP gene into 3.3 Kb pPICZαA expression vector, the synthetic gene was inserted into the intermediate vector (17ADBLEC_HEV-CSP_pMS-RQ), selected in LB plates with Spectinomycin, amplified for plasmid maxi-preparation, using QIAGEN Plasmid Maxi Purification Kit (QIAGEN, Milan, Italy) and recovered through digestion with *BstBI* and *NotI*. The pPICZαA expression vector was digested with the same restriction enzymes to obtain compatible ends for the subsequent cloning of its insert.

The *BstBI-NotI* restriction fragments obtained from both vector and insert were purified on agarose gel, using the QIAEX I Gel Extraction kit (QIAGEN, Milan, Italy), and linked by T4 Ligase to obtain the recombinant plasmid: pPICZα-HEV -CSP of 5.1 Kb. The ligation

product was used to transform competent cells of E. coli XL10 GOLD, selecting transformants in LB agar with Zeocin® (Sigma-Aldrich, Milan, Italy).

Selection of transformants was carried out by mini-preparations of plasmid DNA and digestion with restriction enzymes.

To assess the presence of recombinant plasmids, 14 colonies were examined for pPICZαAHEV-CSP: plasmid DNA was extracted from each colony and analysed on 1% agarose gel. Figure 25 shows that DNA of each colony, except clone # 8, is represented by a 5.1 Kb band which coincides with the size of the previously mentioned recombinant plasmid.

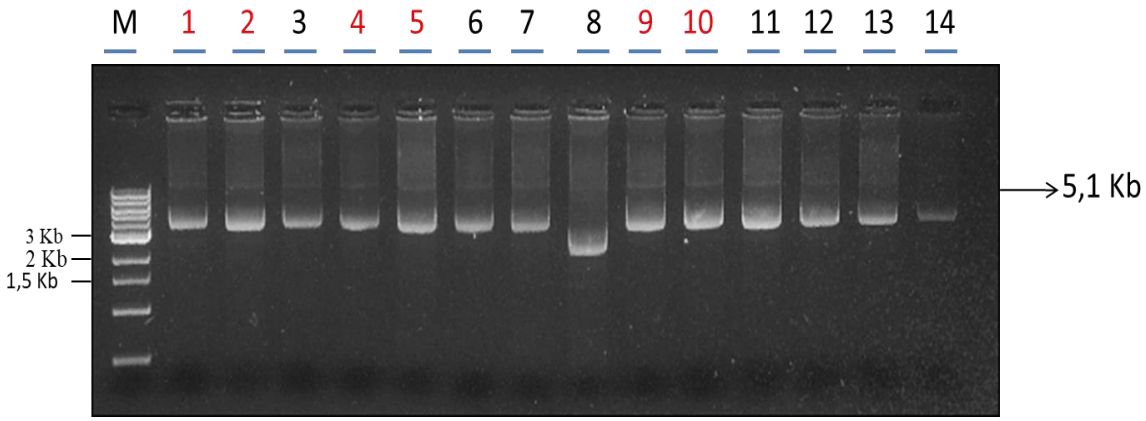


Fig. 25: Screening of transforming bacterial colonies with pPICZαAHEV-CSP recombianant DNA by mini plasmid DNA preparation. Marker: 1 Kb DNA ladder. Clones # 1, # 2, # 4, # 5 # 9 and # 10 highlighted in red were selected for subsequent enzymatic digestions with *BstBI-NotI*

To evaluate the successful integration of HEV-CSP gene into pPICZαA, the enzymatic digestion of clones # 1, # 2, # 4, # 5 # 9 and # 10 was carried out using two cloning enzymes

BstBI and *NotI*. As demonstrated by figure 26, clones # 2, # 4, # 5, # 9 and # 10, show two distinct and separate bands, one relative to vector pPICZαA (3.2 Kb) and the other to HEV-CSP insert (1,8 Kb), indicating the positivity of the clones. Clone # 1 shows also a 5 kb band which corresponds to a not digested plasmid. This information suggests that a partial enzymatic digestion happened and therefore this clone was eliminated.

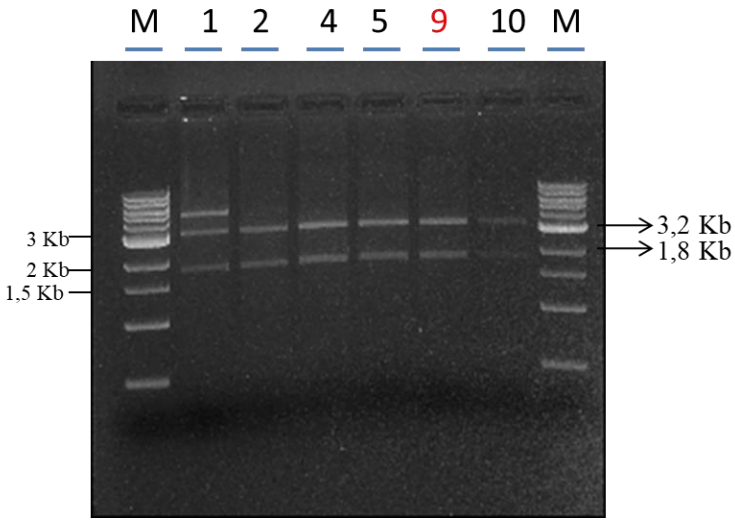


Fig. 26: Enzymatic digestion with *BstBI-NotI*, to evaluate the potential of pPICZαAHEV-CSP plasmid. Marker = 1Kb DNA ladder. The selected clone is # 9

For these reasons we decided to select clone # 9, which was amplified by plasmid maxi-preparation (materials and methods, p 5.6) and then verified by analytical digestion with BstBI-NotI (Fig.27). The sequence was confirmed for automatic sequencing at MWG, resulting 100% homologous to a presumed sequence.

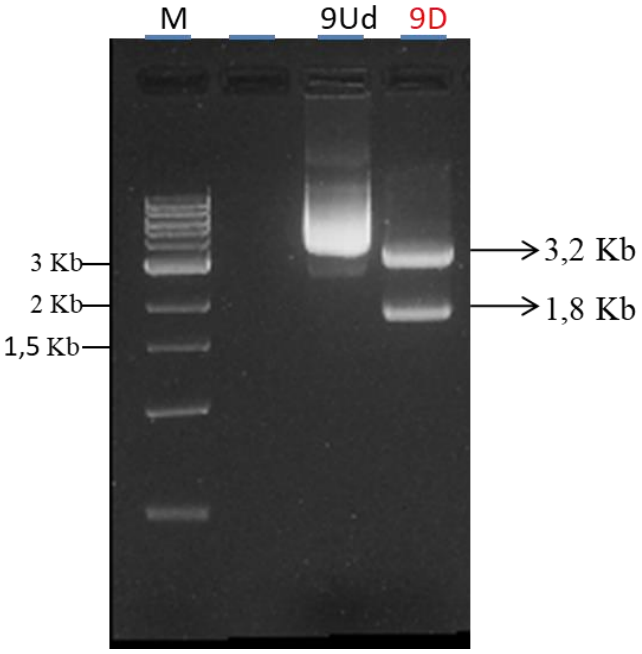


Fig. 27: Maxi preparation of pPICZαAHEV-CSP # 9 plasmid DNA and related digestion with *BstBI-NotI* (Digested = D) and (Undigested = Ud). Marker = 1Kb DNA ladder

6.2 Plasmid Linearization

To facilitate entry and integration of the construct into host cells, pPICZαAHEV-CSP # 9 construct was linearized with *Sac I* restriction enzyme and then maintained in gene sequences as a unique restriction site. An empty pPICZαA vector was also linearized with the same enzyme (Fig. 28). The linearized plasmid DNA was then extracted and purified using the QIAquick® spin Handbook extraction kit (QIAGEN, Milan, Italy).

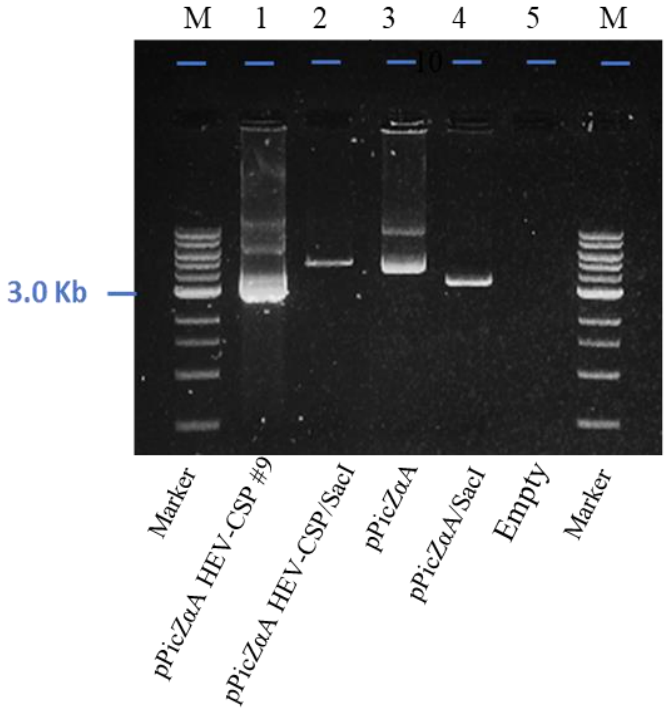


Fig. 28: Linearization of plasmid DNA with *SacI*; PICZαAHEV-CSP and the empty PICZαA plasmid. Marker: 1Kb DNA Ladder

6.3 Transfection of *Pichia Pastoris* KM71 by electroporation

Transfection allows introduction of exogenous cellular in eukaryotic biological material. In this case, this includes cells belonging to *Pichia pastoris* strain KM71. This is a methyltrophic yeast widely used for biotechnological production of human proteins due to the advantages it offers, including: rapid growth rate that allows in a short time to obtain a high cell density; presence of a strong and inducible promoter; expression of heterologous proteins with an almost always correct tertiary structure (Kurtzman CP, 2009).

The first step for HEV-CSP target protein expression is precisely the transfection which allows to obtain a stable transformation of the host genome, integrating then an exogenous sequence in a well-defined gene locus, which in our case corresponds to the locus aox1, placing it under control of its promoter inducible by the presence of methanol. Transfection takes place by electroporation allowing entry of exogenous DNA through use of an electric field, which acts by increasing diameter of the host cell's pores.

In detail, after having made *Pichia pastoris* KM71 cells competent for transformation (Cregg JM et Russel KA, 1998), transfection was carried out with the aid of a Gene Pulser XCell (Bio-Rad, Milan, Italy) electroporator with the following voltage parameters 1,500 V, 25 μF of capacitance and 200 Ω of resistors and cuvettes for electroporation. Immediately after electroporation, cells were added with 1 ml of sorbitol and incubated at 28-30 °C without shaking for 1-2 hours. Once transfection took place, plates were incubated for 2-3 days in YPD (1% yeast extract 2% peptone 2% dextrose, 2% agar) with an appropriate concentration of Zeocin® (Sigma-Aldrich, Milan, Italy).

6.4 Transformants selection by *Replica Plating*

The selection of transformants was carried out by *replica plating*, using different concentrations of Zeocin® (Sigma-Aldrich, Milan, Italy) (1,000 μg/ml; 2,000 μg/ml; 2,500 μg/ml). This allowed for clone # 5 to be selected which was highly resistant to its antibiotic (2,500 μg/ml). This clone was selected because empirically it contains the largest number of copies of the integrated plasmid and, consequently, will express the recombinant protein of interest in a higher concentration. The induction was carried out in 10% methanol.

6.5 Expression and analysis of HEV-CSP fusion protein after methanol induction

The induction curve via methanol performed for five days with daily sampling allows to establish an optimal time of maximum gene expression of the recombinant proteins and allows to select their over productive clones. The gene expression profile of each recombinant protein was evaluated on the total cellular lysates. Western blot analysis, using mouse anti-HEV monoclonal antibody, found that the maximum expression level of the pPICZαAHEV-CSP protein (66 kDa) was after 5 days (D5) of induction in 10% methanol (Fig. 29).

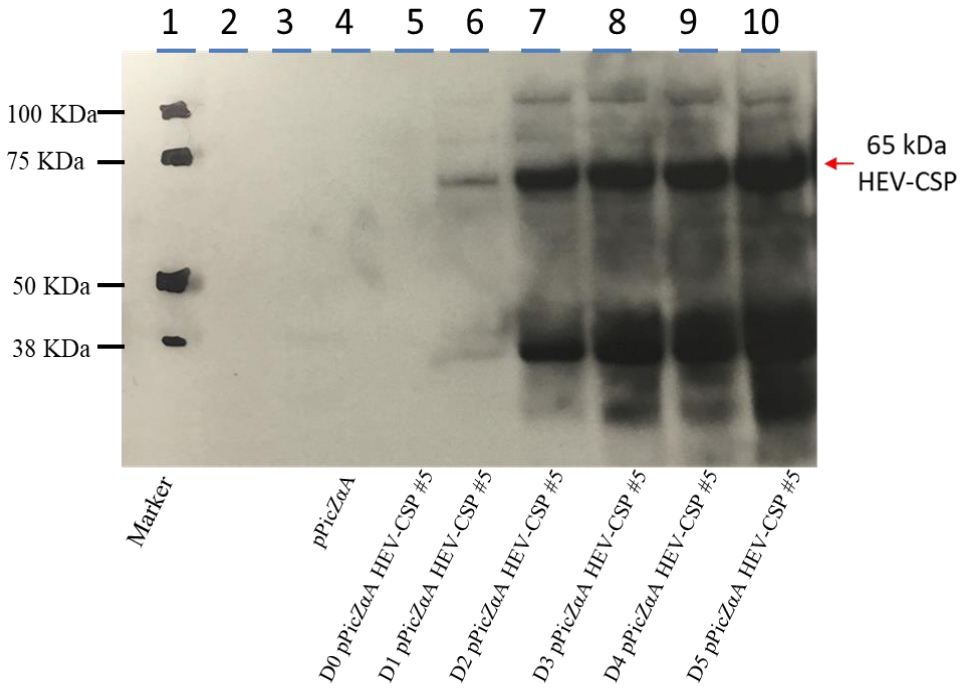


Fig. 29: Recombinant protein expression profile performed on total cell lysates of pPICZαA HEV-CSP from 1 to 5 days of 10% methanol induction. Protein expression levels were compared with those of the control (D0)

6.6 Cloning of recombinant HEV-PfF2 gene into pPICZαA expression vectors

To clone the recombinant HEV-PfF2 gene into a 3.3 Kb pPICZαA expression vector, the synthetic gene, inserted into an intermediate vector (17AEBIBC_HEV-PfF2_ pMA-T), was selected in LB agar plates with Ampicillin, amplified for plasmid maxi-preparation, using the QIAGEN Plasmid Maxi Purification Kit (Milan, Italy) and recovered by digestion with *BstBI* and *NotI*.

The pPICZαA expression vector was then digested with the same restriction enzymes, to obtain also in this case compatible ends for the subsequent cloning steps of the insert. *BstBI-NotI* restriction fragments, obtained from both vector and insert, were purified on agarose gel, using the QIAEX I Gel Extraction kit (QIAGEN, Milan, Italy), and linked by T4 Ligase to obtain a recombinant plasmid known as pPICZα-HEV -PfF2 of 5.0 Kb. This ligation product was used to transform *E. coli* competent cells XL10 GOLD (Stratagene, USA), selecting transformants in LB agar with Zeocin® (Sigma-Aldrich, Milan, Italy).

Selection of transformants was carried out by mini-preparations of plasmid DNA and digestion with restriction enzymes. To assess the presence of recombinant plasmids, 3 colonies were examined for pPICZαAHEV-PfF2: plasmid DNA was extracted from each colony and analysed in 1% agarose gel.

Figure 30 shows that the DNA of each colony, except clone # 15, is represented by a band of height equal to 5 Kb; it coincides with the size of the recombinant taken in examination. Therefore, probably both clones 16 and 18 represent the recombinant plasmid.

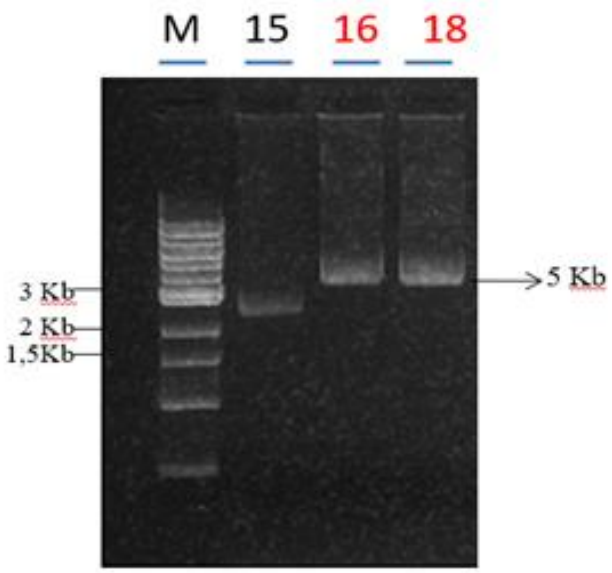


Fig. 30: Screening of transforming bacterial colonies with pPICZαAHEV-PfF2 recombianant DNA by mini plasmid DNA preparation. Marker: 1 Kb DNA ladder. All three clones were selected for subsequent enzymatic digestion with *BstBI* and *NotI*

To evaluate the successful integration of HEV-PfF2 gene into pPICZαA plasmid, enzymatic digestion was performed using *BstBI* and *NotI* cloning enzymes.

Figure 31 highlights a different digestion pattern in these three clones. Specifically, clone # 15 indicates that digestion did not take place because of showing only one band; thus, it also indicated the lack of one of the restriction sites. This incorrect clone was eventually discarded.

Clone # 16 shows a three-band digestion pattern, two with expected sizes (3.2 Kb corresponding to the vector and 1.8 Kb, corresponding to the insert), and one situated high up (about 5 Kb), corresponding to undigested plasmid DNA. For this reason, clone # 16 was

also eliminated due to partial or incomplete digestion. Clone # 18 instead presents two bands with expected sizes: one evidently relating to pPICZαA vector of 3.2Kb and the other, not markedly visible in the figure below, of 1.8 Kb corresponding to the HEV-PfF2 insert.

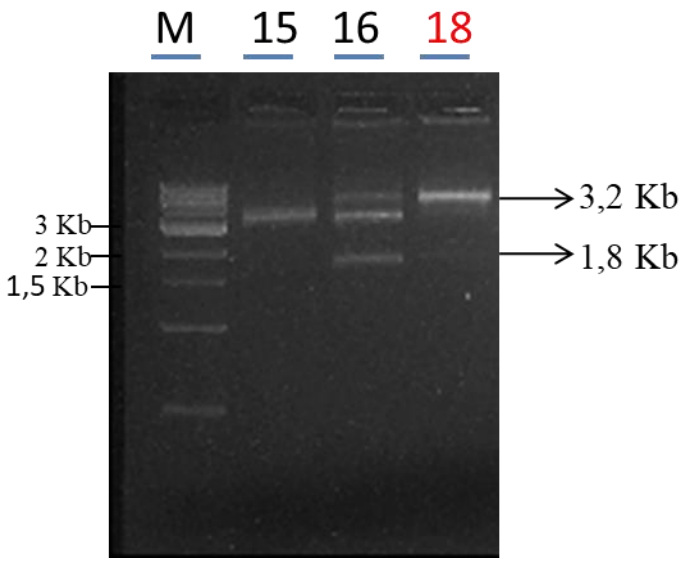


Fig. 31: Enzymatic digestion with *BstBI-NotI*, to evaluate the positivity of plasmid pPICZαAHEV-PfF2. Marker = 1Kb DNA ladder. The selected clone was # 18

For these reasons we decided to select clone # 18, which was then amplified by plasmid maxi-preparation, verified by analytical digestion with *BstBI-NotI* (Fig. 32). Its sequence was confirmed by automatic sequencing at MWG, verifying 100% of homology with this assumed sequence.

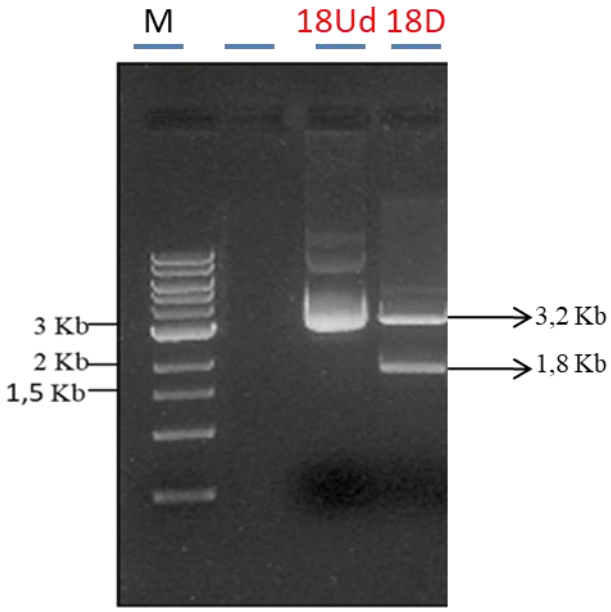


Fig. 32: Maxi plasmid DNA preparation pPICZαAHEV-PfF2 # 18 and its related digestion with *BstBI-NotI* (Digested =D) and (Undigested= Ud). Marker= 1Kb DNA ladder

6.7 Plasmid Linearization

To facilitate entry and integration of pPICZαAHEV-PfF2 # 18 construct into a host cell, the construct was linearized with the *Sac I* enzyme, which was specially maintained in the gene sequences as a unique restriction site. Linearized plasmid DNA was then extracted and purified using the QIAquick® spin Handbook extraction kit (QIAGEN, Milan, Italy) (Fig. 33). Similarly, empty pPICZαA plasmid was also linearized with *SacI*.

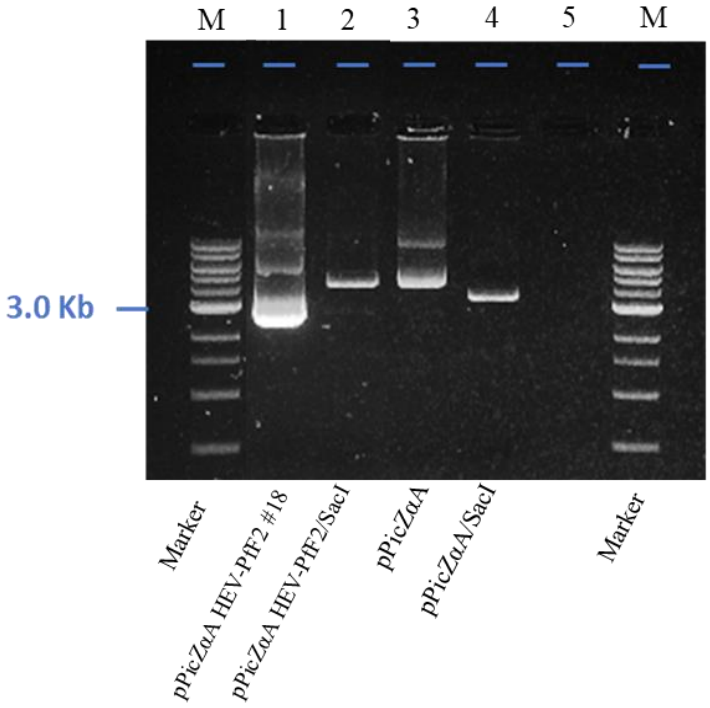


Fig. 33: Linearization with DNA plasmid DNA *SacI*, PICZαAHEV-PfF2# 18 and an empty plasmid PICZαA. Marker: 1Kb DNA Ladder

6.8 Transfection of *Pichia pastoris* KM71 by electroporation

After making *Pichia pastoris* KM71 cells competent for transformation (Cregg JM et Russel KA, 1998), transfection was carried out with the aid of an electroporator (Bio-Rad, Milan, Italy). Once transfection took place, the plates were incubated for 2-3 days in YPD (1% yeast extract 2% peptone 2% dextrose, 2% agar) with a suitable concentration of Zeocin® (Sigma-Aldrich, Milan, Italy).

6.9 Transformants selection by Replica Plating

The selection of transformants was carried out through using replica plating technique and different concentrations of Zeocin® (Sigma-Aldrich, Milan, Italy) (1,000 µg/ml; 2,000 µ/ml; 2,500 µg/ml). This allowed to select clones # 6 and # 7, which were highly resistant to antibiotic (2,500 µg/ml). These clones are those who contain the greatest number of copies of each plasmid and, consequently, will express high concentration of the recombinant proteins. Methanol induction was then carried out.

6.10 Expression and analysis of HEV-PfF2 fusion protein after methanol induction

Western blot analysis, using mouse anti-HEV monoclonal antibody, revealed that the maximum expression of the pPICZαA-HEV-PfF2 protein for both clones # 6 and # 7 (67 kDa) was after 2 days (D2) (Fig. 34) with 10% methanol induction; clone # 7 instead presented a more marked band, probably because expression vector *Pichia pastoris* KM71 had integrated a greater number of copies of recombinant plasmid.

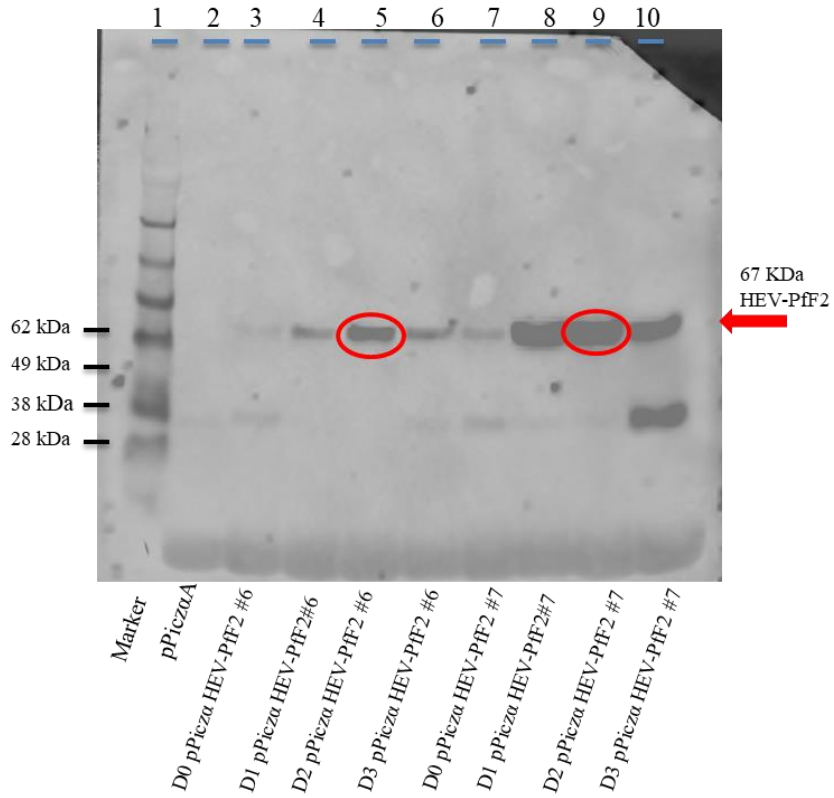


Fig. 34: Recombinant protein expression profile performed on total cell lysates of pPICZαA HEV-PfF2 from 1 to 5 days of induction in 10% methanol. Protein expression levels were evaluated during 5 days of induction in methanol and compared to day 0 where there is a slight endogenous expression of the recombinant protein

6.11 Cloning of recombinant HEV-MSPFu24 gene into pPICZαA expression vector

To clone recombinant HEV-MSPFu24 gene into 3.3 Kb pPICZαA expression vector, a synthetic gene, inserted into the intermediate vector (17AEBIBC_HEV-MSPFu24_pMA-T), was selected in LB agar plates with Ampicillin, amplified for plasmid maxi-preparation, using QIAGEN Plasmid Maxi Purification Kit (QIAGEN, Milan, Italy) and recovered by digestion with *BstBI* and *NotI*.

The pPICZαA expression vector was digested with the same restriction enzymes, to obtain also in this case compatible ends for those subsequent cloning steps of the insert. BstBI-NotI restriction fragments, acquired by both vector and insert, were purified on agarose gel, using QIAEX I Gel Extraction kit (QIAGEN, Milan, Italy). They were then linked by T4 Ligase to obtain recombinant plasmid pPICZαAHEV- MSPFu24 (4.4 Kb).

This ligation product was used to transform *E. Coli* competent cells XL10 GOLD, selecting transformants in LB agar with Zeocin® (Sigma-Aldrich, Milan, Italy). Selection of the transformants was carried out by mini-preparations of plasmid DNA and digestion with restriction enzymes. To evaluate presence of the recombinant plasmids, 8 colonies for pPICZαAHEV-MSPFu24 were examined via mini plasmid DNA preparations.

The screening of bacterial colonies, as shown in figure 35, demonstrated different plasmid DNA sizes, some closer to the desired size (clones # 5, # 6, # 7, # 10 and # 11 of about 4 Kb) and others completely wrong (clones # 8, # 9 and # 12 of sizes over 6 Kb). Clones # 5, #6 and #7 were selected for subsequent restriction analyses, to select a definitive capable of being expanded and used to transfect KM71 yeast cells (Invitrogen, Life Technologies, Milan, Italy).

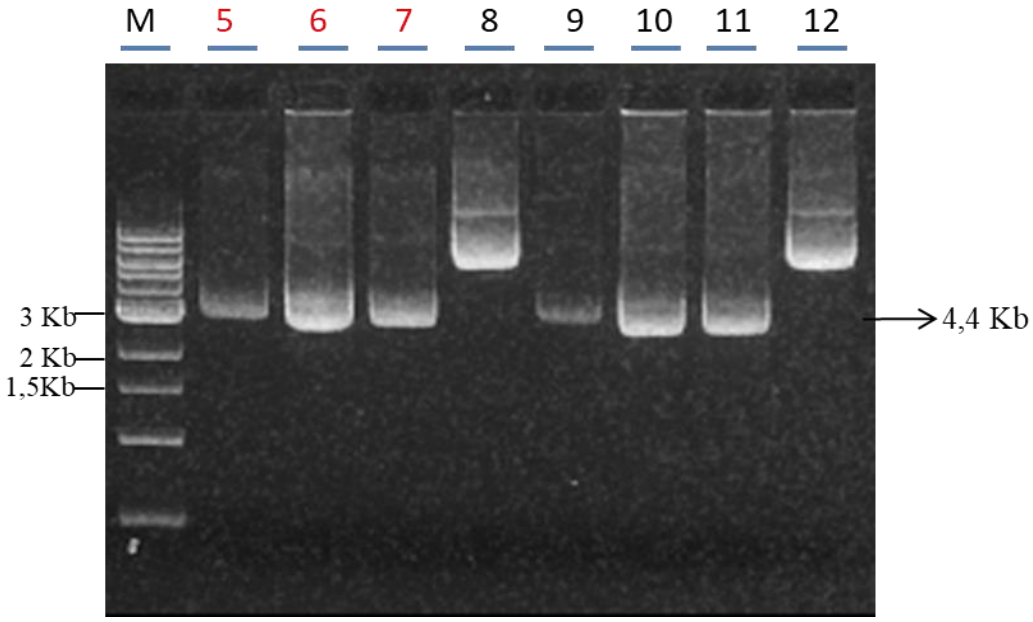


Fig. 35: Screening of transforming bacterial colonies with pPICZαAHEV-MSPFu24 recombinant DNA via mini plasmid DNA preparation. The height of expected DNA band is about 4.4 Kb. To confirm the recombinant clone accuracy, clones 5, 6 and 7, highlighted in red, were selected for subsequent enzymatic digestion with *BstBI-NotI* cloning enzymes. Marker: 1 Kb DNA ladder

Then, to evaluate the exact integration of recombinant HEV-MSPFu24 gene into pPICZaA expression vector, am enzymatic digestion of the three clones was carried out with restriction enzymes used in the previous cloning, namely *BstBI* and *NotI*.

Figure 36 highlighted that all clones were positive as there were two distinct and separate bands, one relating to vector pPICZαA (3.2 Kb) and the other to HEV-MSPFu24 insert (1.2 Kb). Clone # 6 still shows some undigested DNA, as evidenced by the band above 4 Kb.

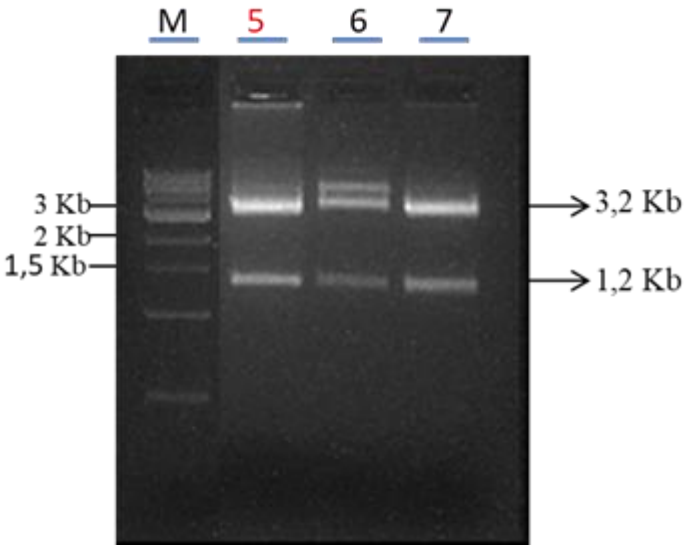


Fig. 36: Enzymatic digestion with *BstBI-NotI*, to evaluate potivity of plasmid pPICZαAHEV-MSPFu24. Marker = 1Kb DNA ladder. The selected clone is # 5

We then selected clone # 5, which was amplified for plasmid maxi-preparation, verified by analytical digestion with *BstBI-NotI* (Fig. 37) and the sequence was confirmed by automatic sequencing at MWG, showing 100% homology with the presumed sequence.

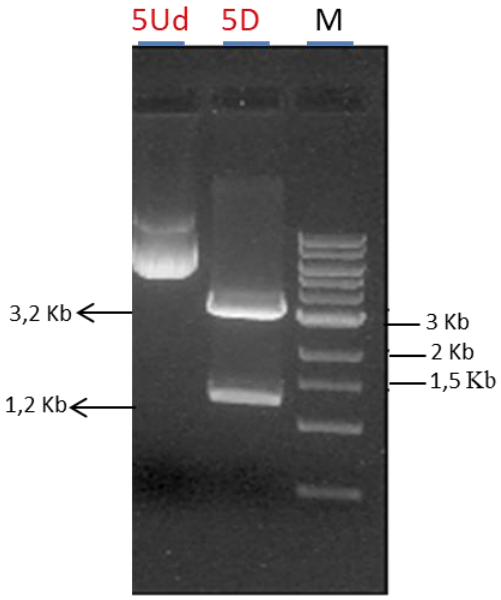


Fig. 37: Maxi plasmid DNA preparation pPICZαAHEV-MSPFu24 # 5 and its related digestion with *BstBI-NotI* (Undigested= Ud). Marker= 1Kb DNA ladder

6.12 Linearization of the plasmid

Recombinant plasmid pPICZαAHEV-MSPFu24 # 5, together with an empty pPICZαA plasmid, was linearized with *Sac I* enzyme, specially maintained in its gene sequences as a unique site of restriction, to facilitate entry and integration of a construct into its host cell. The linearized plasmid DNA was then extracted and purified using the QIAquick® spin Handbook extraction kit (QIAGEN, Milan, Italy) (Fig. 38).

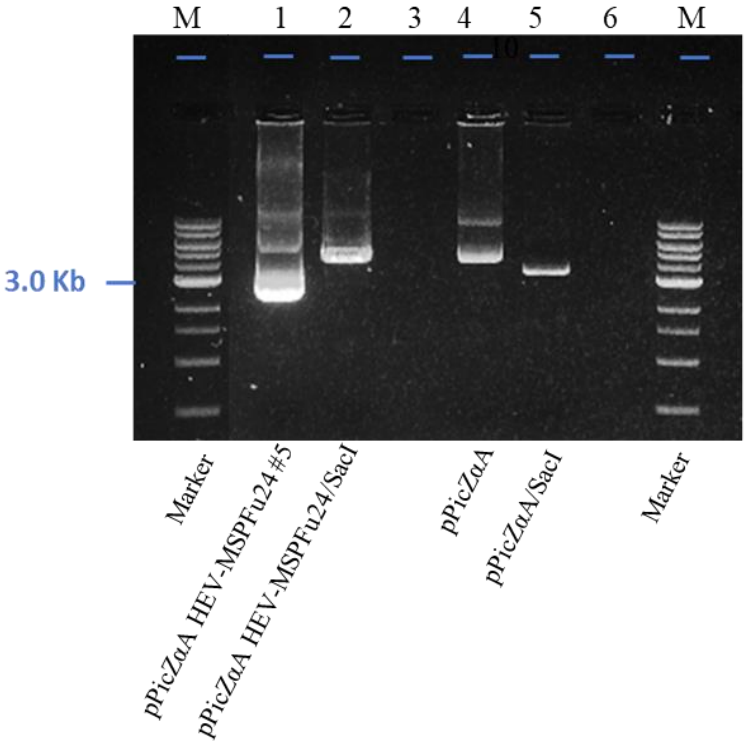


Fig. 38: Linearization with plasmid DNA *SacI;* PICZαAHEV-MSPFu24 and empty plasmid PICZαA. Marker: 1Kb DNA Ladder

6.13 Transfection of Pichia Pastoris KM71 by electroporation

After making *Pichia pastoris* KM71 cells competent for transformation (Cregg JM et Russel KA, 1998), transfection was carried out with the aid of an electroporator. Once transfection took place, the plates were incubated for 2-3 days in YPD (1% yeast extract 2% peptone 2% dextrose, 2% agar) with appropriate concentration of Zeocin® (Sigma-Aldrich, Milan, Italy).

6.14 Transformants selection by Replica Plating

The selection of transformants was conducted by replication plating, using different concentrations of Zeocina® (1,000 µg/ml; 2,000 µg/ml; 2,500 µg/ml). This allowed clone # 3, which was highly resistant to antibiotic (2,500 µg/ml), to be selected. This clone contained the largest number of copies of plasmid and, consequently, will express the recombinant protein of interest in higher concentrations. The induction was performed in 10% methanol.

6.15 Expression and analysis of HEV-MSPFu24 fusion protein after methanol induction

Western blot analysis revealed that the maximum expression level of the pPICZαAHEV-MSPFu24 protein (44 kDa) occurred after 4 days (D4) (Fig. 39).

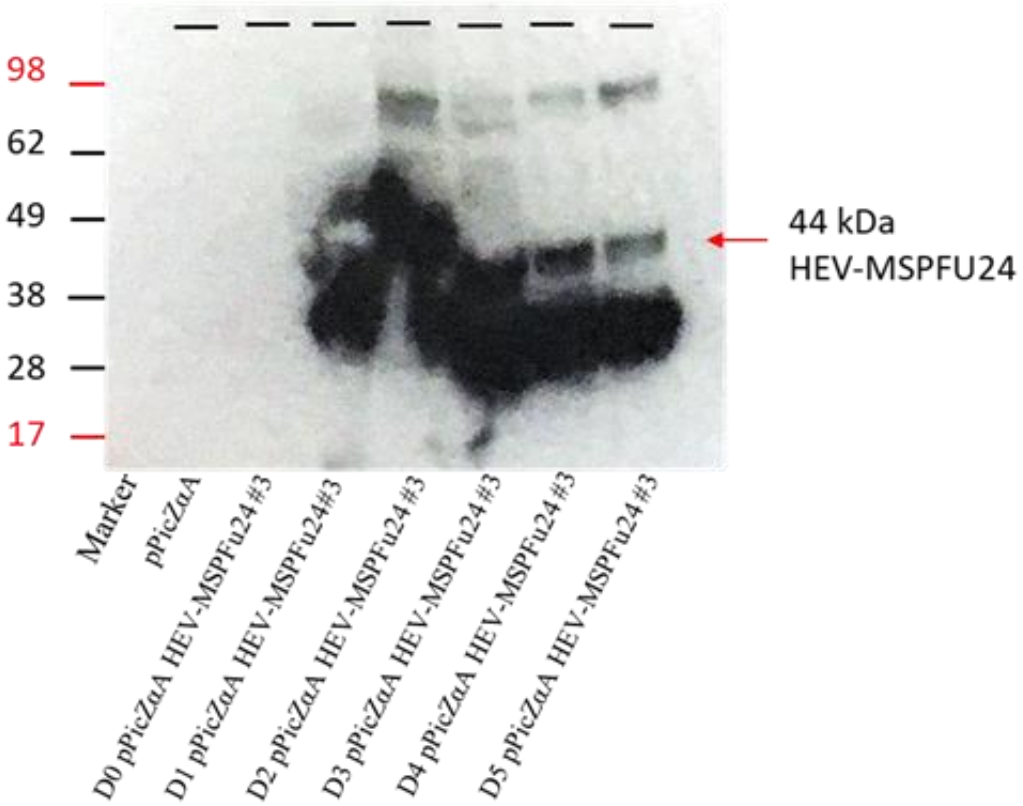


Fig. 39: Recombinant protein expression profile conducted on total cell lysates of pPICZαA HEV-MSPFU24 from 1 to 5 days of methanol induction. Protein expression levels were compared with those of the control (D0)

6.16 Homogenization, Filtration and Ultrafiltration

Once established the day (D) of maximum expression of each single construct, HEV-CSP (D5), HEV-PfF2 (D2) and HEV-MSPFu24 (D4), all cultures were homogenized using the LM20 homogenizer. In particular, 100 ml of each sample in Lysis Buffer (50 mM tris-HCl, 150 mM NaCl, 0.5% Triton, 10 mM EDTA, pH 7.8) and corresponding to an OD (λ=600 nm) of about λ=300 nm was subjected to three cycles of homogenization at high pressure (28,000 Psi) obtaining a uniform reduction in particle size and a halved OD. The samples were centrifuged at 10,000 x g for 30 minutes and the supernatants thus obtained were, in a

first step, filtered through a 22 nm filter and subsequently ultrafiltered and concentrated through AMICON- 2- CENTRIFUGAL FILTER DEVICES® (30K) with a cut -off 2 times smaller than the molecular weight of our proteins, respectively about 66 kDA for HEV-CSP, 67 kDa for HEV-PfF2 and 44 kDa for HEV-MSPFu24.

6.17 Determination of the PI of each recombinant protein

In order to start purification of the proteins of interest, a study was carried out on the respective amino acid sequences of these proteins of interest, using the ExPASY IT resource portal. Molecular weight and isoelectric point (pI) of each recombinant protein were thus determined (Fig. 40).

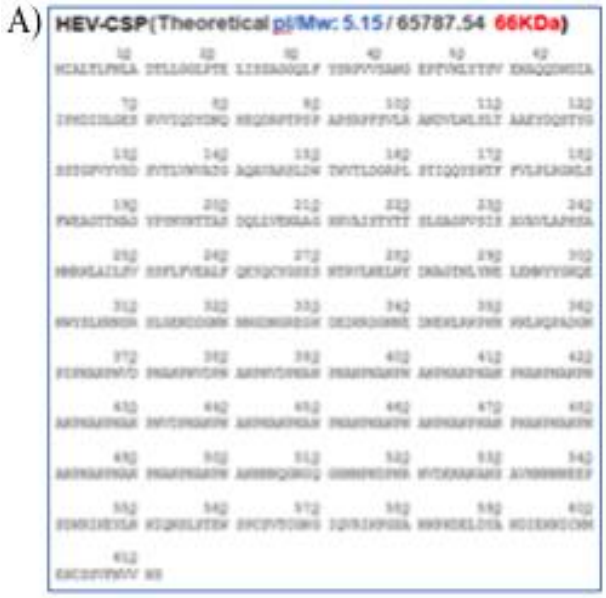


Fig. 40: Amino acid sequence, molecular weight and isoelectric point (pI) of the recombinant proteins: A) HEV-CSP; B) HEV-PfF2; C) HEV-MSPFu24

Purification procedures were performed on the recombinant HEV-CSP protein o adjust optimal experimental conditions and to avoid waste at the same time. For the purification, saline gradients with increasing concentration were used by varying both pH of the sample and that of the Buffer, taking into account pI of the protein in question, equal to 5.15; first

purification step was started with using an anion exchange chromatography (Buffer A: or Binding Buffer 20 mM Tris-HCl) and Buffer B: o Elution Buffer 20 mM Tris-HCl, 1 M NaCl with a pH equal to 6.15 one point higher than its pI, following a gradient linear elution.

6.18 Preliminary phase of purification of HEV-CSP protein on anion exchange column (Hi Trap Q FF) through the AKTA system (GE)

After being diluted in a suitable mobile phase (Buffer A: or Binding Buffer 20 mM Tris-HCl), a crude protein was filtered and loaded into the column at a flow of 1 ml per min. During the process of loading this protein into the column, the "Flow Through", corresponding to any part of protein not bound to a resin, was collected and stored for subsequent analysis.

Furthermore, a fraction corresponding to the Buffer used to wash the “Washing” column was also collected to highlight, in a second analysis, presence of a protein of interest initially bound to the resin but later detached at very low concentrations of washing Buffer.

Finally, the sample was eluted using Buffer B: or Elution Buffer 20 mM Tris-HCl, 1 M NaCl according to a linear gradient, collecting the eluate obtained in 30 fractions of 4 ml each, previously placed in an automatic fraction collector; they were subsequently analysed via Western Blot analysis.

From the chromatogram analysis performed using the UNICORN software, 3 peaks were highlighted: the first well resolved (green box) which includes fractions from 11 to 18. The peak began to elute at a concentration of Buffer B 23.83% and at a conductivity of 5.71 mS/cm and resolved in fraction 18 at an Elution Buffer concentration equal to 42.40%.; two other peaks, partially overlapping but eluted at a very high salt concentration (Fig. 41).

Assuming that our protein of interest could be eluted in the first peak, we analysed by Western blot analysis the fractions 11 to 18 (lane 5-12). Furthermore, to obtain a complete picture on the first purification step, we evaluated the samples corresponding to “Flow Through” (lane 3) and to Washing (lane 4).

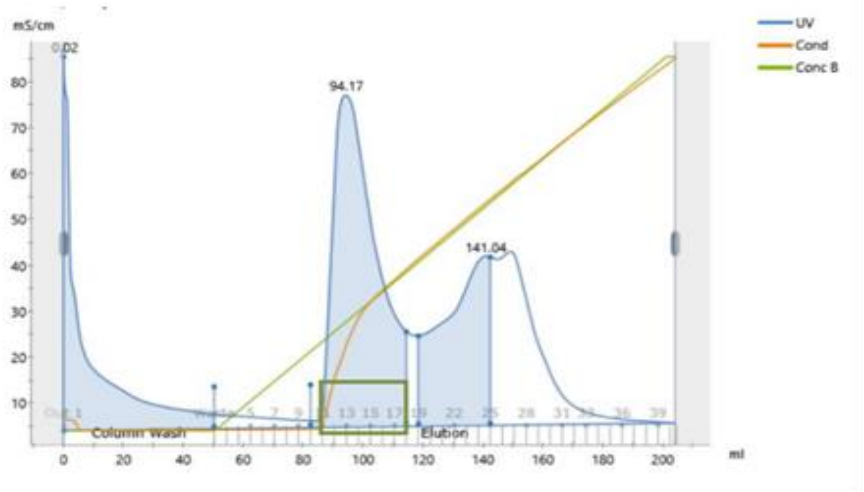


Fig. 41: Chromatogram obtained after purification of the HEV-CSP protein on anion exchange resin, according to linear gradient at pH 6.15

The expression levels of the recombinant HEV-CSP protein were determined using a mouse monoclonal antibody against HEV. As can be seen from Figure 42, the HEV-CSP protein having a MW of about 66 kDa, highlighted as a single band by the specific anti HEV antibody, was not expressed in fractions 11 to 18; instead, it was in those of the "flow through" and of the "washing".

This meant that neither the chosen resin, nor the loading Buffer and the elution Buffer used at pH 6.1 had allowed a binding of HEV-CSP protein to the resin and its subsequent release at a well-determined salt concentration. Therefore, our protein only crossed the column

finding itself into the flow through fractions and in the washing; this led to the conclusion that the chosen conditions were not suitable.

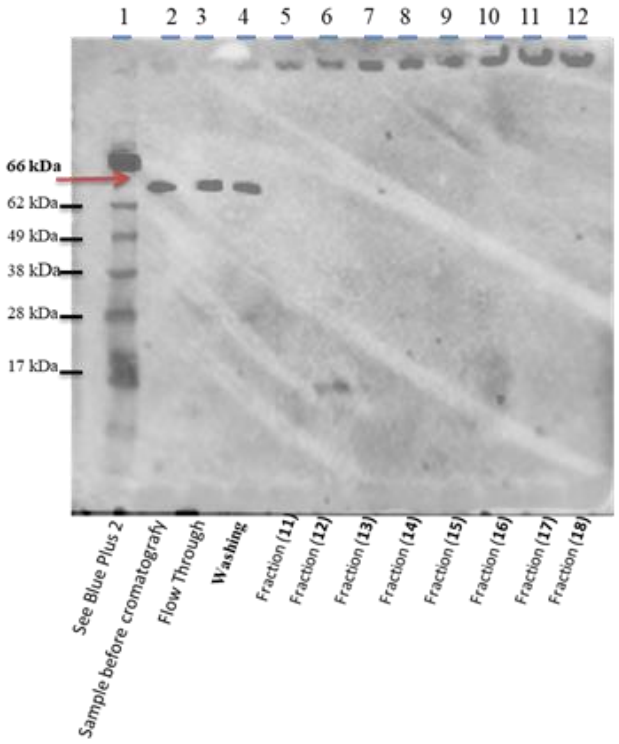


Fig. 42: Immunoblot of the fractions eluted from the anion exchange column through a linear gradient at pH 6.15 Marker: M See blue plus 2. Lane 2: HEV-CSP sample. Lane 3: flow through. Lane 4: washing. Lane 5-12: eluted fractions

6.19 Preliminary phase of purification of the HEV-CSP protein on a cation exchange column (Hi Trap SP FF) through the AKTA system (GE)

The completely negative result obtained from anion exchange chromatography led us to focus attention on cation exchange chromatography, varying the pH by one point towards acidity, via selecting washing and elution Buffers that work at an optimal pH of 4.15. A "HiTrap SP FF" strong cation exchange column was therefore chosen using Buffer A or

Binding Buffer (50 mM Formic Acid, pH 4.15) and Buffer B: or Elution Buffer (50 mM Formic Acid, 1 M NaCl). The pH of our sample was first acidified until reaching the elution pH (4.15), then this sample was suitably filtered and loaded into the column at the rate of 1 ml/min. The cation exchange chromatography was then started according to a linear elution gradient, following a flow of 1 ml/min and collecting 4 ml fractions, each in an automatic fraction collector.

From this conducted chromatogram analysis using the UNICORN software, 1 peak (green box) was highlighted including fractions 8 to 14 (Fig. 43), where our protein could potentially have been eluted. The elution of this peak began at a Buffer B concentration of 47.18% to complete in fraction 14 at a concentration of 95.61% of Buffer B.

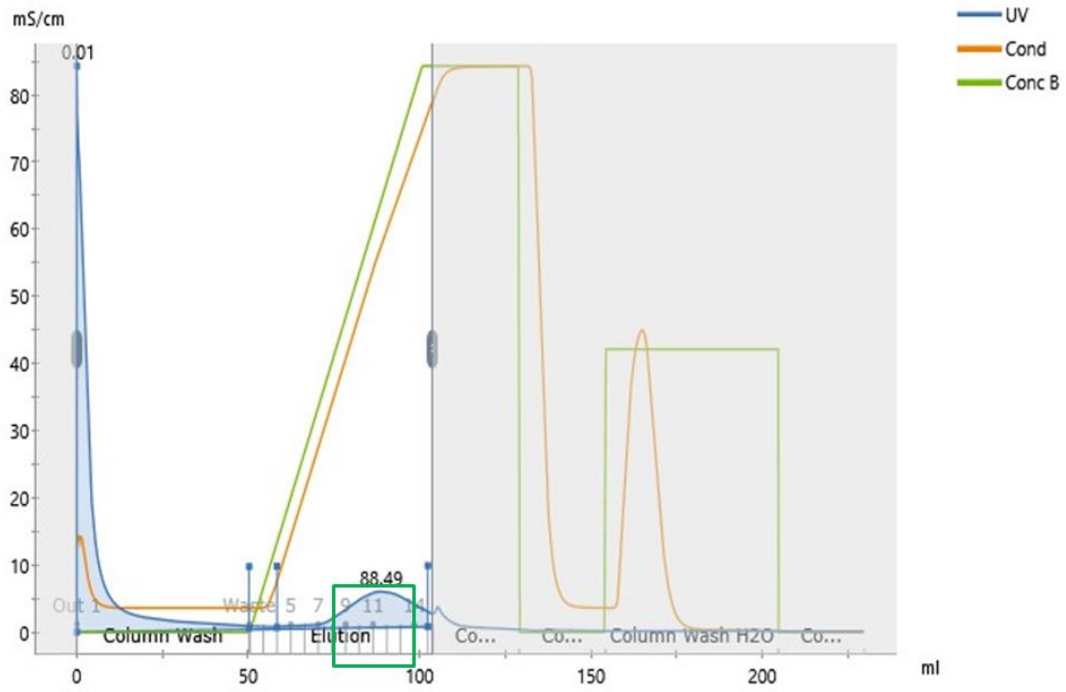


Fig. 43: Chromatogram obtained after purification of HEV-CSP protein on strong cation exchange resin, according to linear gradient at pH 4.15

Fractions from 8 to 13 (Lane 7-12), the one corresponding to "flow through" (Lane 2) and to "washing" (Lane 3) (Fig. 44) were then analysed by Western Blot analysis; using the anti HEV monoclonal antibody as detector of the protein of interest.

As shown in figure 44, HEV-CSP protein of 66 kDa at acid pH (4.1) before being loaded into the column, had the expected band of about 66 kDa, but also two smaller band: one at 14 kDa, which corresponds to a portion of HEV and one at 28 kDa, probably a HEV dimer. We find the whole protein (66 kDa) in flow through, while from fraction 9 to 13 there is an increase in the band at 14 kDa, which reaches its maximum in fraction 12. Probably, an acid pH has favoured proteolysis of our HEV protein CSP right at the melting point of HEV with CSP, in large extent of the loaded sample. The whole protein was not retained by the column and we find it all again in the flow through (lane 2).

Also in this case, the chosen cation exchange resin, the acid pH, the selected Buffers and the elution conditions, according to linear gradient, were not suitable for a purification of the protein under examination.

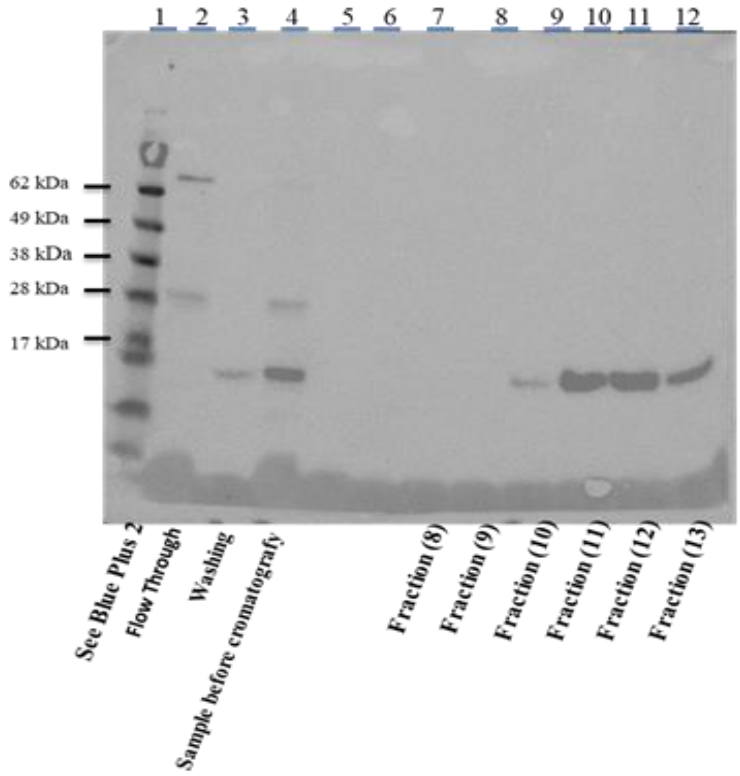


Fig. 44: Immunoblot of the fractions eluted from the cation exchange column through a linear gradient at pH 4.15 Marker: M See blue plus 2. Lane 4: HEV-CSP sample. Lane 2: flow through. Lane 3: washing. Lane 7-12: eluted fractions

6.20 HEV-CSP protein purification on strong anion exchange column (Hi Trap Q FF through the AKTA system (GE)

The results of the previous cation exchange chromatography showed an undesired proteolytic fragmentation of the protein under conditions of acid pH, an event which had not previously occurred under conditions of near neutrality (during the first anion exchange chromatography at pH 6.15). It was therefore decided to return to strong anion exchange chromatography using "Hi Trap Q FF" column (GE Healthcare, Naples, Italy), moving however the elution pH values by another point towards alkalinity: pH 7.0

After adjusting conductivity of the protein to 3.3 milli Siemens (mS) through a series of dilutions in Buffer A (or Binding Buffer, 50 mM Tris-HCl, pH 7.0) (the column needs a maximum conductivity of 5 mS) we proceeded to application of the sample and its subsequent elution, according to a linear purification gradient at a constant flow of 5 ml per minute.

We started with balancing the column to be used by setting the following program called "5 ml QFF PRIME" in Akta Pure system (GE Healthcare Europe GmbH, Life Sciences, Milan, Italy); it consisted in 4 washes:

1. Washing with H_2O

2. Washing with 5 column volumes (cv). of Buffer A (50 mM Tris HCl, pH 7.0)

3. Washing with 5 cv of Buffer B (50 mM Tris-HCl, 1 M NaCl, pH 7.0)

4. Washing with 10 hp Buffer A

Then, the sample of interest was automatically applied through injector A and the "flow through" was collected; a fraction which was subsequently analysed to evaluate any part of the sample not bound to the resin. This column was then washed, with a process known as "washing", conveyed into a collection bottle and stored at 4 °C to be subsequently tested.

The chromatographic process ended with elution of the sample in Buffer B (or Elution Buffer) according to a linear gradient, collecting eluate in 30 tubes, previously placed in an automatic fraction collector. The final graph allows you to view peaks that include fractions in which the target protein is probably present and therefore to see which fractions it necessary to be analysed (Fig. 45).

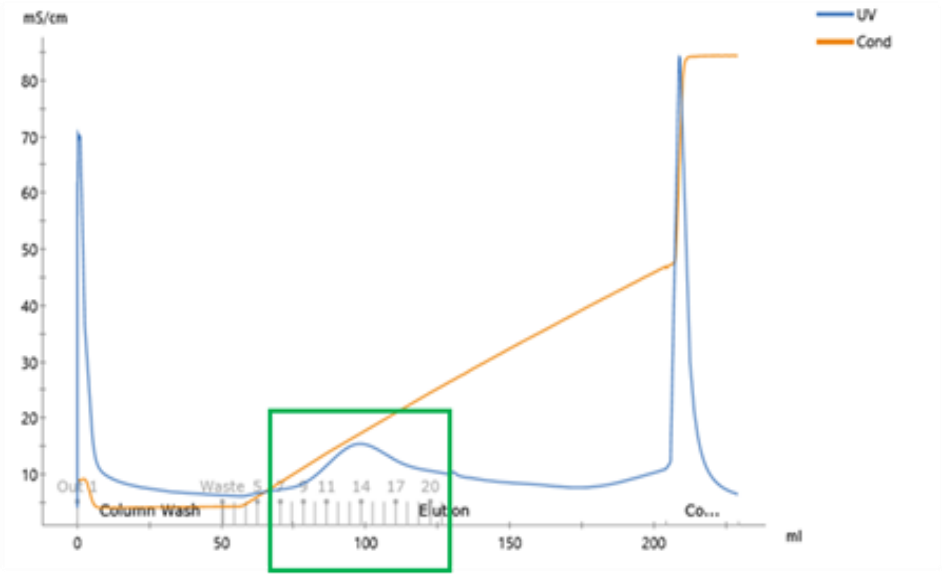


Fig. 45: Chromatogram depicting the first purification step of the HEV-CSP protein on anion exchange resin, second linear gradient in Elution Buffer at pH 7.0

The chromatogram in Figure 45 shows a single peak (in the green box) where all the eluted fractions (from # 4 to # 22) are concentrated according to a linear gradient. These fractions were then analysed by SDS-PAGE and Western blot analysis (Fig. 46) to identify those corresponding to HEV-CSP protein.

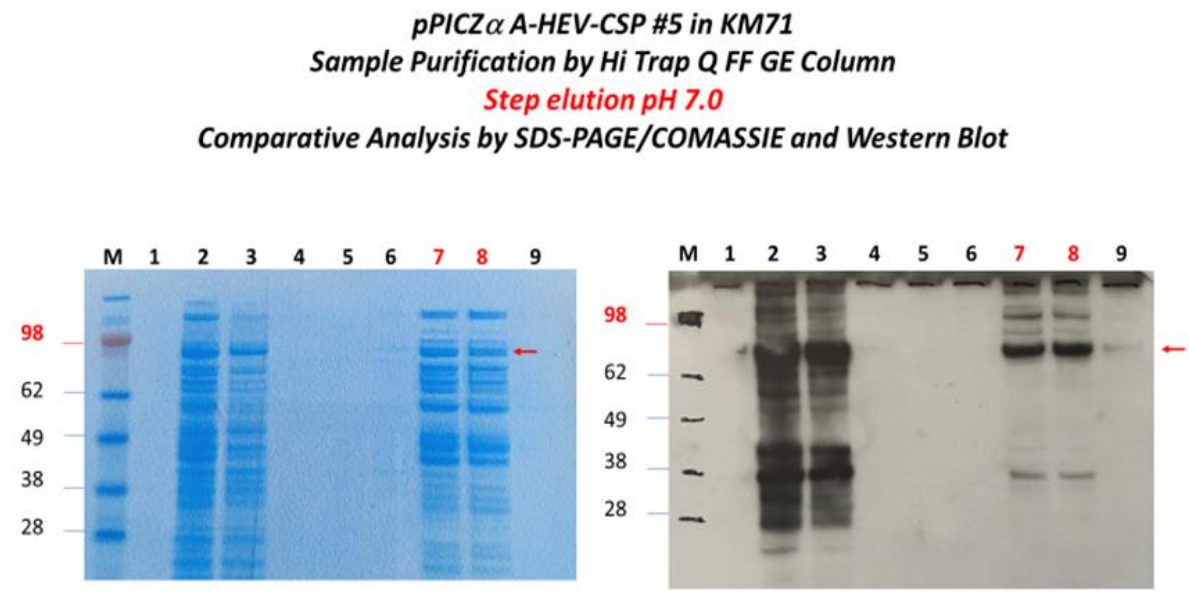


Fig. 46: SDS-PAGE and Western blot Analysis of the fractions eluted from the anion exchange column through a linear gradient at pH 7.0. Marker: M See blue plus 2. Lane 2: HEV-CSP sample. Lane 3: flow through. Lane 4: washing. Lane 5-9: eluted fractions

It is evident that at pH 7.0 the binding of our sample to resin is not optimal, since we find it abundantly in the flow through (Lane 3); this indicates that the chosen chromatographic conditions are not optimal. We find our HEV-CSP protein in fractions 7 and 8, as can be seen from hybridization with the specific antibody against HEV used in the Western blot analysis. Furthermore, many non-specific proteins were eliminated with this purification, as evidenced by comparing lines 2 and 3 with lines 7 and 8 of the Western blot analysis.

To try to separate and purify our target protein and obtain a better resolution, we decided to repeat the same chromatography, but this time using a step gradient designed on the basis of the conductivity of fractions 7 and 8, corresponding to specific percentages of elution Buffer. In detail, 4 elution steps have been established:

1. Corresponding to 6.5% of Buffer B

2. Corresponding to 12.2% of Buffer B

3. Corresponding to 15.5% of Buffer B

4. Corresponding to 100% of Buffer B

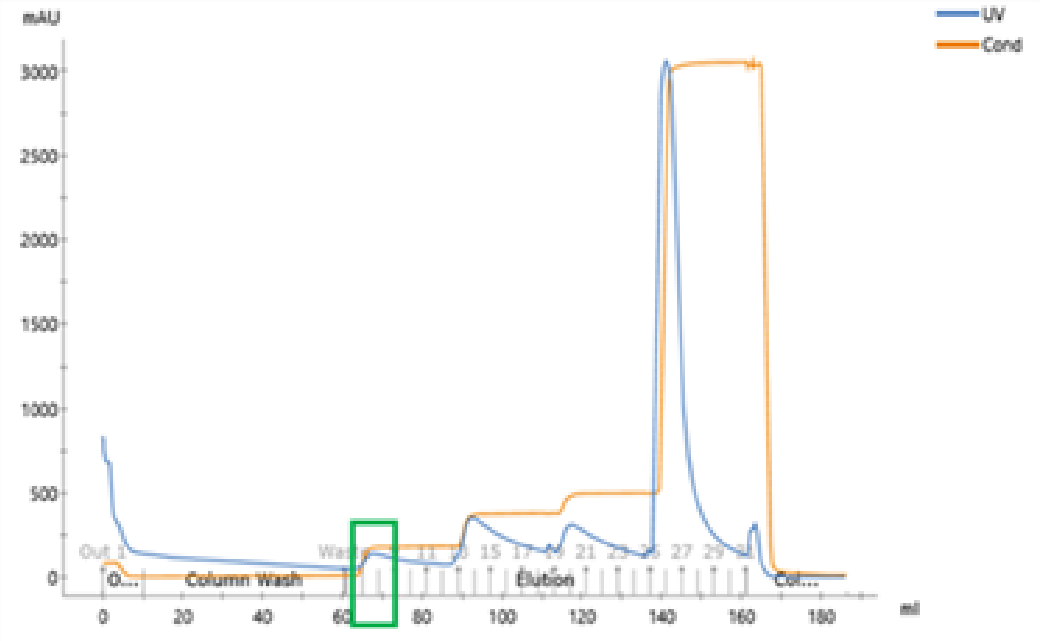


Fig. 47: Chromatogram showing the second purification phase of the HEV-CSP protein on anion exchange resin, according to 4-step gradient corresponding to different percentages of Elution Buffer at pH 7.0

As can be seen from the chromatogram in Figure 47, the peak obtained in linear gradient separation (Fig. 45, green box) resolved into three small but better-defined peaks, with different conductivity and corresponding to various percentages of Elution Buffer used.

The fourth peak, corresponding to elution with 100% of Buffer B and also present in the linear elution chromatogram, represents samples with negative charge, most likely DNA. Our protein is found in fractions 7 and 8 (green box), eluted and better separated at 6.5% of Buffer B. These two fractions were then unified and analysed by SDS-PAGE and Western blot analysis (Fig. 48).

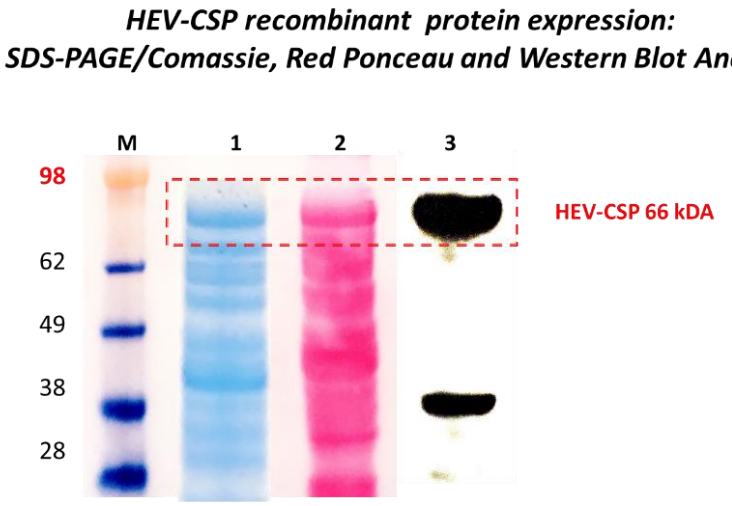


Fig. 48: (1) SDS-PAGE/Comassie, **(2)** Red Ponceau staining and **(3)** Western Blot analysis of fractions 7 and 8 eluted from the anion exchange column through a step gradient at pH 7.0 and combined. Marker: See blue plus 2. The mouse monoclonal antibody used against HEV is LS-667675 LsBio diluted 1:4,000

With this further chromatographic analysis, a better purification of the recombinant HEV-CSP protein was obtained: a non-specific binding of heterologous proteins with the anti HEV antibody is indeed no longer observed, but other steps will be necessary to achieve a purification equal to about 90 %, such as to evaluate the efficacy of this recombinant protein in pre-clinical studies in animal model.

CHAPTER VII

Conclusion

Since malaria presents itself as one of the main parasitic pathologies that affect nowadays different areas of this world, there have been several attempts to produce antimalarial vaccines. An innovative approach consists in the use of recombinant proteins expressed in yeast cell cultures. Yeast cultures could be a valid alternative model to classical expression systems, because they offer the possibility to obtain high levels of expression proteins. In addition, they present a considerable ease whereby production levels can be increased or decreased.

Design studies conducted *in silico* made it possible to appropriately choose antigens expressed in different phases of *Plasmodium falciparum* life cycle, since ultimate goal of this work was the development of a multivalent/multiphasic vaccine candidate against Malaria. These antigens, specifically CSP, PfF2 and MSPFu24, were fused at their 5 'ends with a highly immunogenic part of Hepatitis E capsid protein to enhance their immunogenicity. These recombinant genes (HEV-CSP, HEV-PfF2 and HEV-MSPFu24) with a *codon bias* optimized for subsequent expression in yeast cells, were cloned into the pPICZαA expression plasmid and expressed in yeast cell cultures *Pichia pastoris* KM71 as intracellular proteins recombinants.

Experimental activities continued with the selection at "lab-scale" level of recombinant clones that best expressed their desired proteins to subsequently use them on a large scale in industrial fermenters, after defining the optimal induction conditions. It is relevant to note variation in expression between pre-induction phase and induction phase, by adding methanol, in selected clones: induction for 5 days with 10% methanol proved to be the most suitable in achieving expression optimum for a recombinant HEV-CSP protein.

After two days of induction in 10% of methanol, optimum expression of the HEV-PfF2 protein was reached. Four days of induction in methanol were necessary to observe the optimum expression of the recombinant HEV-MSPFu24 protein. The need to express recombinant HEV-CSP, HEV-PfF2 and HEV-MSPFu24 proteins as intracellular proteins has led to use very energetic cell lysis conditions, such as high-pressure homogenization (28,000 Psi) with consequent presence in lysates of a large number of cellular proteins which determined a very evident background.

We therefore focused our attention on a single recombinant protein, HEV-CSP, and we started a downstream processing, to obtain a partial purification of the protein under examination.

Partial purification of the target protein HEV-CSP was conducted via ion exchange chromatography; first using linear chromatography and then 4-steps chromatography, where different Elution Buffer concentrations were deployed.

Recombinant HEV-CSP heterologous protein is currently in a subsequent purification phase: various strategies are being evaluated to obtain a suitable level of purity, to be suitably formulated in a single "protein-based" vaccine candidate, in presence or absence of adjuvants, in subsequent pre-clinical immunization studies in an animal model.

Acknowledgments

First of all, I would like to thank Prof. Agata Campisi, who followed me day after day throughout my doctoral path, managing to be a maternal figure on one hand and a guide towards my professional growth on the other. Thanks also goes to Dr. Agata Fazzio, who helped me in the development of my doctoral thesis project. I would like to thank Phd coordinator prof. Stefiania Stefani for her support in this research. Other thanks go to the love of my life Julia, who is close to me every day and who supported me greatly throughout my PhD program. Furthermore, I would like to thank all Prof. Rosaria Grasso, Rosalia Pellitteri, Giuseppina Raciti, Ugo Chiacchio and Salvatore Guccione who have followed me along this path, always giving me good advice on how to face this journey full of satisfactions and difficulties. I would also like to thank Dr. Francesca Scuderi who taught me everything about how to continue with my "cloning" process, always giving me help during the different phases of the research. Furthermore, I would like to thank Mr. Gaetano Galatà, who, having always ready jokes and many years of knowledge, has helped me in my human and professional growth. I would also like to thank all the Etna Biotech company staff, namely Epifanio Fichera, Viviana Gianninò, Antonio Di Mauro and last but not least the pivot of "our" company: Reinhard Glueck (who will be missed by everyone) and his wife Carmela Drago, who have allowed my experience at Etna Biotech. Also, I would like to thank my Slovenian tutor Miša Cajnko, a wonderful person who helped me in my journey abroad and my tutor Blaž Likozar, who allowed me to have this fantastic experience at the National Institute of Chemistry. Finally, I would like to

thank my entire family, starting with my mother Licia and my father Raffaele, who have always believed in me, my little brother Marcello who encourages me constantly to never give up and to always believe in myself. I also thank my sister-in-law Gerica and my uncles Lino, Nello, Salvatore and my aunt Monica who have supported me in everyday life. Thanks also to the whole Bisicchia family (Salvatore and Peppe) and Lola who for me are now like a second family. Furthermore, I thank all the undergraduates I have met and who have taught me so much recently and I would like to remember some of them: Ambra, Giada, Mariangela, Alessia, Giuseppe, Andrea, Carmelo, Alessandra, Federica and many many others. Finally, I would like to thank my friends Elena and Tano, who, between a laugh and an encouraging phrase, made my weekends happy and carefree. I also thank my friend Salvo Lombardo who helped me to improve the quality of my English and supported me throughout my experience. All those I have not mentioned will please pardon me but I would add this final thanks to all those who have been part of my path.

References

- Adams JH, Sim BK, Dolan SA, Fang X, Kaslow DC, Miller LH (1992). A family of erythrocyte binding proteins of malaria parasites. Proc Natl Acad Sci USA 89: 7085–7089.
- Arankalle VA, Chobe LP (1999). Hepatitis E virus: can it be transmitted parenterally? J Viral Hepat 6: 161-164.
- Aravind L, Iyer LM, Wellems TE, Miller LH (2003). *Plasmodium* biology: genomic gleanings. Cell 115: 771–785.
- Bell D, Wongsrichanalai C, Barnwell JW (2006). Ensuring quality and access for malaria diagnosis: how can it be achieved?. Nat Rev Microbiol 4: S7-S20.
- Bhandari PL, Raghuveer CV, Rajeev A, Bhandari PD (2008). Comparative study of peripheral blood smear, quantitative buffy coat and modified centrifuged blood smear in malaria diagnosis. Indian J Pathol Microbiol 51: 108-112.
- Bradley DW (1995). Hepatitis E virus: a brief review of the biology, molecular virology, and immunology of a novel virus. J Hepatol 22: 140-145.
- Cerami C, Frevert U, Sinnis P, Takacs B, Clavijo P, Santos MJ, Nussenzweig V (1992). The basolateral domain of the hepatocyte plasma membrane bears receptors for the circumsporozoite protein of *Plasmodium falciparum* sporozoites. Cell 70: 1021–1033.
- Chan M (2016). World malaria report 2015. World Health Organization. Feltrinelli (Ed), pp 1-100.

- Chau TN, Lai ST, Tse C, Ng TK, Leung VK, Lim W, Ng MH (2006). Epidemiology and clinical features of sporadic hepatitis E as compared with hepatitis A. Am J Gastroenterol 101: 292-296.
- Chotivanich K, Silamut K, Day NPJ (2006). Laboratory diagnosis of malaria infection-a short review of methods. Aust J Med Sci 27: 11-15.
- Clendennen TE 3rd, Long GW, Baird JK (1995). QBC and Giemsa-stained thick blood films: diagnostic performance of laboratory technologists. Trans R Soc Trop Med Hyg 89: 183-184.
- Coelho CH, Doritchamou JYA, Zaidi I, Duffy PE (2017). Advances in malaria vaccine development: report from the 2017 malaria vaccine symposium. NPJ Vaccines 30: 2-34.
- Coppi A, Natarajan R, Pradel G, Bennett BL, James ER, Roggero MA,Corradin G, Persson C, Tewari R, Sinnis P (2011). The malaria circumsporozoite protein has two functional domains, each with distinct roles as sporozoites journey from mosquito to mammalian host. J Exp Med 208: 341-356.
- Coppi A, Pinzon-Ortiz C, Hutter C, Sinnis P (2005). The *Plasmodium* circumsporozoite protein is proteolytically processed during cell invasion. J Exp Med 201: 27-33.
- Coursaget P, Buisson Y, N'Gawara MN, Van Cuyck-Gandre H, Roue R (1998). Role of hepatitis E virus in sporadic cases of acute and fulminant hepatitis in an endemic area (Chad). Am J Trop Med Hyg 58: 330-334.
- Cregg JM, Russell KA (1998). Trasformation. In: Higgins DR, Cregg JM (Eds) "Pichia Protocols Methods in Molecular Biology™". Humana Press, p. 103.

- Das K, Agarwal A, Andrew R, Frösner GG, Kar P (2000). Role of hepatitis E and other hepatotropic virus in aetiology of sporadic acute viral hepatitis: a hospital based study from urban Delhi. Eur J Epidemiol 16: 937-940.
- De Carneri I (2012). Parassitologia medica e diagnostica parassitologica. In: Genchi C, Pozio E (Eds). Casa Editrice Ambrosiana, pp. 1-432.
- De Cock KM, Bradley DW, Sandford NL, Govindarajan S, Maynard JE, Redeker AG (1987). Epidemic non-A, non-B hepatitis in patients from Pakistan. Ann Intern Med 106: 227-230.
- Deepak G, Singh V, Chitnis C (2013). Blood stage malaria vaccine. Nat Med 19: 168–178.
- Holder AA (2009). The carboxy-terminus of merozoite surface protein 1: structure, specific antibodies and immunity to malaria. Parasitol 136: 1445–1456.
- Holland CA, Kiechle FL (2005). Point-of-care molecular diagnostic systems-past, present and future. Curr Opin Microbiol 8: 504- 509.
- Khuroo MS, Dar MY (1992). Hepatitis E: evidence for person-to-person transmission and inability of low dose immune serum globulin from an indian source to prevent it. Indian J Gastroenterol 11: 113-116.
- Khuroo MS, Kamili S, Khuroo MS (2009). Clinical course and duration of viremia in vertically transmitted hepatitis E virus (HEV) infection in babies born to HEV-infected mothers. J Viral Hepat 16: 519-523.
- Klotz FW, Orlandi PA, Reuter G, Cohen SJ, Haynes JD, Schauer R, Howard RJ, Palese P, Miller LH (1992). Binding of *Plasmodium falciparum* 175-kilodalton erythrocyte binding antigen and invasion of murine erythrocytes requires N-

acetylneuraminic acid but not its O-acetylated form. Mol Biochem Parasitol 51: 49–54.

- Kurtzman CP (2009). Biotechnological strains of *Komagataella (Pichia) pastoris* are *Komagataella phaffi* as determined from multigene sequence analysis". J Ind Microbial Biotecnol 36: 1435-1438.

- Kwo PY, Schlauder GG, Carpenter HA, Murphy PJ, Rosenblatt JE, Dawson GJ, Mast EE, Krawczynski K, Balan V (1997). Acute hepatitis E by a new isolate acquired in the United States. Mayo Clin Proc 72: 1133-1136.

- Langhorne J, Ndungu FM, Sponaas AM, Marsh K (2008). Immunity to malaria: more questions than answers. Nat Immunol 9: 725–732.

- Lu L, Li C, Hagedorn CH (2006). Phylogenetic analysis of global Phylogenetic analysis of global hepatitis E virus sequences: genetic diversity, subtypes and zoonosis. Rev Med Virol 16: 5-36.

- Mawson AR (2013). The pathogenesis of malaria: a new perspective. Pathog Glob Health. 107: 122-129.

- Mayor A, Bir N, Sawhney R, Singh S, Pattnaik P, Singh SK, Sharma A, Chitnis CE (2005). Receptor-binding residues lie in central regions of Duffy-binding-like domains involved in red cell invasion and cytoadherence by malaria parasites. Blood 105: 2557-2563.

- Mazumdar S, Mukherjee P, Yazdani SS, Jain SK, Mohmmed A, Chauhan VS (2010). *Plasmodium falciparum* merozoite surface protein 1 (MSP-1)-MSP-3 chimeric protein: immunogenicity determined with human-compatible adjuvants and induction of protective immune response. Infect Immun 78: 872-883.

- McCutchan TF, Kissinger JC, Touray MG, Rogers MJ, Li J, Sullivan M, Braga EM, Krettli AU, Miller LH (1996). Comparison of circumsporozoite proteins from avian and mammalian malarias: biological and phylogenetic implications. Proc Natl Acad Sci U S A 93: 11889-11894.
- McMorrow ML, Masanja MI, Abdulla SM, Kahigwa E, Kachur SP (2008). Challenges in routine implementation and quality control of rapid diagnostic tests for malaria-Rufiji District, Tanzania. Am J Trop Med Hyg 79: 385-390.
- Meng XJ, Purcell RH, Halbur PG, Lehman JR, Webb DM, Tsareva TS, Haynes JS, Thacker BJ, Emerson SU (1997). A novel virus in swine is closely related to the human hepatitis E virus. Proc Natl Acad Sci USA 94: 9860-9865.
- Morassin B, Fabre R, Berry A, Magnaval JF (2002). One year's experience with the polymerase chain reaction as a routine method for the diagnosis of imported malaria. Am J Trop Med Hyg 66: 503- 508.
- Mwangi TW, Mohammed M, Dayo H, Snow RW, Marsh K (2005). Clinical algorithms for malaria diagnosis lack utility among people of different age groups. Trop Med Int Health 10: 530-536.
- Myung JM, Marshall P, Sinnis P (2004). The *Plasmodium* circumsporozoite protein is involved in mosquito salivary gland invasion by sporozoites. Mol Biochem Parasitol 133: 53-59.
- Ngasala B, Mubi M, Warsame M, Petzold MG, Massele AY, Gustafsson LL, Tomson G, Premji Z, Bjorkman A (2008). Impact of training in clinical and microscopy diagnosis of childhood malaria on antimalarial drug prescription and health outcome at primary health care level in Tanzania: a randomized controlled trial. Malar J 7: 199-210.

- Olotu A, Fegan G, Wambua J, Nyangweso G, Awuondo KO, Leach A, Lievens M, Leboulleux D, Njuguna P, Peshu N, Marsh K, Bejon P (2013). Four-year efficacy of RTS,S/AS01E and its interaction with malaria exposure. N Engl J Med 368: 1111-1120.
- Olotu A, Fegan G, Wambua J, Nyangweso G, Leach A, Lievens M, Kaslow DC, Njuguna P, Marsh K, Bejon P (2016). Seven-year efficacy of RTS,S/AS01 malaria vaccine among young african children. N Engl J Med. 374: 2519-2529.
- Orlandi PA, Klotz FW, Haynes JD (1992). A malaria invasion receptor, the 175-kilodalton erythrocyte binding antigen of *Plasmodium falciparum* recognizes the terminal Neu5Ac(alpha 2-3)Gal- sequences of glycophorin A. J Cell Biol 116: 901-909.
- Rathore D, Hrstka SC, Sacci JB Jr, De la Vega P, Linhardt RJ, Kumar S, McCutchan TF (2003). Molecular mechanism of host specificity in *Plasmodium falciparum* infection: role of circumsporozoite protein. J Biol Chem 278: 40905-40910
- Robson SC, Adams S, Brink N, Woodruff B, Bradley D (1992). Hospital outbreak of hepatitis E. Lancet 339: 1424-1425.
- Schlauder GG, Mushahwar IK (2001). Genetic heterogeneity of hepatitis E virus. J Med Virol 65: 282-292.
- Shi Q, Lynch MM, Romero M, Burns JM Jr (2007). Enhanced protection against malaria by a chimeric merozoite surface protein vaccine. Infect Immun 75: 1349-1358.
- Shrestha MP, Scott RM, Joshi DM, Mammen MP, Thapa GB, Thapa N, Myint KS, Fourneau M, Kuschner RA, Shrestha SK, David MP, Seriwatana J, Vaughn DW,

Safary A, Endy TP, Innis BL (2007). Safety and efficacy of a recombinant hepatitis E vaccine. N Engl J Med 356: 895-903.

- Sidjanski SP, Vanderberg JP, Sinnis P (1997). Anopheles stephensi salivary glands bear receptors for region I of the circumsporozoite protein of *Plasmodium falciparum*. Mol Biochem Parasitol 90: 33-41.
- Sim BK, Orlandi PA, Haynes JD, Klotz FW, Carter JM, Camus D, Zegans ME, Chulay JD (1990). Primary structure of the 175K *Plasmodium falciparum* erythrocyte binding antigen and identification of a peptide which elicits antibodies that inhibit malaria merozoite invasion. J Cell Biol 111: 1877-1884.
- Tamborrini M, Stoffel SA, Westerfeld N, Amacker M, Theisen M, Zurbriggen R, Pluschke G (2011). Immunogenicity of a virosomally formulated *Plasmodium falciparum* GLURP-MSP3 chimeric protein-based malaria vaccine candidate in comparison to adjuvanted formulations. Malar J 10:359-370.
- Teshale EH, Hu DJ, Holmberg SD (2010). The two faces of Hepatitis E virus. Clin Infect Dis 51: 328-334.
- Teshale EH, Hu DJ (2011). Hepatitis E: epidemiology and prevention. World J Hepatol 3: 285-291.
- Tolia NH, Eric J, Enemark EJ, Sim BK, Joshua-Tor L (2005). Structural basis for the EBA-175 erythrocyte invasion pathway of the malaria parasite *Plasmodium falciparum*. Cell 122: 183-193.
- Velázquez O, Stetler HC, Avila C, Ornelas G, Alvarez C, Hadler SC, Bradley DW, Sepúlveda J (1990). Epidemic transmission of enterically transmitted non-A, non-B hepatitis in Mexico, 1986-1987. JAMA 263: 3281-3285.

- Viswanathan R (1957). A review of the literature on the epidemiology of infectious hepatitis. Indian J Med Res 45: 145-155.

Sitography

- http://www.salute.gov.it/portale/temi/p2_6.jsp?lingua=italiano&id=1983&area=Malattie%20infettive&menu=viaggiatori
- http://apicomplexity.blogspot.com/2013/
- https://www.malariavaccine.org/malaria-and-vaccines/vaccine-development/life-cycle-malaria-parasite
- https://www.uspharmacist.com/article/malaria-treatment-and-prevention

Index

La borsa di dottorato è stata cofinanziata con risorse del
Programma Operativo Nazionale Ricerca e Innovazione 2014-2020 (CCI 2014IT16M2OP005),
Fondo Sociale Europeo, Azione I.1 “Dottorati Innovativi con caratterizzazione Industriale”

UNIONE EUROPEA
Fondo Sociale Europeo