Hemagglutinin-esterase (HE) from Infectious Salmon Anemia virus (ISAv): Characteristics and attempted expression

by

Viviana Carolina Cerda Llanos

B.Sc., University of Concepcion, 2009

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

In the
Department of Molecular Biology and Biochemistry
Faculty of Science

SIMON FRASER UNIVERSITY

Spring 2015

Approval

Name: **Viviana Carolina Cerda Llanos**

Degree: **Master of Science**

Title: ***Hemagglutinin-esterase (HE) from Infectious Salmon Anemia virus (ISAv): Characteristics and attempted expression***

Examining Committee: **Chair:** Dr. Nancy Hawkins
Associate Professor

Dr. Andrew Bennet
Senior Supervisor
Professor

Dr. William Davidson
Supervisor
Professor

Dr. Dipankar Sen
Supervisor
Professor

Dr. Rosemary Cornell
Internal Examiner
Professor
Department of Molecular Biology and Biochemistry

Date Defended/Approved: April 2nd, 2015

Partial Copyright Licence

Abstract

Infectious salmon anemia virus (ISAv) is a pathogen that mainly affects Atlantic salmon (*Salmo salar*), which are commonly grown in the aquaculture industry. The resulting disease, infectious salmon anemia, has caused large financial losses for this industry. ISAv is a member of the *Orthomyxoviridae* family, as are the influenza type A, B and C viruses, and it belongs to the genus Isavirus. ISAv has a single dual-functional surface protein that is involved in the interaction with the host cells, namely: a hemagglutinin-esterase (HE). The HE protein binds preferentially to 4-*O*,5-N-diacetylneuraminic acid residues that are present on the target cells, which in this case are salmon erythrocytes. As is the case for influenza, the binding of HE protein to 4-*O*-acetylsialosides (receptor function) triggers infection in salmon, whereas the esterase active site, which possesses the 'receptor destroying' activity hydrolyzes the 4-*O*-acetyl groups, aiding the release of the viral progeny to infect neighbouring cells. This thesis reports the cloning and the attempted expression of the haemagglutinin-esterase from two different ISAv strains, that is a Canadian and a Norwegian strain, using baculovirus expression systems.

Keywords: Aquaculture; Infectious salmon anemia virus (ISAv); Hemagglutinin esterase (HE); Atlantic salmon (*Salmo salar*); baculovirus

Dedication

This thesis is dedicated to all the people who supported me throughout this journey, my family and friends.

Acknowledgements

First of all, I would like to thank my supervisor Dr. Andrew Bennet for taking me into his research family and for giving me this exciting project. I thank him for always being optimistic and encouraging, even when the scenario was not the best.

I would also like to thank Dr. William Davidson, who was always someone I could talk to about professional and personal topics. Dr. Davison made the graduate school journey possible; from my arrival from my home country, Chile, to the completion of my Masters program.

I should not forget to thank Dr. Dipankar Sen, from giving me helpful suggestions and for asking very interesting questions on the meetings we had every year, that were useful for troubleshooting the issues that this project may have had. Also, Dr. Rosemary Cornell, my internal examiner, who played a very important role in the development of my program. She facilitated most of the materials that I needed to start working, helped solving the problems that I encountered, and also for giving me helpful insights in the writing of this thesis.

So far I have thanked my supervisors and superiors, but now I would like to turn to my colleagues from the Bennet laboratory, past and current. For the past members, Dr. Fahimeh Shidmoossavee, Dr. Lydia Cheng, Dr. Saswati Chakladar, Miss Kobi Khazaei, and Dr. Jefferson Chan. I thank them for taking me into the group so warmly. I learned so much with them from day one, and it was an honour to have had the opportunity to work with all of them, especially Dr. Shidmoossavee, who collaborated in this project, helped me with chemistry and enzyme kinetics whenever I was having problems, for listening and for all the advice. And for the current members, Natalia Sannikova, Saeideh Shamsi, Marco Farren-Dai, Matthew Deen, Yumeela Ganga-Sah, Cinzia Colombo, Christopher Adamson and Xiaoyang Shan. I thank them for all the good times and for making the lab a fun place to work.

I should definitely thank Dr. Anita Muller, Dr. Turhan Markusen from the Veterinarian Institute in Oslo and Dr. Børre Robertsen from the University of Tromsø,

Norway, for facilitating important of the materials for this project. Special thanks to Dr. Muller, who was very supportive and always willing help whenever I had doubts and questions.

Finally, I would like to finish with two important people: First, my very good friend Nevena Cekic, who was always by my side supporting me in every possible situation. She has become one of my best friends and a person I can always rely on, no matter what; and second, my dearest boyfriend Tristan Gillis, who came to my life in the right moment, always supporting me on everything and for giving me a kind of happiness that I never thought I could have.

To all of them and to all the people I might be forgetting, thank you for being a part of my professional and personal life.

Table of Contents

List of Tables

List of Figures

List of Acronyms

4-OAc MUNANA	4-methylumbelliferyl *N*-acetyl-4-*O*-acetyl-α-D-neuraminic acid
*Ac*MNPV	*Autographa californica* multiple nucleopolyhedrovirus
BEV	Baculovirus expression vector
BEVs	Baculovirus expression vector system
GalNAc	*N*-acetylgalactosamine
GlcNAc	*N*-acetylglucosamine
HE	Hemagglutinin-esterase
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid, N-(2-Hydroxyethyl)piperazine-N′-(2-ethanesulfonic acid)
HTV	High titer virus
ISA	Infectious Salmon Anemia
ISAv	Infectious Salmon Anemia virus
Man	mannose
MUNANA	4-methylumbelliferyl *N*-acetyl-α-D-neuraminic acid
*Mv*NA	*Micromonospora viridifaciens* neuraminidase
NaCl	Sodium chloride
NaF	Sodium fluoride
NaPPi	Sodium pyrophosphate
NaPi	Sodium phosphate
Neu5Ac	*N*-acetylneuraminic acid
Neu5,9Ac_2	5-*N*-acetyl-9-*O*-acetylneuraminic acid
ORF	Open reading frame
p-NPA	*p*-nitrophenyl acetate
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide electrophoresis
Tris	Tris-(hydroxymethyl)aminomethane
UV	Ultraviolet

Chapter 1. Introduction

Aquaculture is defined as the rearing of aquatic animals or the cultivation of aquatic plants for food. With regard to fish, this industry is a developing component of the overall strategy towards sustainable fisheries, and as such has a high economic importance for the countries that participate in this business. Over the last few decades, the marine aquaculture industry has undergone a dramatic increase in size, which was triggered by a high demand for farmed products including fish, crustaceans, molluscs, and seaweeds. The aquaculture of salmon has been a particular growth industry and it is now one of the leading marine species being farmed.

Among all marine species being cultured, salmon is one of the most important and since 1985 this industry has experienced an annual growth in production of 16%. Norway is the main farmed salmon producer with over 50% of the world's industry, which primarily involves Atlantic salmon (*Salmo salar*) [1]. From 1985, total worldwide salmon production greatly increased [2, 3] as at least 24 countries, according to the Food and Agriculture Organization (FAO), became producers while Norway, Chile, Scotland and Canada were the major producers of farmed fish (FAO 2005-2014). Salmon farming can be sub-divided into three main categories that relate to the species being farmed, with these being the chinook salmon (*Oncorhynchus tshawytscha*), the coho salmon (*Oncorhynchus kisutch*) and the atlantic salmon, the latter representing almost 90% of the total production [4].

1.1. Aquaculture and disease.

As with every developing industry, salmon aquaculture has encountered problems during its growth phase and is still struggling either to prevent or to fix these issues. For instance, the industry kept cages or pens that are located in the natural environment, for example in a bay near the shore or in a sea loch. As a result, nutrients, pathogens and chemicals are released into the marine environment [4] and this results

in a major problem that farmed salmon encounter being the spread of infectious diseases. These include parasites [5], fungi [6], bacteria [7], and viruses [8].

This thesis will be focused mainly on one viral pathogen of the atlantic salmon that is named Infectious Salmon Anaemia virus or ISAv.

1.2. ISAv Infectivity

Infectious Salmon Anemia (ISA) is a disease that affects atlantic salmon, although it has also been reported in other salmonids, such as the Chilean Pacific coho salmon (*Oncorhynchus kisutch*) and the Irish Rainbow trout (*Oncorhynchus mykiss*) [9, 10], however, the disease is asymptomatic in these two species. The symptoms of this disease in farmed atlantic salmon include anaemia, accumulation of fluid in the peritoneal cavity (ascites), swelling of the spleen and liver, skin haemorrhages, pale gills, and intestinal congestion [11, 12]. Ever since the first reported outbreak of ISA, which occurred in Norway in 1984 [13], this disease has caused dramatic losses in farmed salmon production and as such lead to a financial crisis in the Norwegian industry. Moreover, ISA has been reported in other salmon producing countries including Canada, Chile, Scotland and the United States [14, 15].

The causative agent of this disease is the Infectious Salmon Anemia virus (ISAv). This virus belongs to the *Isavirus* genus, and it has been classified as a member of the *Orthomyxoviridae* family [16] because it has been shown to have similar biochemical, morphological, and physiochemical characteristics, as well as a high genomic homology, to the various influenza viruses [17-19].

1.3. Relationship between ISAv and Influenza viruses

1.3.1. Influenza viruses.

One of the most important human diseases is Influenza, causing an annual mortality of approximately 36,000 in the United States according to Thompson et al.,

2003. The organism responsible for this disease is a viral agent. An influenza type A virus was responsible for the 1918 Spanish flu pandemic in which up to 20 million people died worldwide [20]. The influenza virus is also responsible for the yearly influenza epidemics during which the death toll is typically an approximate 50,000 people [21].

The structure of the Influenza type A virus.

The *Orthomyxoviridae* family are enveloped viruses and they possess between six and eight segments of linear negative sense single stranded RNA. In the case of influenza virus, these RNA segments encode for at least eleven proteins. These proteins include nucleoprotein (NP), polymerases PB1, PB2 and PA that interact with the viral RNA [10]; and two spike-like glycoproteins that each play important roles in the viral replication and infectivity cycles. These two proteins are hemagglutinin (HA) and neuraminidase (NA), which is also called sialidase (S). The HA protein, which is a functional homotrimer, has receptor binding and fusion activities and it participates in the first stage of the infection cycle by recognizing a cell surface receptor, sialic acid or neuraminic acid. The virus fuses to the cell membrane by the multivalent binding of several HAs to the sialic acids present on the cell surface, allowing the entry of the pathogen into the host cell. The viral neuraminidase, which is a homotetrameric enzyme, cleaves terminal sialic acid residues from the cell surface glycoconjugates and thus it is known as the ‘receptor destroying enzyme’ or ‘RDE’. These cleavage events are important for release of the newly emerged viral progeny into the environment, thereby allowing the progeny to infect neighbouring cells [21].

1.3.2. Sialic acids

Many years ago, two researchers working independently discovered that salivary mucins and brain glycolipids released a negatively charged carbohydrate upon mild acid treatment. In 1936, Gunnar Blix isolated this compound from bovine submaxillary mucin and named it “sialic acid”, after the Greek word *sialo,* for saliva. Simultaneously, Ernst Klenk discovered it performing autopsies in brain tissues of Tay-Sach patients and he called this sugar neuraminic acid [22]. Both researchers agreed that the scientific community can use both names as synonymous, when referring to

this structure and also for the enzyme that cleaves the glycosidic linkage between the oligo (or poly) -saccharide and the sialic acid (sialidase or neuraminidase) [23]. The most common sialic acid is α-D-*N*-acetylneuraminic acid or Neu5Ac (5-acetamido-3,5-dideoxy-D-*glycero*-D-*galact*o-non-2-ulosonic acid) shown in Figure 1.1 and it is often the terminal residue of various *N*-glycans, *O*-glycans, and glycosphingolipids that are found in animal cells.

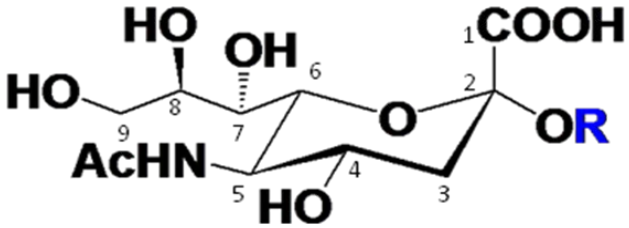

Figure 1.1 Generic structure of α-D-sialosides (*N*-acetylneuraminides) with the carbon atoms numbered.
The R group represents the various glycoconjugates to which the sialic acid is attached. Numbers show the position of the carbon atoms

Sialic acids are often located at the termini on cellular glycoconjugates and it is this prominent position along with their negative charge that result in these carbohydrates participating in numerous important intermolecular and intercellular biological interactions. For instance, such as those occurring in the neuronal cells of vertebrates [24]. In addition, pathogens can recognize sialic acids present on their target cells, therefore they also have a receptor function [25].

Interaction of sialic acid with proteins and biological importance

There is a group of proteins that have an affinity with carbohydrates such as sialic acid. These proteins are called lectins and they can be categorized as vertebrate pathogen lectins, vertebrate endogenous lectins and lectins from other sources such as plants, although the biological importance of this last group still remains unknown.

Some pathogens such as virus and bacteria can develop lectins that recognize sialic acids on host surface glycolipids/proteins and use such sugars as the binding receptor. These lectins are used specifically, in the case of viruses, to enter the host cell and replicate once inside. Once these proteins recognize and bind to the sialic acids they agglutinate the host cells, mainly erythrocytes. Because of this, they are also

known as hemagglutinins. The Influenza A virus hemagglutinin is the most studied of this family [26-28]. The receptor specificity for this lectin is based mostly on the different sialic acids that the host cell expresses [29]. In addition, viral hemagglutinins recognize sialic acids that have glycosidic bonds from C2 to the O3 or O6 of galactose (Gal) or *N*-acetylgalactosamine (GalNAc) forming α-2,3 and α-2,6 glycoconjugates [30]. As for the case of bacteria, they can express lectins on their surface, which recognize sialic acids from the host cell. These bacterial lectins are called adhesins. The expression of the adhesins are strain dependent and they may determine the type of tissue the bacteria can infect [31], as is the case for *Helicobacter pylori* and *Escherichia coli* K99 strains. *H. pylori* recognizes sialic acids in the host cell, however, some strains like *E.coli K99 and Pasteurella haemolytica* recognize Neu5Gcα2-3Gal*β*1-4Glc found in pigs intestinal tract [32] and GlcNAc [33] structures respectively. Bacterial lectins can also be secreted in soluble form, although in this state, they are generally toxins. These toxins are also sialic acid dependent, such as the ones secreted by *Clostridium tetani* and *Clostridium botulinum* both agents responsible for tetanus and botulism respectively [30].

Similar to pathogens, vertebrates can express endogenous lectins that can also recognize sialic acids, many of which are found in the immune system. The most common endogenous lectins are siglecs, factor H and selectins. Siglecs are the most abundant superfamily of sialic acids binding lectins [34]. It has been shown that a variety of siglecs can be expressed in human cells related to the immune system such as siglec-1, expressed in macrophages; siglec-2/CD22, expressed in B-cells and in charge of regulating their activation; siglec-3, expressed in monocytes; and siglec-4a, expressed in Schwann cells and oligodendroglia [35].

Factor H are lectins that play a role in the early events of the innate immune system response, the alternative pathway of complement. It possesses 20 short consensus repeats constituted by approximately 60 amino acids per unit and recognizes sialic acids that are unsubstituted on carbons 7 and 9 [36]. The binding of the Factor H lectins to the cell surface glycoconjugates that have terminal sialic acids, assures that the external pathogen is targeted and prevents ‘self’ attacks from the immune system to their own cells. Lastly, selectins are lectins that recognize carbohydrates that contains sialic acids and fucose called sialyl Lewis x (Le^x) and sialyl Lewis a (Le^a) [37]. There are three different types, namely E- L- and P- selectins, all

involved in leukocyte trafficking (initiating inflammation) and therefore in cancer related diseases [38].

1.3.3. Infectious Salmon Anemia virus (ISAv).

Structure of ISAv.

ISAv belongs to the new genus *Isavirus* [39] and it shares similar structural characteristics with influenza type A virus as it also belongs to the *Orthomyxoviridae* family. This virus is the first member from this family that has been characterized in fish. ISAv is an enveloped virus and has genome consisting of a single stranded RNA, that is cut into eight negative sense segments that encode for at least 10 proteins [40, 41].

On the surface of ISAv there are two major proteins that are involved in the viral infection cycle, and these are fusion (F) protein that is encoded by gene segment 5 of its RNA, and a hemagglutinin-esterase (HE) encoded by gene segment 6 [12]. The dual function hemagglutinin-esterase protein provides the virus with a receptor binding activity (hemagglutinin domain) and a receptor destroying activity (RDA) (esterase domain) [42]. Essentially, the HE protein has one domain to recognize and bind to sialic acids for uptake into cells, and a separate esterase domain for enabling the release of mature virus from the plasma membrane of the host cell. This is the major difference between the influenza viruses and ISAv, since the HE protein possesses the two biological functions that require two proteins (HA and NA) in the influenza virus. This is shown schematically in Figure 1.2.

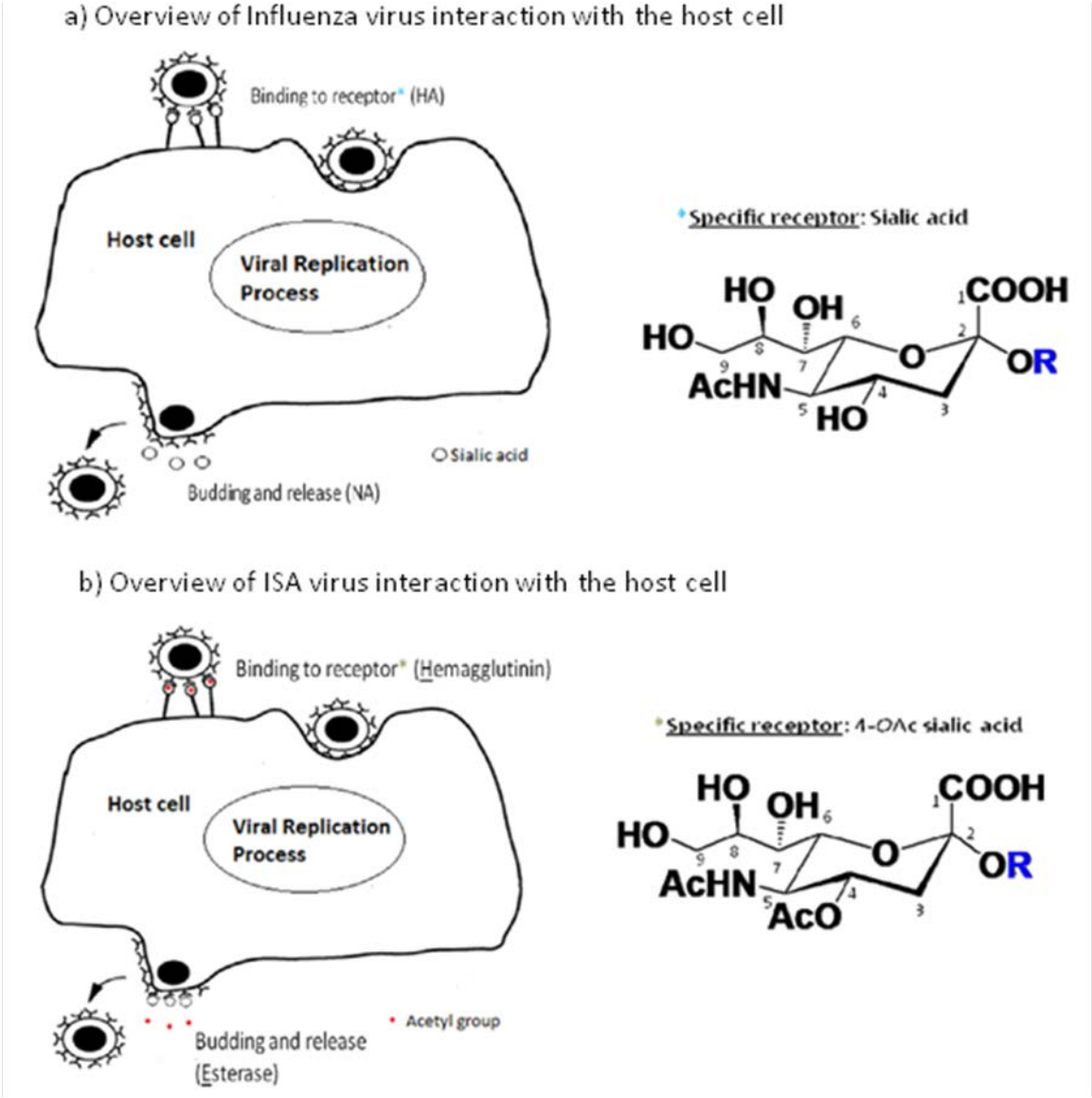

Figure 1.2 Schematic representation of the influenza and ISA virus infection. **a)** Influenza A virus infection in the host cell. The proteins involved in binding (hemagglutinin (HA)) and cleavage (neuraminidase (NA)) of the receptor are shown. **b)** ISA virus infection in the host cell. The proteins involved in binding (hemagglutinin) and cleavage (esterase) of the receptor are shown. Major differences between the two processes can be noted: the specific receptors for both viruses and what is released by the action of the NA (sialic acid groups) and esterase enzymes (acetyl groups) for Influenza and ISA viruses, respectively.(Modified from Amano et al., 2005)

Figure 1.4a shows the structure of the dual functional HE dimeric protein. This three dimensional structure was reported by Muller et al, 2010 [43]. It was obtained by performing a comparative alignment between the HE sequence of ISAv and the HE sequence of the porcine torovirus (pToV HE), since it shares great amino acid similarity [43]. The structures generated from this analysis correspond to the ones matching to the pToV HE protein. These include the receptor-binding portion (hemagglutinin domain), the receptor-destroying side (esterase domain), and also a membrane-proximal domain, which has been reported to be involved in protein trafficking to the cell membrane [44, 45].

The ISAv HE is located on the surface of the viral particles. It has a lower expression level than does the hemagglutinin lectin that occurs on the surface of influenza viral particles and HE has an approximate molecular weight of 43 kDa [16, 46]. It has been suggested that an acetylesterase is responsible for its receptor destroying activity. This activity was demonstrated by the catalyzed hydrolysis of *p*-nitrophenyl acetate (*p*-NPA) by the purified protein from ISAv [16] and by the inhibition of hydrolysis of this substrate using the serine esterase inhibitors diisopropyl fluorophosphate (DFP) and 3,4-dichloroisocoumarin (DCIC) [10]. Moreover, a mutation of a serine (Ser_{32}) to an alanine (Ala_{32}) residue in the HE sequence from the Norwegian isolate Glesvaer2/90 showed that the serine residue is vital for the acetylesterase activity of the protein [43]. Similarities exist between the amino acid sequences of hemagglutinins from Influenza C and ISAv. Specifically, the similarity is that both proteins contain a typical esterase catalytic triad. The general mechanism of action for esterases is represented in Figure 1.3, when a serine residue is the nucleophilic residue of the catalytic triad.

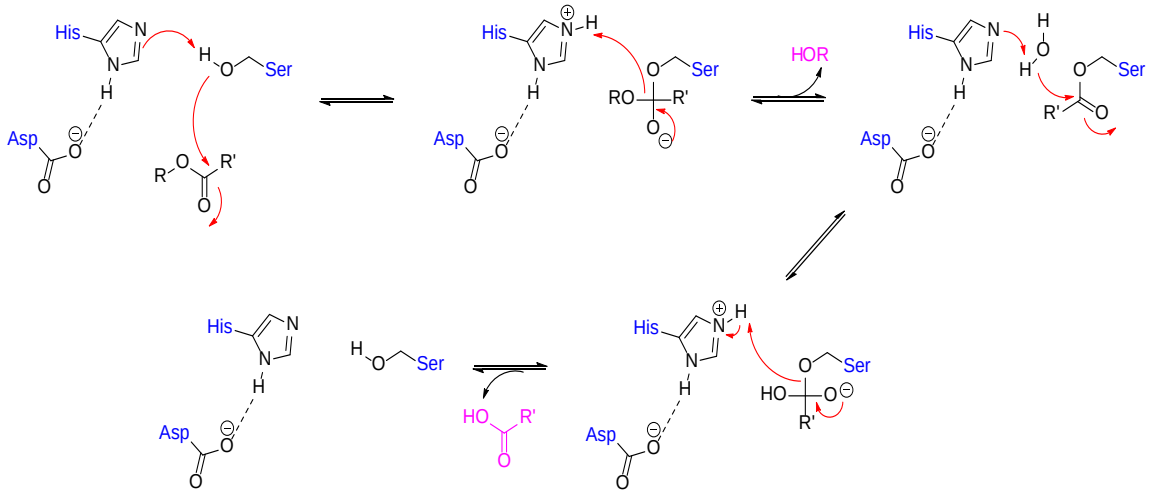

Figure 1.3 Generally accepted mechanism of serine esterases that proceeds via an acyl-enzyme intermediates
Shown in the figure are the catalytic triad (Ser-His-Asp) and the functional groups interacting with the ester substrate.

For both viruses, the amino acids composing this triad are Serine-Histidine-Aspartic acid (Ser-His-Asp). Interestingly, the location of these residues within the respective proteins differ, they are Ser_{57}, His_{355} and Asp_{352} for the Influenza C virus and Ser_{32}, His_{264} and Asp_{261} for the ISA viral proteins, specifically for the Norwegian isolate Glesvaer 2/90 [10, 47]. Figure 1.4b shows the catalytic triad of the ISAv HE protein from the pToV HE in an enhanced view of the esterase domain. A linear map of the ISAv HE protein is shown on Figure 1.4c.

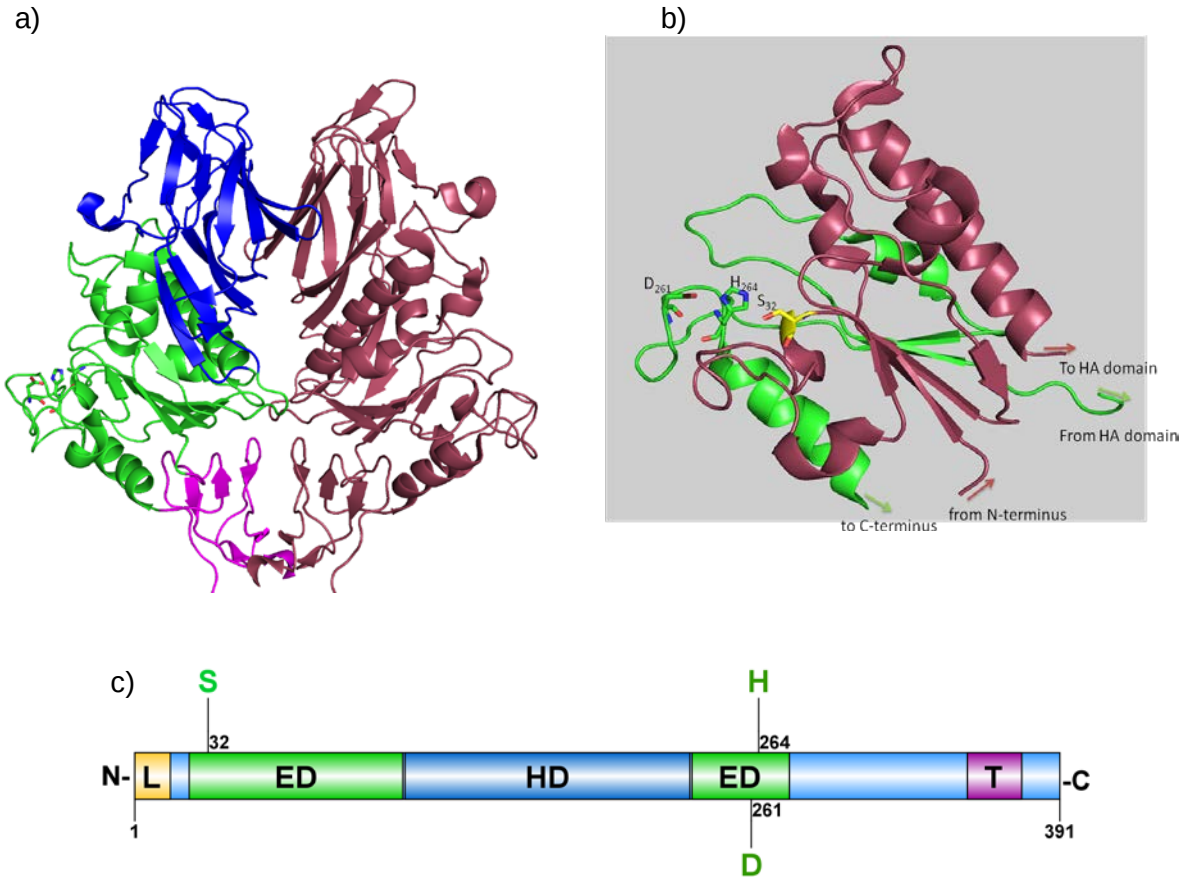

Figure 1.4 Structure of the hemagglutinin-esterase from porcine torovirus. **a)** The dimeric structure of the porcine torovirus (pToV) HE protein (pdb 3I1K).The left monomer has the hemagglutinin domain shown in blue, the esterase domain in green, which includes the three catalytic amino acid residues, and the membrane-proximal domain is shown in pink. The right monomer of the protein is shown in copper. **b)** An enlarged representation of the esterase domain is colour coded. The domain segment close to the N-terminal is shown in red. The second domain segment, which is closer to the C-terminal of the protein, is shown in green. The amino acids (S_{32}-H_{264}-D_{261}) from the catalytic triad are also depicted. **c)** Protein map of ISAv HE protein showing the conserved domains, ED (esterase domain), HD (hemagglutinin domain), T (transmembrane domain), and L (leader sequence). The amino acid residues from the catalytic triad are shown explicitly.

Previous work has shown that the esterase activity for ISAv HE protein results in hydrolysis of an isomeric acetylated sialic acid to generate sialic acid [16, 47]. It has also been described that the substrate specificity of the ISAV HE protein shows a preference for the specific residue 4-*O* acetylated sialic acid, shown in Figure 1.4 [19]. This is another factor that differentiates ISA virus from influenza viruses, since the specific receptor for the latter are sialic acids (Neu5Ac) and 5-*N*-acetyl-9-*O*-acetyl sialic acid (Neu5,9Ac_2) for influenza type C virus, specifically. The receptor for ISA virus HE protein can be found in endothelial and endocardial cells as well as in polymorphonuclear leukocytes of fish [16, 48] and specifically for Atlantic salmon, in endothelial cells lining the blood vessels [49].

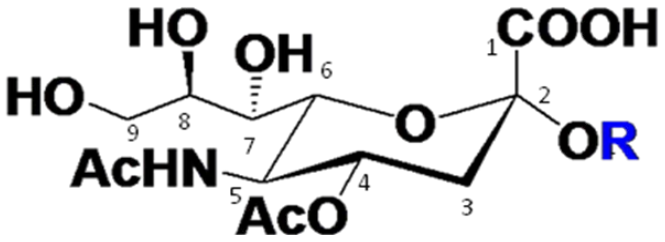

Figure 1.5 Structure of a generic 4-*O*-acetylated-α-sialoside.
Such a structure can also be called 5-*N*-acetyl-4-*O*-acetyl-α-neuraminic acid or Neu4,5Ac_2. The R group in blue represents a glycoconjugate to which it is attached.

HE is a glycoprotein that requires glycosylation at the time it is being expressed. These events are called post-translational modifications or PTMs and represent a key step for the maturation and proper folding of a protein, in order for it to be functional.

1.4. Post-translational modifications and glycosylation

After the mRNA is translated into the primary structure of a protein it must go through several maturation processes in order that it can fold properly and undergo final processing to become a fully functional protein. These series of covalent events alter the properties of the final protein via proteolytic cleavage and addition of modifying groups onto specific amino acids residues from the polypeptide chain. They are grouped together under the name post-translational modifications or PTMs. Activity, turnover and protein-protein interactions can also be determined by these PTMs [50].

There are many different post-translational modifications that play important roles in signal transduction [51], of which phosphorylation is the most common. However, in this thesis I will discuss glycosyl PTMs only. In general, glycosylation is the addition of sugar residues to an amino acid side chain of a protein [52]. The glycosidic linkages between the two molecules may occur on amino acid residues such as asparagine (Asn), serine (Ser) or threonine (Thr) and there may be multiple such modifications on a single protein. The glycosylation pathway in eukaryotic cells such as mammalian, insect and plants mainly occurs in the rough endoplasmic reticulum (ER) where folding, assembly and the basic modification takes place before being transported to the Golgi (ER-Golgi pathway) to complete more complex covalent modifications such as *N*-,*O*-glycosylations, sulfation and palmitylation [53, 54]. If the initial glycosylation in the ER is successful, the protein is delivered to the Golgi via an intermediate compartment. However, not all of the glycosylations take place in the ER. Some proteins are bound to glycans in the cytoplasm, nucleus and in some cases in the plasma membrane as it is the case of prokaryotic cells, such as bacteria [55].

The types of protein glycosylations can be categorized in five main categories, such as *N*-, *O*- and *C*- linked glycosylations, glycosylphosphatidylinositol (GPI) anchored and phosphoglycosylation. Formation of *N*-linked glycan PTMs are covalently attached to a side chain amide nitrogen atoms of asparagines; *O*-linked means that glycans are attached to the hydroxyl group of a serine or a threonine residue and *C*-linked glycosylation corresponds to an α-mannose that is covalently linked to the indole ring of a tryptophan residue.

The target protein of my thesis undergoes an *N*-glycosylation PTM after expression that is required for full functionality. Since there are several types of glycosylation PTMs, I will only describe *N*-type for the sake of brevity.

1.4.1. *N*-Glycosylation

N-glycoslylation occurs in the ER and the Golgi apparatus of eukaryotic cells, and it can be described as trimming reactions performed by glycosidases and also branching and elongations reactions, mediated by glycosyltransferases [56]. Until

recently, it was believed that *N*-glycosylation occurred only in eukaryotic cells, however it has been demonstrated that organisms belonging to the Eubacteria and Archea domains also undergo this process [57, 58]. *N*-glycosylated proteins such as pilins and flagellins can be found in *Pseudomonas aeruginosa* [59], *Caulobacter crescentus* [60] and in different Campylobacter species [61], although their glycans are very different from the ones found in eukaryotes [62, 63].

In eukaryotic cells it has been shown that differences in the structure of the *N*-glycans exist between insect and mammalian cells. Specifically, insect cells display less complex glycans when compared to mammalian cells [64, 65], an observation that suggests that the latter has much more intricate trimming and elongation pathways than do insects [66]. The initial steps are common for both cell lines (Fig 1.5) and they are carried out in the ER. The preassembled oligosaccharide containing glucose (Glc), mannose (Man) and *N*-acetylglucosamine (GlcNAc) ($Glc_3Man_9GlcNAc_2$) formed in the cytosolic face of ER, is transferred from the lipid anchor dolichol pyrophosphate into the lumen face of the ER by the RAFT1 protein, where it binds to the side chain amidic nitrogen atom present in the recognition sequence Asn-X-Thr/Ser (being X any amino acid except proline (Pro)) by the enzymatic action of a oligosaccharyltransferase [67]. This lipid anchor is synthesized through the same metabolic route as cholesterol by the phosphorylation of dolichol.

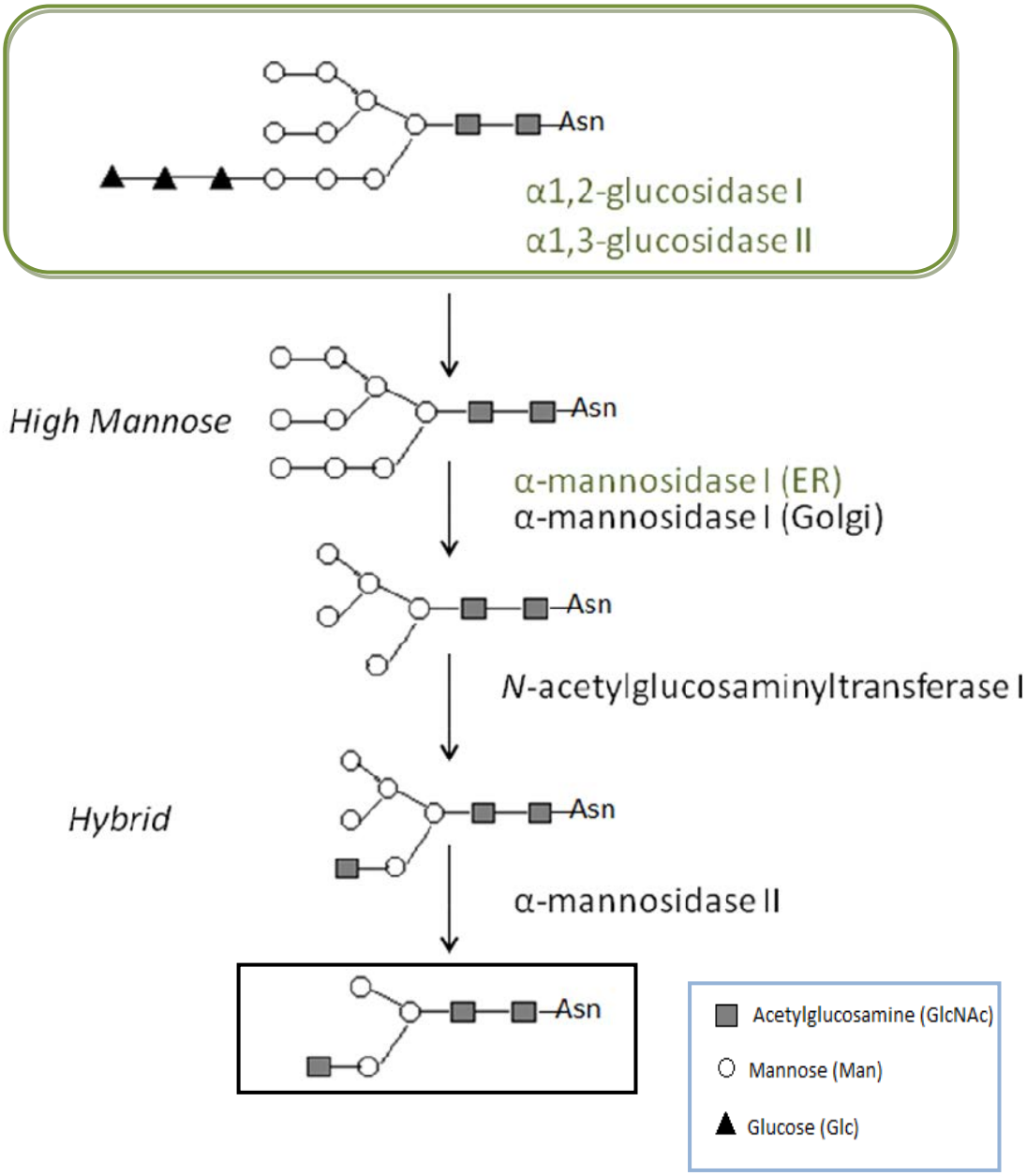

Figure 1.6 Initial steps of *N*-glycoside trimming process in eukaryotic cells that are common to insect and mammalian cells.
The green figure represents the ER. The items inside the green figure correspond to the events that happen in the ER. The structures and the enzymes outside the green figure correspond to those found in the golgi apparatus. The glycan inside the black square is $GlcNAcMan_3GlcNAc_2$, the last common intermediate for both types of cells. (Modified from Shi et al., 2007)

N-glycan processing starts via a series of trimming reactions by the action of α-glucosidase I, which removes the terminal 1,2-linked glucose from the preassembled moiety. This reaction is followed by a glucosidase II catalyzed removal of the two 1,3-linked glucose units to give the '*high mannose'* $Man_9GlcNAc_2$ glycan. The subsequent trimming reactions involve the removal of one mannose unit from the previous glycans followed by the cleavage of all the remaining 1,2-linked mannose residues, by the action of the class II and I α-mannosidases which are located in the ER and in the Golgi apparatus respectively [68]. The end product of these reactions is a $Man_5GlcNAc_2$ complex, which by the addition of another GlcNAc via *N*-acetylglucosaminyltransferase I it becomes the "*hybrid"* $GlcNAcMan_5GlcNAc_2$. This hybrid intermediate leads to the formation of a common substrate for *N*-glycosylation for insect and mammalian cells generating $GlcNAcMan_3GlcNAc_2$, after the removal of the two more mannose residues by α-mannosidase II. The mammalian and insect cell pathway have in common a core α-1,6-fucosyltransferase that adds a 1,6-fucose residue onto the GlcNAc attached to the Asn from the previous structure. Interestingly, this addition does not happen to all of the $GlcNAcMan3GlcNAc_2$ structures. A GlcNAc residue can be removed from this insect-mammalian common intermediate substrate by an *N*-acetylglucosaminidase as in the case of insect cells, or it can be added to the structure by the action of an *N*-acetylglucosaminyltransferase II in mammalian cells. A core α-1,3-fucosyltransferase recognizes the 1,6 fucosylated structure, and then adds another fucose to the corresponding position of the GlcNAc. The resultant glycan cannot be produced or used in mammalian cells, since it is allergenic for mammals [69, 70]. This also constitutes a major difference between the mammalian and insect glycosylation pathways, because the following trimming reaction only applies to the insect cell pathway, generating paucimannosidic *N*-glycans as final the products (Figure 1.6).

Insect cells have been used to recombinantly express HE glycoproteins that originate from various viruses [71, 72] including ISA virus [43]. The most common system used in these insect cell expressions is the 'Baculovirus expression system'.

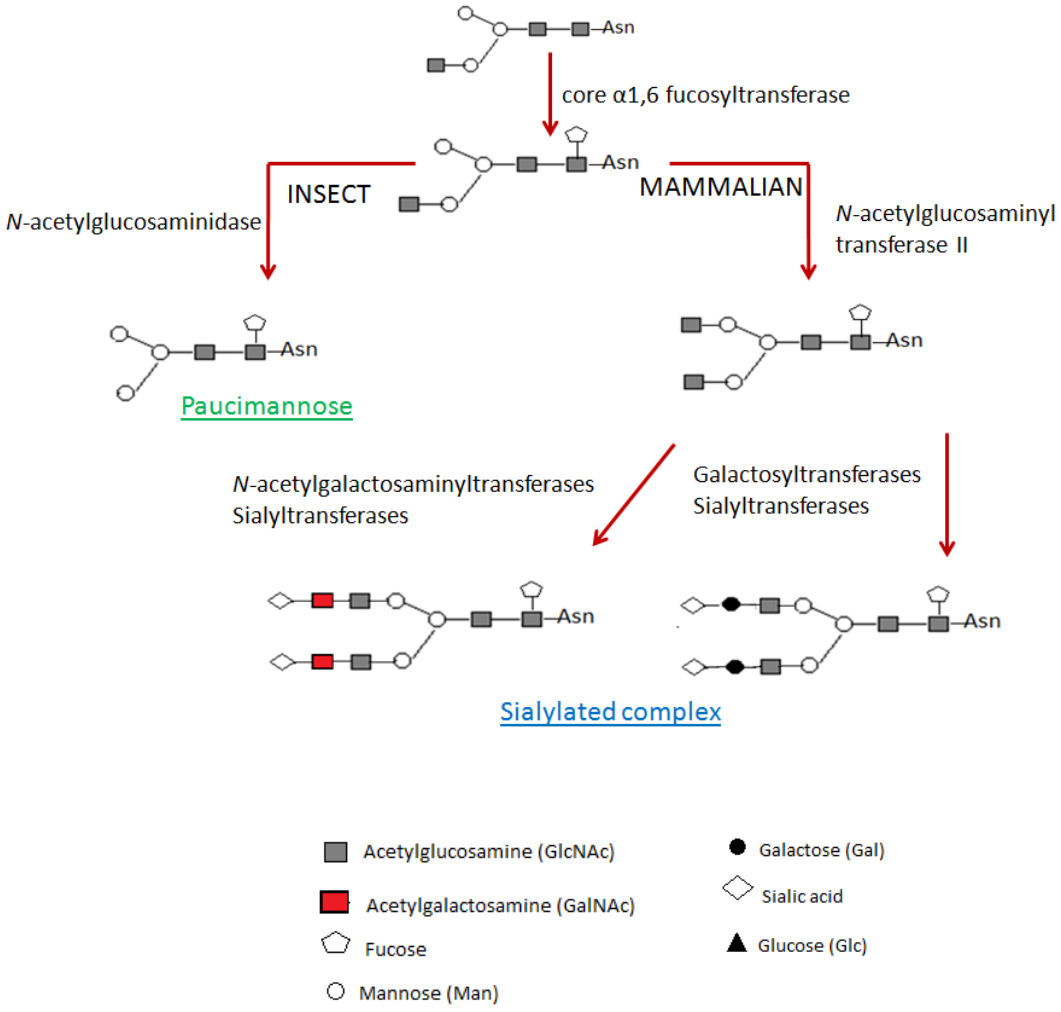

Figure 1.7 Different products of *N*-Glycosylation in insect and mammalian cells. The *N*-glycosylation pathway differs between insect and mammalian cells. The last common substrate from the previous trimming reactions undergoes different enzymatic modifications. In insect cells, the synthesis of the final glycan paucimannose is obtained in one simple step, whereas in mammalian cells, an extra step is required in order to obtain sialylated complexes. (Modified from Shi et al., 2007)

1.5. Baculovirus Expression system.

Several papers have reported the expression of the hemagglutinin-esterase protein in insect cell lines from the Lepidoptera order, such as *Trichoplusia ni*, also called HighFive, and *Spodoptera frugiperda* or *Sf9* cells using a system called Baculovirus [43, 73].

A particular group of viruses called baculoviruses possess double stranded, circular DNA. Their genomes, often found as rod-shaped nucleocapsids, range from 80 to 108 kb in size and they encode for between 90 and 180 genes. Usually they can be found as two types, occluded virions (OVs) and budded virions (BVs). OVs are covered by a clear matrix forming occlusion bodies (OBs) which are infectious if ingested orally, whereas the BVs are secreted enveloped with a nucleocapsid when the infected cell bursts [74].

Based on occlusion body morphology, baculoviruses can be divided into two genera: nucleopolyhedriviruses (NPVs) and granuloviruses (GVs). NPVs may contain either single (SNPV) or multiple (MNPV) nucleocapsids per envelope, while the latter only has one per envelope [75]. The sizes of the NPVs occlusion bodies range from 0.6 to 2 µm and they are called polyhedral, hence their major occlusion protein is called polyhedrin [74], which is a structural protein found in the matrix of OBs where the virions are located [76].

In nature, these pathogens infect insects that are often qualified as plagues such as the velvet bean caterpillar *Anticarsia gemmatalis* [77], codling moth *Cydia pomonella* [78] found in apples, and the cotton bollworm *Helicoverpa armigera* [79]. As a result, these viruses can be very useful for pest control. In biotechnology, they have been used as vectors for foreign gene expression for the production of recombinant proteins.

Advantages of this system include that it is: safe to humans since the virus is not pathogenic to mammals, easy to use and to scale up, can give high levels of expressed recombinant protein, which can be 250 times greater than the corresponding

mammalian cell systems [80]. For this purpose, the most common baculoviruses utilized as vectors for recombinant protein expression in insect cells are the *Autographa californica* multiple nucleopolyhedrovirus (*Ac*MNPV), and the *Bombyx mori* nucleopolyhedrovirus (*Bm*NPV).

1.5.1. Insect cells, Baculoviruses and Transfection Vectors

In the baculovirus-insect cell system, the virus acts as the vector carrying the gene of a recombinant protein and the latter is the host, in which the protein will be expressed. The resulting protein should have the appropriate post-translational modification.

Insect cells.

As was reviewed previously, protein *N*-glycosylation in insect cells deviates from the main mammalian pathway to generate different final glycans. The main difference between them is that in insect cells these products are not terminally sialylated. This lack of sialylation could be the reason why the baculoviruses are highly specialized to take advantage of the insect cell machinery, and therefore are widely used for the production of recombinant proteins [56].

The host range of baculoviruses are invertebrates, mainly insects [81]. Insect cells are the most common cell line used for this purposes, specifically from the Lepidoptera order. These include cells from the larvae stage of the Armyworm moth *Spodoptera frugiperda or Sf21* [82]*;* a substrain from the same cells called *Sf9* [83] and *Trichoplusia ni* or High Five, from the cabbage moth [84, 85]. Due to the differences in glycosylation, transgenic insect cell lines derived from *Sf9* and HighFive were developed. These mutant cell lines include genes encoding enzymes that are required to perform mammalian glycosylation, such as bovine β-1,4-galactosyl transferase [86] and rat α-2,6-sialyltransferase [87]. Consequently, sialylated products can be obtained. The reason why insect cells are used in the baculovirus system can be explained because they can be easily propagated in large quantities and they are considered to be biosafety level 1. In addition, they are the baculoviruses natural host cells.

Baculovirus DNA

Baculovirus gene expression in its host, most commonly in insect cells, is carried out in four different stages: Immediate-early, delayed-early, late and very late [88]. Immediately after infection, a host-endogenous RNA polymerase II transcribes immediate-early genes. These genes activate delayed-early and late genes which consequently are transcribed in the delayed-early phase. The encoded proteins after this transcription, such as DNA polymerase and helicase, are necessary for DNA replication and the expression of the genes from the late stage [89, 90]. Proteins involved in the virion assembly and budding are produced in the late stage, as well as DNA replication. Lastly, two important proteins are created in large quantities: polyhedrin and p10. As previously mentioned, polyhedrin is part of the structure of the OBs. The p10 protein is involved in the release of the OBs from the nucleus of the cells, in order to continue with the infection [91]. These proteins are not required for the production of baculovirus in the form that will generate infection in insect cells; however their genes can be used as promoters to express a foreign protein. This can be accomplished by inserting a transfer vector containing these promoters and the recombinant protein into the viral DNA.

Transfer vectors.

Since the size of the baculovirus genome is too large (about 130 kb) [92] to incorporate foreign genes into it easily, the gene is inserted into bacterial plasmids or “bacmids”, which can be replicated in bacteria. These are the transfer vectors or baculovirus expression vectors (BEVs), and they often possess a lac Z gene.

The deletion of specific genes in BEVs allows the expression of foreign genes. Generally, BEVs utilizes the polyhedrin promoter from the *Autographa californica* multiple nucleopolyhedrovirus (*Ac*MNPV) which has also been modified for foreign gene insertion [92]. Since this promoter is viral in nature it needs critical viral factors for gene transcription. The first step towards gene transcription and expression is the transfection of the BEV and the baculovirus DNA into the insect cells.

Once the transfer vector and the baculovirus DNA are inserted into the insect cells they exchange nucleic sequences via a process called homologous recombination. This causes allelic replacement, which incorporates the recombinant gene carried by the vector into the baculovirus genome [74], thus aiding the process of

gene expression. This homologous recombination method, however, produces a low yield of recombinant viruses (or occlusion negative viruses (*occ*$^{-}$)) inside the host cells due to a low recombination frequency (approximately 0.1%) [93]. In addition to this, a mixture of parental (occlusion positive viruses (*occ*$^{+}$)) and the recombinant viruses are also created. In order to identify *occ*$^{-}$ from a pool of *occ*+, plaque purification assays was needed, a tedious system that required microscopy skills and good quality microscopes. The development of a strategy that involved the linearization of the baculovirus DNA was the solution to this problem.

Within the polyhedrin locus, there is a unique *Bsu36* I restriction site [94]. When the baculovirus DNA is cut at this restriction site, the polyhedrin promoter becomes inactive. The transfer vector must contain the same promoter sequence, therefore when combined; the recombination process will generate a circular viral DNA with an active promoter. This system increases the recombination frequency up to 25% and allows the production of recombinant viruses only (Figure 1.7). When having more *Bsu36* I restriction sites within the viral DNA, the recombination frequency can improve up to 90% [95].

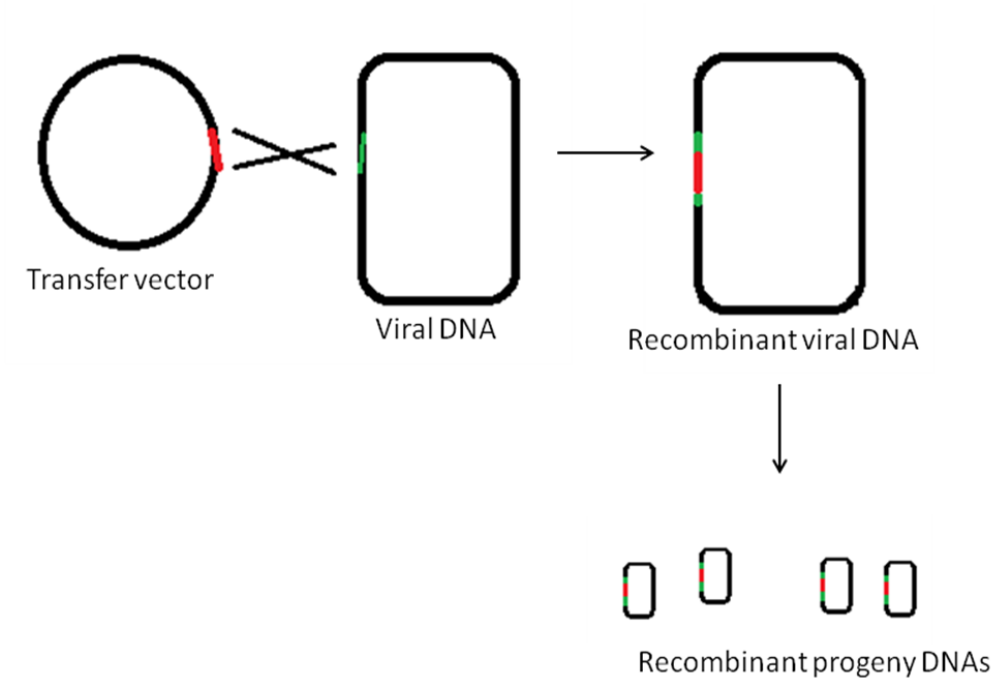

Figure 1.8 Schematic representation of the homologous recombination between linearized baculovirus DNA and the transfer vector.
The transfer vector carrying the gene of interest undergoes homologous recombination with the baculoviral DNA to produce recombinant viral DNA. This modified DNA will be the recombinant viral progeny when the budded virions are released from the infected cell. The progeny will carry the gene of interest in their DNA. (Modified from Kitts et al., 1990)

Furthermore, this additional gene has been shown to be required for viral propagation [96] and it encodes a structural protein [94] called ORF1629. In *Ac*MNPV, a double digestion with Bsu36 I removes a gene coding sequence such as *β*-galactosidase and also a part of ORF1629 making it unable to form infectious viral particles. The homologous recombination process will restore the deletion of the ORF1629 and also will insert the foreign gene that will be expressed instead of the *β*-galactosidase. In addition, another gene called ORF603 encodes a protein not essential for viral replication and that a foreign gene can be expressed under the control of its promoter [97]. This gene is also utilized with the same objective as ORF1629.

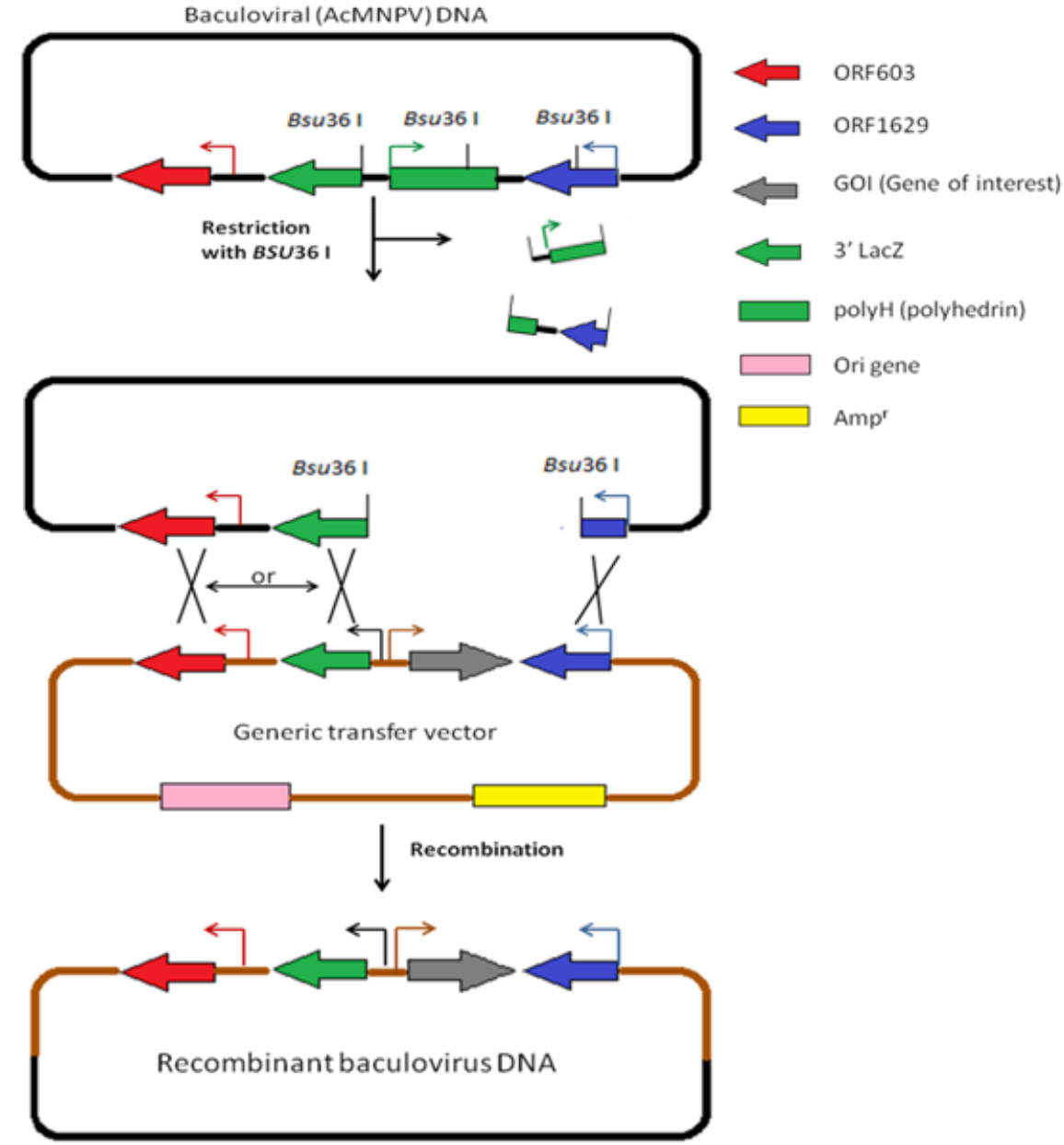

Figure 1.9 Simplified example of the homologous recombination process between the baculoviral DNA and a transfer vector from the BacNBlue manual.
After enzymatic restriction with *Bsu*36 I, parts of the cut genes from the baculoviral DNA are replaced with genes from a transfer vector carrying the gene of interest. This is done via homologous recombination. As a result, a recombinant baculovirus DNA with the gene of interest incorporated is generated. The coloured arrows show the direction where genes with corresponding colours are transcribed. The viral polyhedrin promoter is replaced with the polyhedrin located prior to the Gene Of Interest (GOI) from the transfer vector and it is represented here as a brown arrow. (Modified from BacNBlue transfection kit manual, Invitrogen)

Some baculovirus genomes, which are commercially available such as BaculoGold (Pharmingen), Bac-to-Bac and BacNBlue (Life Technologies), have been engineered to incorporate the previously described strategies and methods. Figure 1.8 shows an example of how homologous recombination should take place inside the insect cells once the baculovirus DNA and the BEV are transfected. This diagram was modified and simplified from the BacNBlue kit manual, which was followed during in the experiments that I performed during the course of this research.

1.6. Objectives of this thesis

The main objective of this thesis was to find chemical compounds that can reduce ISAv infectivity, by inhibiting its hemagglutinin-esterase (HE) protein. This could be accomplished by achieving the following specific objectives:

- To clone the HE protein from a Canadian strain (Back Bay 810/9/99) and a Norwegian strain (Glesvaer2/90) of ISAv.
- To express the HE proteins recombinantly using a Baculovirus Expression Vector System (BEVS).
- To develop a specific activity assay for the esterase from HE protein.
- To assess kinetically inhibition of the esterase activity of HE using various 4-substituted sialosides.

Chapter 2. MATERIALS AND METHODS

2.1. Subcloning and modifications of hemagglutinin-esterase (HE)

2.1.1. Origin of hemagglutinin-esterase from the viral strains Back Bay 98 (Bay 810/9/99) and Glesvaer 2/90.

Hemagglutinin-esterase (HE) from viral strain Back Bay 98 (Bay 810/9/99).

The cDNA containing the full length of the high polymorphic region (HPR0) of hemagglutinin-esterase was obtained from the Canadian viral isolate Back Bay 810/9/99, New Brunswick, and it was kindly donated by Dr. Nellie Gagne (Fisheries and Oceans Canada), who had already cloned it into a pET 101/D-TOPO vector. This construct was named *pET101_HE*. The orientation of the gene in the plasmid allows for expression of the T7 promoter, and therefore replication of the gene in *E. coli*.

Hemagglutinin-esterase (HE) from viral strain Glesvaer 2/90. The cDNA of the HPR0 of hemagglutinin-esterase protein from the Norwegian viral isolate Glesvaer 2/90, which had been cloned in a pEGFP vector, was kindly donated by Anita Muller and Turhan Markussen (National Veterinarian Institute, Oslo, Norway). This construct was named *pEGFP_HE_Gles2/90*. The orientation of the gene in the plasmid allows expression of the T7 promoter, and therefore replication of the gene in *E. coli*. In order to perform western blot analysis, a rabbit anti-HE immune peptide antiserum for this specific strain was obtained from Dr. Børre Robertsen from the Fish Immunology and Vaccinology Research Group, Norwegian College of Fishery Science, University of Tromsø, Norway. The antiserum was prepared as described in Krossøy et al. [42].

2.1.2. Subcloning of Hemagglutinin-esterase (HE) into a pVL1392 vector to generate *pVC_HE6* construct

Source of reagents. The baculovirus expression vector pVL1392 was kindly provided by Dr. Rosemary Cornell from the Molecular Biology and Biochemistry Department at Simon Fraser University, as well as the Baculovirus DNA BaculoGold (Pharmingen) and the insect cell lines *Spodoptera frugiperda* (*Sf9*) and *Trichoplusia ni (T. Ni or HighFive).* The mammalian Chinese Hamster Ovarian (CHO) cells were provided by Dr. Wesley Zandberg, a Postdoctoral fellow from Dr. David Vocadlo's laboratory, at Simon Fraser University. All of these materials were required for the expression of the HE dual function protein, which is described in this chapter.

Construction of pVC_HE6. Figure 2.1 shows the schematic construction of *pVC_HE6*. The cDNA of HE from *pET101_HE* was amplified by PCR using the whole construct as the DNA template, 0.2 mM of dNTPs, ~0.1 U/µL of *Pfu* polymerase (Thermo Scientific), 1 x of *Pfu* buffer and 0.5 mM of the primer set HE_F and HE_R (Table 1).The primers were design with the appropriate restriction sites in order to facilitate subcloning into the pVL1392 vector. PCR conditions were as follows: 95 °C for 30 seconds, 65 °C for 45 seconds, and 72 °C for 45 seconds. This was repeated for 35 cycles. The extension step was carried out at 72 °C for 10 minutes. The PCR product was concentrated using a speed-vac and approximately 0.5 µg was digested with 1 U of both *Bam*HI and *Pst*I using 1 x of the appropriate buffer in 10 µL reaction volume.

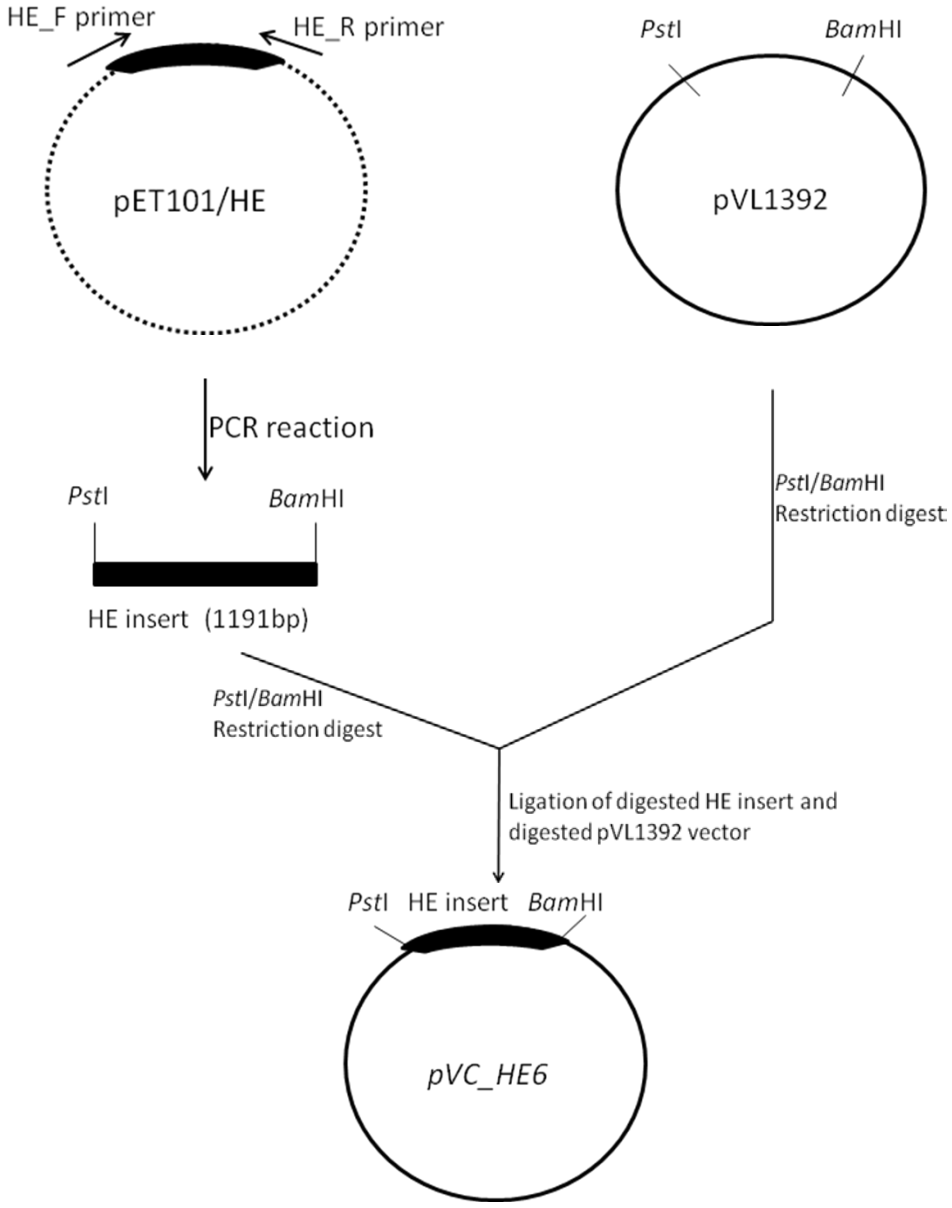

Figure 2.1 Preparation of *pVC_HE6* construct.

DNA purification from agarose gels and ligation of BamHI/PstI digested pVL1392 and PCR product of Hemagglutinin-esterase (HE). The digested DNA was loaded into a 1% agarose gel containing Sybr safe DNA gel stain (Life technologies) and run at 50 V for approximately 1 hour. The single band with the expected size of approximately 1.1 Kb, which was visualized with a green light box, was excised with a scalpel and purified from the agarose using a QIAquick Gel Extraction Kit from QIAGEN. The purified DNA was concentrated using a speed-vac for 25 minutes at 45 °C, and resuspended in 10 µL of water. Determination of the concentrations of DNA was made using a nanodrop spectrometer. The vector and the gene, both previously digested with *Bam*HI and *Pst*I restriction enzymes, were ligated in a 3.5:1 (insert:vector) ratio, with 1 U of T4 DNA ligase (Thermo Scientific) and 1 x of the appropriate buffer in a 15 µL reaction volume. The ligation mix was incubated at room temperature for 1 hour followed by a second incubation of 5 minutes at 70 °C to inactivate the enzyme. The transformation procedure was performed as described above for the pET101_HE construct. Approximately 45 ng of DNA from the ligation mix was added to ~100 µL of BL21Gold(DE3) *E. coli* competent cells. The mixture was incubated on ice for 30 min, heat-shocked at 42 °C for 25 seconds, and then placed back on ice for 2 minutes. After these treatments, bacterial growth and colony selection were performed as described in the '*Transformation and proliferation of pET101_HE plasmid*' section on page 40. The construct obtained at this step was named *pVC_HE6.*

Table 1. Primers used in the PCR reactions for cloning procedures and transmembrane deletions in the pVC_HE7 and pVC_HEGles2/90 constructs.

Construct name	Primer name	Primer sequence (5'-->3')
pVC_HE6	HE_F	GGG *CTG CAG* ATG GCA CGA TTC ATA
	HE_R	GGG *GGA TCC* TTA TGC AGG CCC AGC
pVC_HE7	R_HE_TMdel	CCC CCC *GGA TCC* TTA GTG GTG GTG GTG GTG GTG CAT AGA GAT GAA GAT GTT
pVC_HEGles2/90	F_HEGles2/90	GGG *GAA TTC* ATG GCA CGA TTC ATA ATT TT
	R_HEGles2/90	GGG *CCC GGG* TTA GTG GTG GTG GTG GTG GTG ACC CAT AGT TTG GTT CAG

Underlined bases: Base pairs added to ensure primer annealing with template sequence. Italics bases: Restriction sites. *Pst*I (blue), *Bam*HI (red), *EcoR*I (purple), *Sma*I (uncoloured). Coloured bases: Start codon (green), stop codon (black), Histag (orange).

2.1.3. Subcloning of HE into a pVL1392 vector to generate *pVC_HE7* construct

This construct was created to delete the coding region for a transmembrane domain present in the cDNA of hemagglutinin-esterase. In addition, the specific nucleotides to generate a 6x Histag, in order to facilitate purification, and a stop codon were added to the 3' end of the DNA.

Deletion of the coding region for the transmembrane domain (TMD) from the cDNA of HE from pVC_HE7. A transmembrane region (amino acids 356-378), which was identified as such by the online tool SMART (Simple Modular Architecture Research Tool) program, was deleted from the *pVC_HE6* construct by changing the DNA primers HE_F and R_HE_TMdel (Table 1) and by using the same PCR conditions as described above, using *pET101_HE* as DNA template. In addition, a 6 x Histag was added to the C-terminus of the protein, to facilitate purification and this was followed by a stop codon (Fig 2.2). The procedure was carried out in a similar manner as described in the 2.2 '*Construction of pVC_HE6*' section. This addition of the nucleotides was included in the design of the reverse primer. After sequence analysis, clone number seven was used for further experiments; therefore the final construct was named *pVC_HE7*.

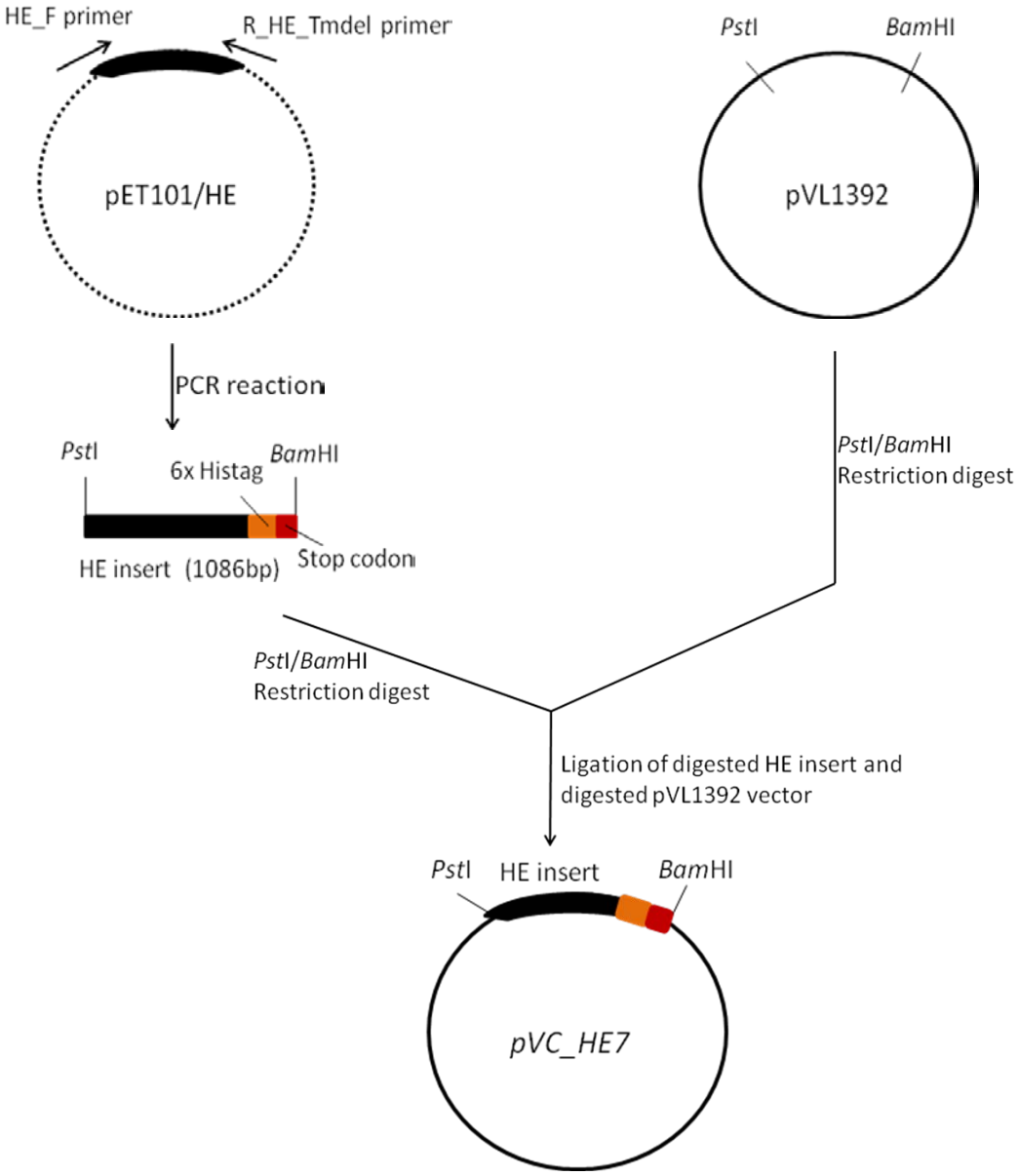

Figure 2.2 Preparation of the *pVC_HE7* construct.

2.1.4. Subcloning of HE into a pVL1392 vector to generate *pVC_HEGles2/90* construct

This construct was made to be used for expression trials with Chinese Hamster Ovary (CHO) cells. The preparation of this construct is schematically shown in Figure 2.3.

The cDNA of HE from the *pEGFP_HE_Gles2/90* construct was amplified by PCR using the whole construct as the DNA template with the primer set F_HEGles2/90 and R_HEGles2/90 (Table 1). These primers were designed to delete the TMD region and add a 6x Histag to the cDNA of HE. The PCR and the digestion conditions used were the same as those for the *pVC_HE6* construct, with the exception that the restriction enzymes used in this case were 1 unit of both *EcoR*I and *Sma*I.

DNA purification from the agarose gels and ligation of EcoRI/SmaI digested pVL1392 and PCR product of *Hemagglutinin-esterase (HE)* The digested DNA was loaded onto a 1% agarose gel containing Sybr safe DNA gel stain (Life technologies) and run at 50 V for approximately 1 hour. The single bands with the expected size (~1.1 Kb), which were visualized using a green light box, were excised with a scalpel and purified from the agarose using a GeneJET Gel Extraction Kit (Thermo Scientific). The extracted DNA was concentrated using a speed-vac for 25 minutes at 45 °C, and resuspended in 10 µL of water. Determination of the concentrations of DNA was made using a nanodrop spectrometer. The vector and the gene were ligated in a 4:1 (insert:vector) ratio, with 1 U of T4 DNA ligase and 1 x of the appropriate buffer in a 10 µL reaction volume. The ligation mixture was incubated at 22 °C for 1 hour followed by a second incubation of 5 minutes at 70 °C, which inactivated the ligase. Transformation of the construct was performed in a similar manner to the pET101_HE construct, as described above. Approximately 40 ng of DNA from the ligation mix was added to ~100 µL of 5-alpha *E. coli* competent cells (New England Biolabs). The mixture was incubated on ice for 30 minutes, heat-shocked at 42 °C for 30 seconds, and then incubated on ice for another 5 minutes. After these incubations, the procedure remained the same as described in the '*Transformation and proliferation of pET101_HE plasmid*' section on page 40.

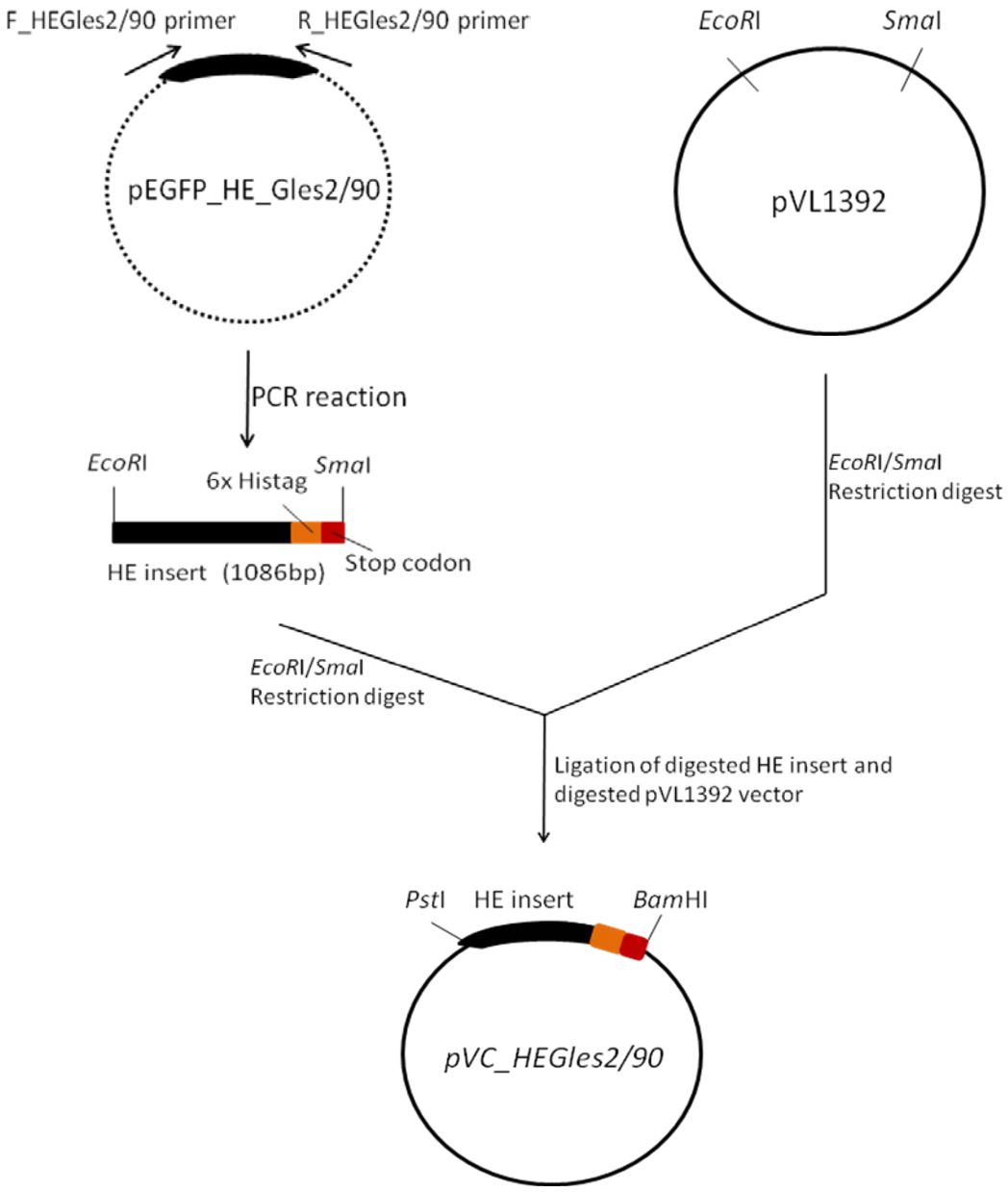

Figure 2.3 Preparation of the *pVC_HEGles2/90* construct.

2.2. Cell culture and propagation.

Growth and propagation of Spodoptera frugiperda (Sf9) cell line. Approximately 1 x 10^7 cells were thawed from liquid nitrogen storage and 1 mL of the cell stock was added drop wise into a T-75 flask containing 20 mL of previously warmed (in a 28 °C water bath) Sf900 II Insect Cell Culture Media (Life Technologies) that had been supplemented with 10% of heat inactivated fetal bovine serum (FBS) (Life Technologies). The flask containing the cells was incubated horizontally at 27 °C in a sealed container until monolayer confluency reached ~70-80% by doing 1:3 and/or 1:6 passages every two days. Once this percentage was achieved, several passages were made in order to keep the cells dividing. Each passage was made with 20 mL of the same media supplemented with 5% FBS and everyone was discontinued after reaching the required cell density for the planned experiment.

Growth and propagation of Trichoplusia ni (T. ni or HighFive) cells. The growth of *T. ni* cells was performed in a similar manner to that detailed above for *Sf9* cells, with a few modifications. These insect cells were grown in Excell 405 Serum-free media for insect cells (Sigma-Aldrich) that was supplemented with 5% of heat inactivated FBS (Life Technologies) which had been previously warmed to 28 °C in a water bath. All passages were made with 20 mL of the same media supplemented with 2% FBS and discontinued after reaching the required cell density for the designed experiment.

2.3. Expression of HE using Baculovirus Systems in Insect Cells.

2.3.1. Baculovirus expression using BaculoGold DNA.

Cotransfection of insect cells (T. ni) with the plasmid and BaculoGold viral DNA.

Step 1. Transfection with BaculoGold DNA and transfer vector (i.e.: pVC_HE7). Cells from the starting cultures described above were counted using a hematocytometer to give an estimated cell count per mL of culture medium, and these were transferred into two 6 cm petri dishes in order to yield 2 x 10^6 cells in each petri dish. Both dishes, one used as a negative control and the other as the reaction (infection) plate, were incubated at 27 °C for at least 1 hour or until a 50% confluency was achieved. In parallel, transfection buffer A (Excell 405 media plus 10% FBS) and transfection buffer B (25 mM HEPES, 125 mM CaCl2, 140 mM NaCl, pH 7.0) were prepared. A DNA mixture was prepared as follows: In a sterile 1.5 mL tube, 0.5 µg of the viral BaculoGold DNA was mixed with 5 µg of transfer vector *pVC_HE7* and this mixture was incubated for 5 minutes at room temperature. Then, transfection buffer B (1 mL) was added to the DNA mix.

After the 50% cell confluency had been achieved in both petri dishes, the media was removed from the cells by use of an aspirator. For one plate, the negative control, 1 mL of Excell media was added. For the second plate, the infection, 1 mL of transfection buffer A was added to the cells, followed by a drop wise addition of the DNA mixture. Both dishes were incubated at 27 °C for 4 hours. Following this incubation, the media containing the DNA mixture was aspirated from the plates containing the transfected cells. These cells were washed with 3 mL of fresh media previously warmed in a 28 °C water bath for about 30 seconds. After the media was discarded via aspiration, 4 mL of serum free media was added to the cells, which were then incubated at 27 °C for 4-5 days.

Step 2. Harvest of media (‘*co-transfection sup’*). The media in which the cells were cultured was used to break the monolayer by flushing the cells from the plate

using a 5 mL pipette, in order to obtain a cell/media mixture. This solution was transferred into 15 mL sterile tubes, and centrifuged at 1100 rpm for 5 minutes. The resultant supernatant here was called the 'co-transfection 'sup'. This 'sup' was transferred into a new clean 15 mL tube, wrapped in aluminum foil and stored in the dark at 4 °C, until required for the next step. The cell pellet was discarded at this stage.

Amplification of co-transfection sup to obtain viral stock (AMP1).

Step 1. Infection with the 'co-transfection sup'. All reactions were carried out on 10 cm petri dishes. Insect cells were plated with 2% FBS supplemented Excell media at a cell number of 5 x 10^6 into two dishes as described above, and incubated for 1 hour or until ~50% confluency was achieved. The media was then aspirated from the cells and 10 mL of fresh Excell (without FBS) was added. One plate was untreated (uninfected) to be used as the negative control. For the infection plate, a 5 mL mixture of diluted soup (0.5 mL of sup + 5 mL of media) was added. Both plates were incubated at 27 °C for 3 days. Within 2-3 days, signs of infection might be visible by an inspection using a standard microscope. If the cells are infected by the baculovirus they show an increase in diameter, change in their morphology, and some may float free in the media.

Step 2. Harvest of cells / media (Amp1). By using a 10 mL pipette, the cells were washed down with the same culture media, transferred to clean 15 mL tubes and spun down at 2000 rpm for 5 minutes. The supernatant, called Amp1, was transferred into a new 15 mL tube, wrapped in aluminum foil and kept in the dark at 4 °C until further use. A 1-2 mL aliquot was also taken and stored at -80 °C. The cell pellet was washed with 5 mL of cold PBS and spun down at 2000 rpm for 5 minutes. The supernatant was discarded and the pellet was wrapped in aluminum foil, and kept at -80 °C until further use.

Amplification of High Titer Virus (HTV)

Step 1. Infection of *T. ni* cells with Amp1. This step was performed in either multiple 10 cm petri dishes or 2-3 plates of 15 cm each, with 10 or 25 mL of culture media, respectively. Cells were plated in the same manner as described above. In parallel, 100-200 µL of Amp1 was diluted into 10 mL of serum free media per plate. Once the cell confluency was ~50%, the media was extracted from the cells by

aspiration and replaced with the diluted Amp1 media. Cells were incubated for 3-4 days at 20 °C.

Step 2. Harvest of media containing High Titre Virus (HTV). Cells were washed down with a 10 mL pipette with the same culture media and transferred to 15 or 50 mL tubes, depending on the volume of the culture media. The cell/media culture was spun down at 2000 rpm for 5 minutes and the supernatant at this step is called High Titer Virus (HTV). The HTV solution was wrapped in aluminum foil and stored at 4 °C for future protein expressions analysis. The pellet was treated as described in step 2 from Amp1, and kept for protein analysis.

2.3.2. Baculovirus expression using BacNBlue DNA.

Cotransfection of insect cells (T. ni and Sf9) with plasmid and BacNBlue viral DNA.

With a few modifications, the process of transfection was done in a similar manner to that reported in the previous section when BaculoGold DNA was used.

Step 1. Transfection with BacNBlue DNA and transfer vector (i.e. pVC_HE7 and *pVC_HEGles2/90).* Cells were transfected once they reached ~90% confluency and grown in Excell 405 and Sf900II media without FBS for *T. ni* and *Sf9* cells, respectively. In 1.5 mL tubes, a DNA mixture containing the 1 µg/µL of transfer vectors *pVC_HE7* and *pVC_HEGles2/90*, and 250 µL of the corresponding medium for each cell line (without FBS) was prepared. In a separate tube, 30 µL of Cellfectin II Reagent was diluted into 250 µL of media, also without supplements and incubated at room temperature for 15-30 minutes. The ‘transfection mixture’ was prepared by combining the diluted DNA with the cellfectin mix, and the resultant solution was incubated at room temperature for 15 minutes. The transfection mix was added drop wise to the confluent cells and the negative control plate was left untreated (uninfected). Transfected cells and negative controls were incubated at 27 °C for 4 hours. Following this incubation period, the cell culture media was removed by aspiration and replaced with fresh media (no FBS added), and then incubated again at 27 °C for 3 days.

Step 2. Harvest of media (P1). The cell culture media was extracted with a 5 mL pipette and stored at 4 °C in the dark. This is a low scale viral stock or 'P1', which was used to initiate a second round of infection.

Amplification of P1 to obtain medium scale viral stock (P2).

Step 1. Infection of insect cells with P1. *T. ni* and *Sf9* cells were seeded into 10 cm plates with the respective media in the presence of FBS (2% for *T. ni* and 5% for *Sf9*). Once the confluency was approximately 80%, the media was aspirated and exchanged with 10 mL of fresh media plus 10% FBS. Twenty microlitres of P1 was added to the cells and then incubated at 27 °C for 5-10 days, or until the cells were observed to be 90-100% lysed.

Step 2. Harvest of media (P2). The infected insect cells were washed using a 10 mL pipette and spun down at 2000 rpm for 5 minutes. The resulting supernatant was called 'P_2'. One millilitre was transferred and kept in the dark at -80 °C for long term storage and another 4 mL were kept at 4 °C also in the dark as a reserve stock. The remainder 5 mL of P2 was used in the production of 'High Titre Virus' stock and the cell pellet was washed with cold PBS, spun down and kept in the dark at 4 °C.

Amplification of High Titer Virus (HTV)

Step 1. Infection with P2. Cells were grown in multiple 10 cm plates or in 2-3 plates of 15 cm using the respective media with FBS, as described above. Once they reached a confluency of approximately 80%, 5 mL of the P2 stock generated at the previous step was added drop wise after fresh media plus 10% FBS was exchanged. Negative control plates were left untreated. From this step of the infection, cells were incubated at 20 °C for 5-7 days.

Step 2. Harvest of media containing High Titer Virus (HTV). When it was observed using a microscope that the cells were approximately 90-100% lysed they were harvested using a 10 mL pipette for transfer to either a 15 or 50 mL tube. These samples were spun down at 2000 rpm for 5 minutes and the supernatant (HTV) was collected and stored in the dark at 4 °C in order to analyze protein expression levels. The cell pellet was also stored in this manner and kept for protein analysis.

HTV samples preparation for analysis of HE expression

Cell media. Cell debris for all HTV samples from the expressions previously described was removed from media by differential centrifugation. After each HTV harvest, 0.5 µg/mL of leupeptin (Sigma) was added to the samples and then spun down at 9000 rpm for 25 minutes, followed by a second centrifugation at 11000 rpm for another 25 minutes at 4 °C. The supernatant was filtered through a 0.45 µm low protein binding filter and then concentrated using a 30 kDa cut off filter Amicon tubes (Milipore) by doing a series of spins for 25 minutes at 3500 rpm at 4 °C. Finally, the concentrated medium was exchanged with 20 mM HEPES, 500 mM NaCl, pH 7.5 using the same approach as for concentration. Ten percent glycerol was added to the final solutions, and stored at -20°C.

Cell pellet. In order to perform protein analysis on the cell pellet, it was resuspended in lysis buffer containing 10 mM Tris, 130 mM NaCl, 1 % Triton X-100, 10 mM NaF (to inhibit kinases),10 mM NaPPi, 10 mM NaPi, (to inhibit phosphatases), pH 7.5. Following sonication at 15–50% amplitude for 2 minutes, 20 seconds on, 40 seconds off, the lysed cells were spun down at 14000 rpm for 10 minutes at 4 °C and the supernatant was concentrated and the buffer exchanged as described above.

2.4. Expression of HE using a transfection system in mammalian cells

This expression was performed by following the instructions that came with the Lipotransfectin kit (Life Technologies).

Step 1. Transfection of CHO cells with lipofectamin. The Chinese hamster ovary (CHO) cells were grown in two 10 cm plates in a 37 °C incubator in the presence of CO_2. Once the cells were 100% confluent, the media was exchanged by aspiration with 5 mL of fresh DMEM media without FBS. A set of reactions were first prepared in 1.5 mL tubes: Control 1 and Transfection 1, both containing 1 mL of OptiMEM reagent plus 30 µL of lipofectamin. A second set was also prepared in 1.5 mL tubes: Control 2 and Transfection 2. Control 2 contains 1 mL of OptiMEM , 6.4 µL of P3000 reagent, while

Transfection 2 has 1 mL of OptiMEM, 5.4 µL of P3000 reagent and 1 µL (2.7 µg/µL) of the unmodified *pEGFP_HE_Gles2/90* construct. Control 1 and 2 were mixed together, as were transfection 1 and 2, and these two mixtures were incubated at room temperature for 5 minutes. Part of the resulting mixtures (2 mL) were added drop wise to the cells and incubated at 37 °C for 3-4 days.

Step 2. Harvest of transfected cells. After the incubation period, the cell culture media was removed by aspiration and the cells were washed with 10 mL of cold PBS which was also aspirated. One millilitre of Trypsin/EDTA solution (Gibco) mixed with 3-4 mL of cold PBS was added to the cells then incubated at 37 °C for 1 minute, in order to detach the cell monolayer from the plate. The resulting cells were transferred into a 15 mL tube with a 10 mL pipette and spun down at 2000 rpm for 5 minutes. The supernatant was discarded and the cell pellet was carefully transferred into clean 1.5 mL tubes and then resuspended in 200 µL of a solution containing 1 x PBS and 1% SDS. The cells were lysed using a sonicator at 50% amplitude for 2 minutes, 20 seconds on and 40 seconds off, then spun down at 14000 rpm for 10 minutes at 4 °C. The supernatant was analyzed by SDS electrophoresis and western blotting.

2.5. Enzyme activity assays

Activity assays were carried out in a Cary 300 UV spectrophotometer (Agilent) equipped with a Peltier temperature controller. Crude extracts from the various expression experiments were used with various concentrations of *p*-nitrophenyl acetate (*p*-NPA) (Sigma) as substrate. As this substrate readily hydrolyzes in water, stock solutions of *p*-NPA were prepared in 1,4–dioxane.

To assay different concentrations of substrate, 5 µL of crude extract was incubated for 5 minutes at room temperature in 20 mM HEPES buffer, pH 8 and 150 mM of NaCl and enough ddH_2O to obtain 0.1 or 0.5 mL of total volume. For 1 mL of total reaction volume, 25 µL of crude extract was added and treated in the same manner as described above. The reaction was started by the addition of 0.1, 0.25, 0.5, 0.75, 1 and 3 mM of *p*-NPA and the absorbance change at 405 nm was monitored for 15 minutes at 20 °C.

2.6. Neuraminidase selection for esterase activity assay

All reactions were carried out in 100 mM acetate buffer, pH 5.5 supplemented with 0.01% of BSA (total reaction volume = 75 µL). Three different neuraminidases were tested, and these were from: *Micromonospora viridifaciens* (*Mv*NA), a mutant form of this neuraminidase (Y370G *Mv*NA) [98] and *Clostridium perfringens* (*Cp*NA) (Sigma) at final concentrations of approximately 3.3 x 10^{-7} µg/µL, 4.25 x 10^{-1} µg/µL and 1:10 dilution from stock, respectively.

The substrates 4-*O*Ac MUNANA and MUNANA were assayed at various concentrations (0.5, 1, 2, 3, 4, 5, 6, 8, 10, 14, 20 µM). The rate of hydrolysis was monitored by fluorescence spectroscopy and the data was depicted in Michaelis-Menten plots of the change in fluorescence intensity per second (int/s) as a function of substrate concentration. Each reaction was carried out for 5 minutes in a Cary Eclipse Spectrofluorimeter equipped with a Peltier temperature controller that was adjusted to 25 °C, using 450.9 nm emission and 322 nm excitation wavelengths.

A calibration curve was also performed by taking single readings of 4-methylumbelliferone (4-MU) in the assay buffer at the following concentrations: 0.5, 1, 2, 3, 5, 8, 10 µM, using the same spectrophotometer as above.

2.7. General procedures

Transformation and proliferation of pET101_HE plasmid.

The *pET101_HE* construct was transformed into BL21Gold(DE3) *E. coli* competent cells (Stratagene) by following the experimental conditions recommended by the manufacturer. Briefly, 40 ng of the closed plasmid was mixed with 100 µL of competent cells and incubated on ice for 30 minutes. The mixture was then heat shocked at 42 °C for 30 seconds, and placed on ice again for 2 minutes. Nine hundred microliters of pre-heated (at 42 °C) SOC or LB media was added to the DNA/cell mixture without antibiotics, and then shaken in an incubator at 37 °C for 1 hour at 225 rpm. The resulting cells were streaked onto plates containing LB agar supplemented

with 100 µg/mL ampicilin (amp) and then incubated overnight in at 37 °C. Bacterial colonies were picked manually and grown overnight in LB/amp at 37 °C, at 180 rpm.

Plasmid purification and insert check procedure.

After taking 700 µl of the media for glycerol stock preparations, the remaining media from the overnight cell growth was spun down at 3500 rpm for 10 minutes. The plasmid DNA from the pellet was extracted and purified using QIAprep Spin Miniprep Kit (QIAGEN) and DNA quantification was estimated by use of a nanodrop spectrometer. In order to verify the presence of the HE gene in the purified plasmid an insert check was performed. Basically, 1-5 µg of DNA was digested with 1 unit of corresponding restriction enzymes (Thermo Scientific), and 1 x of the appropriate buffer in a 10 µL reaction volume. The solution was incubated for 1 hour at 37 °C, then loaded onto a 1% agarose gel containing ethidium bromide and ran at 90 V for approximately 30 minutes. Gel bands were visualized by using a UV light box. All constructs were sequenced by Eurofins MWG Operon.

SDS-gel electrophoresis and Western Blot preparation.

Sodium dodecyl sulphate (SDS)-page gels were made using a BioRad casting apparatus. Briefly, 10 and 12% bis/acrylamide gels were made in the presence of SDS, ammonium persulfate (APS), and water. For the 5% stacking layer, 1 M of tris-HCl pH 6.8 was used and for the resolving layers, 1.5 M of tris-HCl pH 8.8 was added to the mix. Samples were prepared by mixing 2 µL of protein sample with 18 µL of water and 1 x Laemmli buffer and boiled for 5 minutes. A total of 20 µL was loaded into each well.

Western Blots were performed using the procedure as reported by Muller, 2008 [73] with slight modifications. Briefly, SDS gels were run at 100 V for approximately 45 minutes. The proteins in the gel matrix were transferred onto a polyvinylidene fluoride (PVDF) membrane using 1 L of a previously prepared 1 x transfer buffer (200 mL methanol, 100 mL of 10 x transfer buffer stock (250 mM tris-base, 2 M glycine) and 700 mL of ddH_2O) at 90 V for 1 hour, at 4 °C. Blocking solution was made of 5% w/v non-fat dry milk in 1 x Tween-PBS (T-PBS) buffer (1 x PBS, 0.1% Tween 20) and used to incubate the membrane at room temperature (rt) for 30 minutes, followed by an overnight binding at 4 °C with the primary antibody rabbit anti-HE peptide immune

serum (prepared according to Krossøy, 2001 [42], using MGDSRSDQSRVNPQSC and CPKMVKDFDQTSLGNT as peptides) diluted 1:2000 in 10 mL of blocking solution. After three 5 minutes washes with T-PBS buffer at rt, the membrane was incubated for 3 hours with the goat anti-rabbit HRP secondary antibody which had been diluted 1:2000 in 10 mL of blocking solution. The membrane was again washed three times for 5 minutes each and then incubated for 5 minutes with SuperSignal West Pico Chemiluminescent Substrate (Pierce), a chemiluminescent substrate solution. The membrane was then exposed to an autoradiography film.

Chapter 3. RESULTS AND DISCUSSION

3.1. Subcloning and modifications of HE constructs

Subcloning of pET101_HE into pVL1392 vector, to create *pVC_HE6* construct. The cDNA of hemagglutinin-esterase (HE) contained in the *pET101_HE* construct was amplified in order to obtain several copies of the gene with specific restriction sites (*Pst*I/*Bam*HI), to facilitate subcloning into pVL1392 vector and obtain the *pVC_HE6* construct. The PCR of the *pET101_HE* construct showed a clean band at the expected size in the agarose gel (Figure 3.1). This shows that the cDNA of HE was amplified successfully.

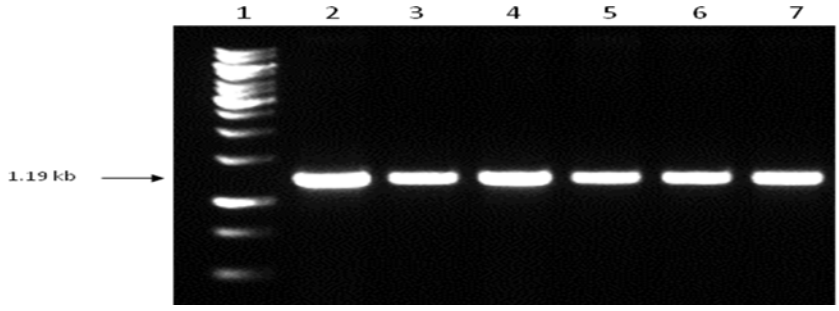

Figure 3.1 PCR products for HE, using *pET101_HE* as the DNA template.
The expected size is 1190 bp (1.19 kb). Lane 1: 1 Kb Ladder; Lanes 2-7: Positive PCR products.

The collected and concentrated PCR product was ligated into the pVL1392 transfer vector to obtain a construct to co-transfect into insect cells with the modified baculovirus. The insert check performed showed eight positive clones (Figure 3.2). In order to decide which of the clones to use for expression, all clones had their DNA

sequenced to confirm that the gene was in a correct reading frame and that no mutations were present within the cDNA.

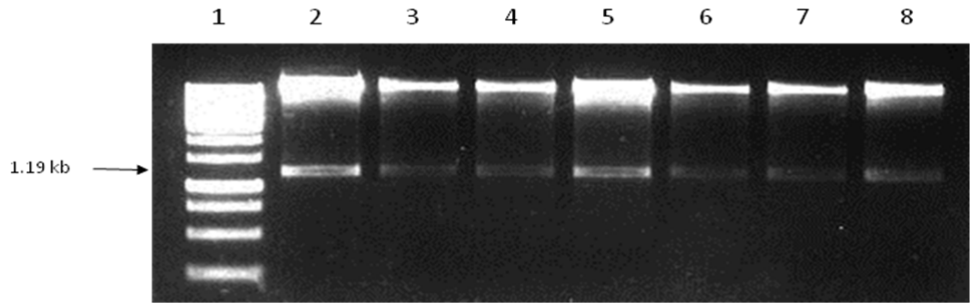

Figure 3.2 Insert check of *pVC_HE6* construct.
All digested plasmids show the presence of the cDNA of HE with the expected size (1.19 Kb). Lane 1: 1 kb ladder; Lanes 2-8: Positive clones.

All clones were sequenced and were shown to possess the same sequence. That is, the cDNA was ligated successfully, in frame with the plasmid and contained no mutations within the HE sequence. Figure 3.3 shows the sequence alignment of clone 6, which was chosen to carry out further experiments. This transient construct was named *pVC_HE6.*

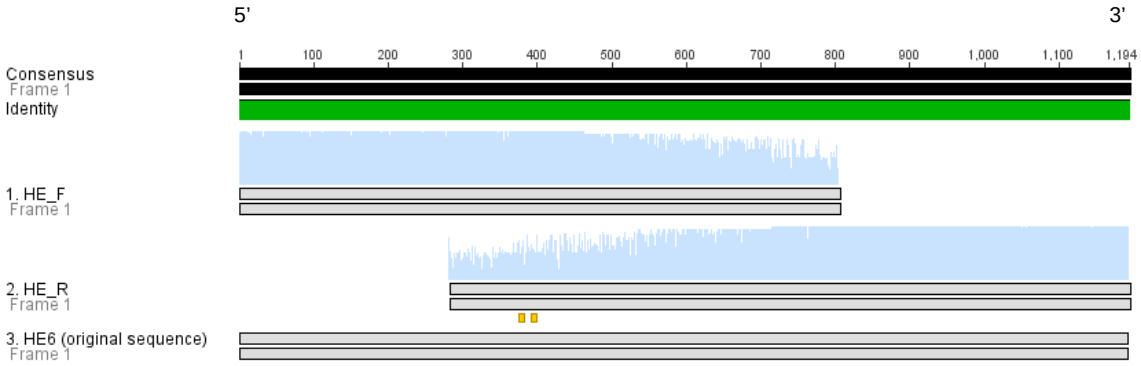

Figure 3.3 Sequence alignment of clone 6, *pVC_HE6*.
The green bar shows whether the cloned sequence is identical to that of the template DNA. A full green bar means that the sequence of the clone exactly matches with the original. HE6 is the original sequence from the *pET 101/D-TOPO*, HE_F is the extraction of the forward sequence; HE_R is the extraction of the reverse sequence. The yellow squares represent the sites where there was a manual correction in the base pairs after the alignment when this was analyzed in the Geneious program.

The amino acid sequence of *pVC_HE6* was then submitted for analysis of possible domain structures within the resulting expressed protein to the online SMART tool. This program allows the identification of domains in proteins that share amino acid residues that are involved in functionality and tertiary structures using known sequences that are stored in the database. It was found that the program suggested HE sequence has two clear domains, which were labelled as '*Low complexity region'* (from amino acid 179 - 191) and a '*transmembrane domain'* (from amino acid 356 - 378). This analysis can be seen pictorially in Figure 3.4.

The Low complexity regions or LCRs are sequences within a protein that display a few amino acids repetitively, which means that such regions show little variability. Typically, these residues are found clustered or spaced periodically within the region [99]. Many theories have been proposed regarding the functionality of LCRs. For example, it has been suggested that in eukaryotes, LCRs promote mRNA stability [100] and protein-protein interactions [101]. Others have suggested that LCRs frequently occurs at the antigenic loci, and this LCRs may have a pathogenesis role involving antigen diversification [102]. In contrast, researches have hypothesised that LCRs do not have a specific function and they are just simple spacers between the important structural features of a protein [103]. As a result, we cannot conclude precisely what the role of the putative LCRs does in HE and its function.

In contrast, the function of the *transmembrane domain* or TMD is known. This region is composed of approximately 25 hydrophobic amino acid residues and it plays an important role in the cell functions. Proteins carrying TMDs or membrane proteins are involved in cellular signalling, differentiation, proliferation and adhesion of cells; reception of growth factors such as hormones, protein channelling and two-way membrane transporters [104]. In general, membrane proteins are produced in the ER of eukaryotic cells and then transported out through the ER membrane via translocons.

Since purification of the HE protein would be facilitated if it was secreted from the cell after expression, we removed the bases encoding the TMD from the cDNA via PCR, before the gene was co-transfected into insect cells. This was performed following the rationale described in Figure 3.4.

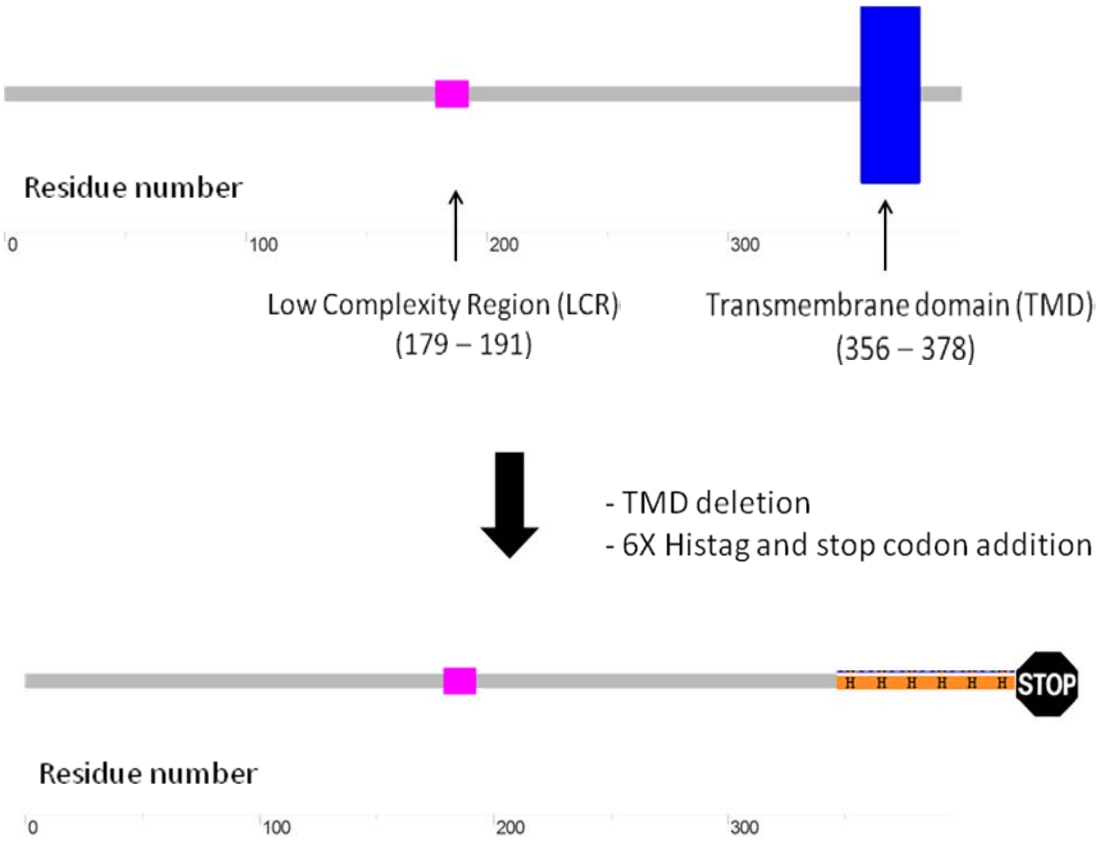

Figure 3.4 Illustration of the TMD deletion method and the 6X Histidine tag (Histag) addition.
The analysis of the amino acid sequence of the HE protein showed a low complexity region (LCR) from amino acids (179-191) and a transmembrane domain (TMD) from residues (356- 378). The last 22 amino acids from the *N*-terminus of the protein were replaced by a 6X Histag followed by a stop codon.

Deletion of the TMD from the protein sequence will result in a soluble protein, which would make the protein to be secreted out from the cell to the cell expression media, facilitating its harvest. Because of this, a TMD deletion procedure was carried out, and in addition, a 6x Histag was added to facilitate purification.

TMD deletion and preparation of pVC_HE7. The deletion of the TMD was realised by performing the PCR experiment with the specific primers as detailed in Table 1. The PCR reaction was successful as can be seen in Figure 3.5a, where the resulting DNA products have the expected size (1086 bp or ~1.1Kb). After ligation and transformation an insert check was performed in order to verify that the PCR products were ligated into the pVL1392 vector using the restriction sites *Pst*I and *Bam*HI, incorporated from the primers. Five clones displayed positive sequences and clone named 'seven' was chosen for the further experiments. This clone was therefore named *pVC_HE7*.

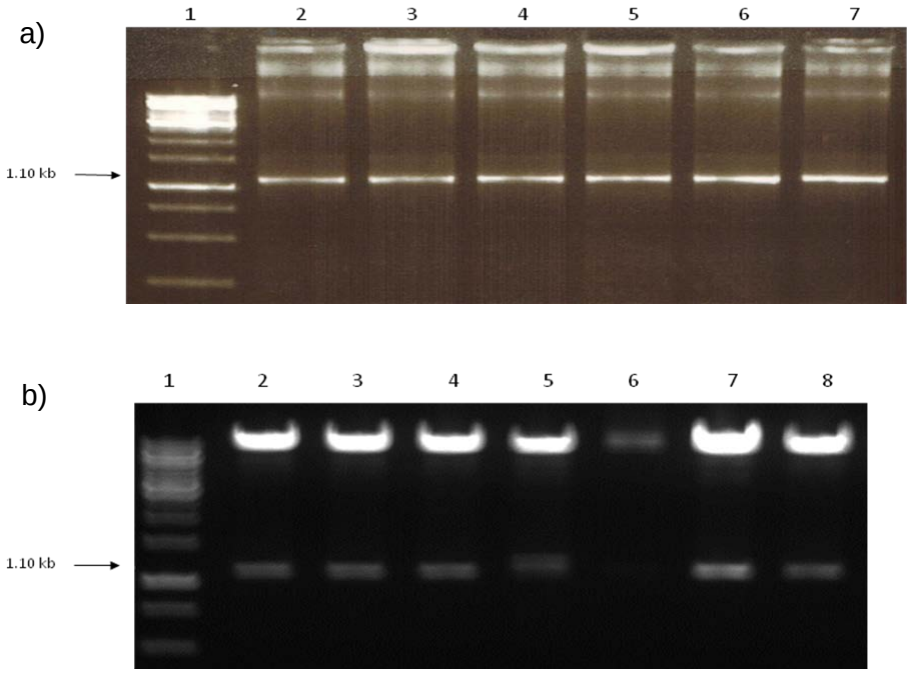

Figure 3.5. DNA gel images of TMD deletion from *pVC_HE6* construct.
The expected size was 1086 bp (~1.1 kb). **a)** PCR products. Lane 1: 1 Kb ladder, Lanes 2-7: Positive PCR products. **b)** Insert check. Lane 1: 1 Kb Ladder; Lanes 2-4, 7, 8: Positive clones with deleted TMD. Lane 5: Negative clone where TMD remains in the PCR product. Lane 6: Negative clone where insert was not ligated into plasmid.

Sequencing analysis shows an almost perfect alignment with the original sequence (Figure 3.6a). All base pairs are aligned, except for the 3' end of the sequence where gaps can be observed. This was expected since the modification of the DNA was precisely at that end. Figure 3.6b shows an enlargement of the

modification zone and it can be seen that the TMD is no longer in the sequence and that it has been successfully replaced by a 6X Histag and a stop codon.

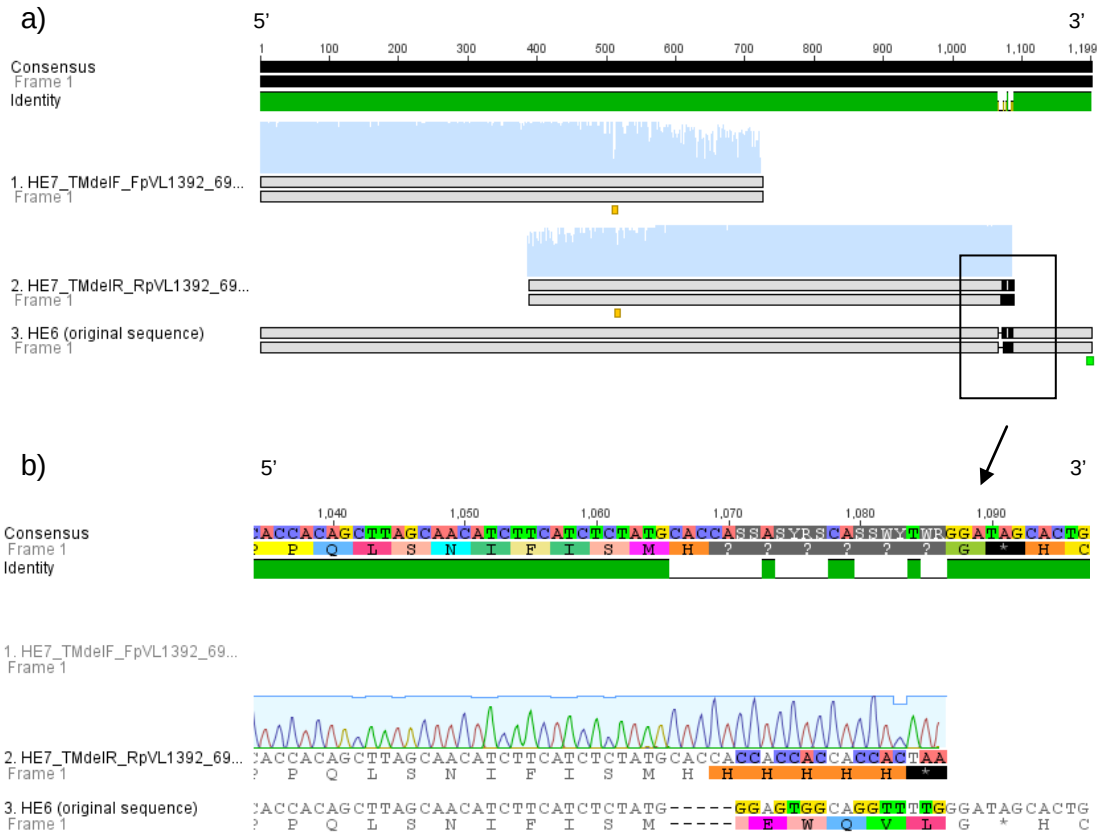

Figure 3.6 Sequence analysis of the TMD deletion experiment that gave the *pVC_HE7* construct.
Panel a shows the DNA sequence alignment of clone seven with the original sequence of HE. The green bar represents a perfect alignment between the two sequences, while the white gaps represent different bases in the sequence. Panel b shows an enlargement of the TMD modification at the 5' end of the reverse sequence. The black square represents the region that is shown explicitly. A 6X Histag is clearly seen followed by a stop codon. The yellow squares represent the sites where there was a manual correction in the base pairs after the alignment when this was analyzed in the Geneious program

3.2. Protein analyses of HE using the *pVC_HE7* construct

An expression trial with the pVC_HE7 construct using the BaculoGold system and the *HighFive* cell line was performed. A protein analysis was made by SDS-page to check for expression, however, the protein was not present in the cell media or in the cell pellet, since the expected band size of ~42 KDa was not visible. This experiment was repeated numerous times using on each occasion either several plates containing 10 mL of culture media or two plates with 25 mL media. I did this in order to determine if protein expression levels were too low to be measured. The SDS-page analyses showed that in both the cell media or in the cell pellet no significant amount of HE protein was present, an observation that was independent of the volume of culture media. This leads us to believe that the expression was not taking place inside the cell. These results were similar in all trials.

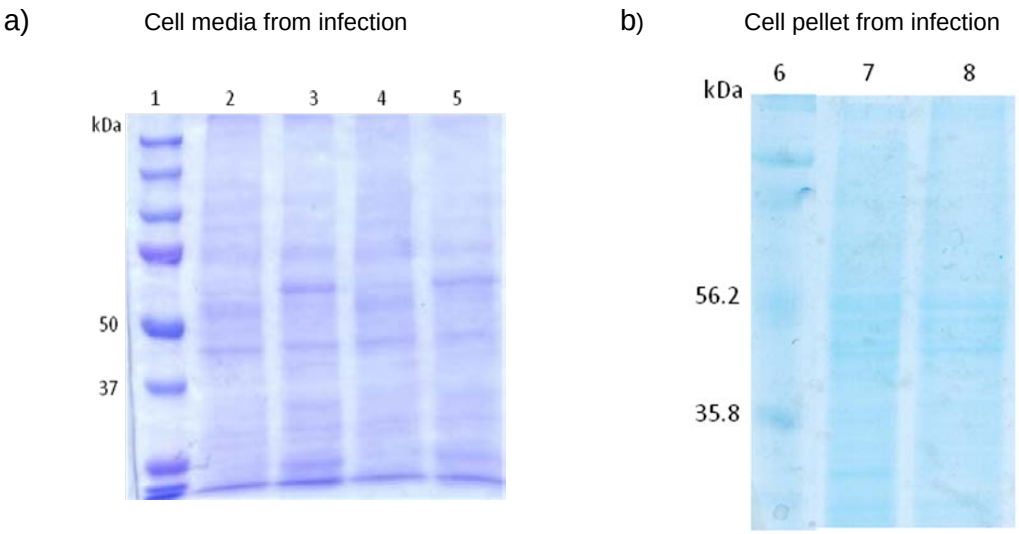

Figure 3.7 SDS-Page of crude extracts from expression using BaculoGold viral DNA. **a)** cell media and **b)** cell pellets. Crude extracts were taken from the HTV stage. Lane 1: Protein ladder. Lane 2: Expression of HE using HighFive cells in ~60 mL of culture media; Lane 3: Negative control of expression using HighFive cells in~40 mL of media; Lane 4: Expression of HE using HighFive cells in ~75 mL of culture media; Lane 5: Negative control of expression using HighFive cells in ~75 mL of media. Lane 6: Protein ladder; Lane 7: Expression of HE using HighFive cells in pellet; Lane 6: Negative control of HighFive cells in pellet.

Since the expression of HE construct *pVC_HE7* using the BaculoGold system did not work, I decided to try with a different cDNA of the gene. As described above, the first expression trial used the gene of the protein from a Canadian strain of ISAv. The first reported ISA outbreak occurred in Norway, while others have been reported in Europe, South and North America [14, 15]. It is known that the ISA disease is caused by different strains of the virus. The reason of the variety if ISAv isolates is still unknown [105] however, Blake et al.(1999) reported they have certain degree of differences between sequences of these strain. In this study different strains from North America and Europe were analysed [106]. For the genotypes of these strains, there is an established functional relationship involving the pathogenicity of ISAv in cell culture and the length of the HE stem [107]. The cDNAs of the viral strains vary between segments and specifically for segment 6, which encodes for HE protein, the Canadian ISAv isolate from Back Bay, New Brunswick, Canada shares approximately 85% of protein sequence similarity with isolates from Norway and Scotland and a 99.2% with another North American isolate from Maine [108]. In addition, the cDNA of HE from the Norwegian strain has two *N*-glycosylation sites at amino acids 333 and 349 [46] whereas the Canadian (or North American) strain has only one, located at amino acid 156 [108]. Although the differences between sequences in ISAv isolates from different geographical areas are small, it is of interest to find out if they have any effect on the expression of the protein, thus it was decided to try the expression of HE using the cDNA from a different strain, specifically the Norwegian isolate Glesvaer 2/90.

3.3. Subcloning of *pEGFP_HE_Gles2/90* into *pVC_HEGles2/90* construct.

The PCR products correspond to the cDNA of HE from the *pEGFP_HE_Gles2/90* construct (Figure 3.8), which was used as the DNA template. A higher band of about 7 Kb is seen in all lanes, which may indicate a possible hybridization of the primers to another segment of the plasmid. Since the primers for this PCR reaction were design to exclude the TMD from the cDNA, the expected size of the PCR product is lower than for the *pVC_HE6* construct.

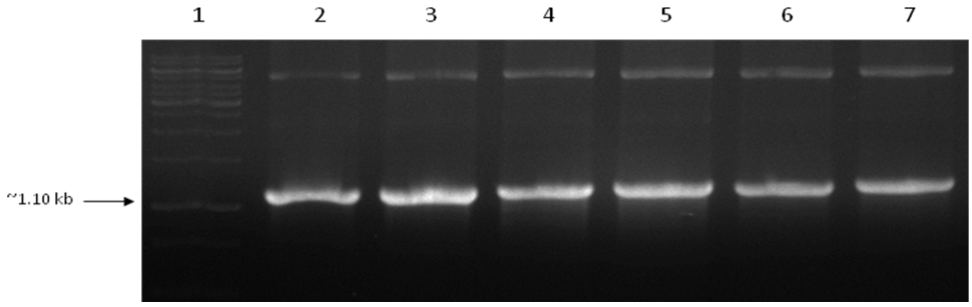

Figure 3.8 PCR products for HE, using *pEGFP_HE_Gles2/90* as DNA template. The expected size was 1086 bp (1.1 kb). Lane 1: 1 Kb Ladder; Lanes 2-7: Positive PCR products.

After ligation of the PCR product using *EcoR*I and *Sma*I restriction sites, an insert check was performed in order to verify that it successfully merged with the pVL1392 vector. Figure 3.9 shows the digestion of the ligated products with the mentioned restriction enzymes. It can be clearly seen that three of the five clones were positive as the gel showed a DNA band of the expected size. There is a very faint band in lane 2, which shows that this is a positive clone, although due to the image resolution it cannot be easily seen.

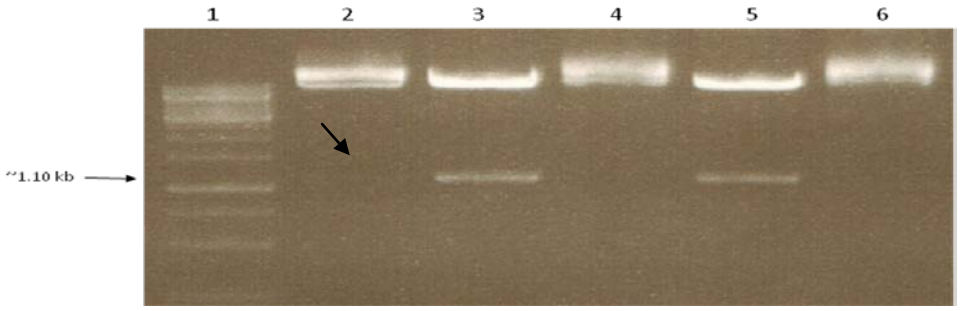

Figure 3.9 Insert check of *pVC_HEGles2/90* construct. Plasmid digestion shows the presence of the HE insert with the expected size only in three clones. Lane 1: 1 kb ladder; Lanes 2, 3, 5: Positive clones (named 1, 2 and 4, respectively); Lanes 4, 6: Negative clones. The arrow indicates the faint band in lane 2 for one of the positive clones

As I had done earlier with the *pVC_HE7* construct, all three positive clones of the *pVC_HEGles2/90* construct were sequenced in order to verify that the insert was ligated into the correct reading frame and that no mutations had occurred during the PCR process. The sequencing analysis is displayed in Figure 3.10 and this shows the alignment of the original sequence of HE with one of the positive clones.

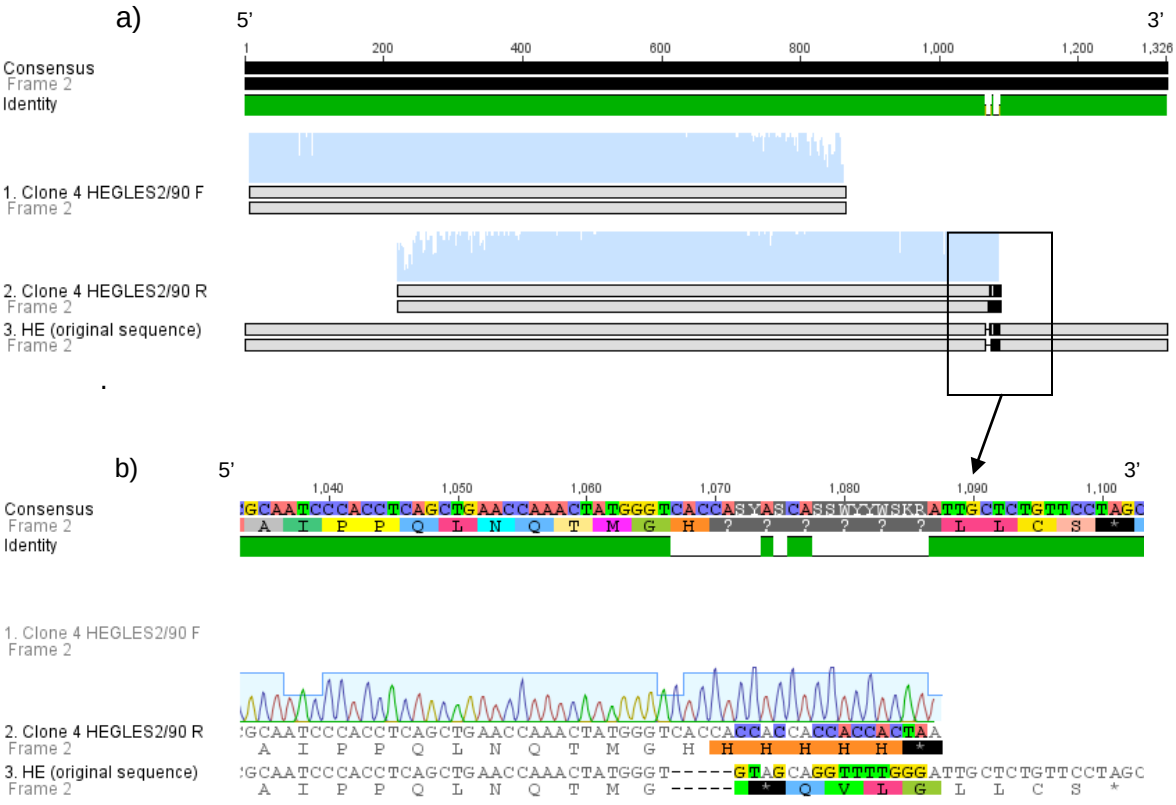

Figure 3.10 Sequence analysis of TMD deletion from *the pVC_HEGles2/90* construct. Panel a shows the sequence alignment of one of the positive clones with the original sequence of HE. The green bar represents a perfect alignment between them, except for the gaps shown at the 3' end of the sequence. Panel b shows an enhancement of the TMD modification at the 5' end of the reverse sequence. The black square represents the region that is enhanced. A 6X Histag is clearly seen followed by a stop codon.

3.4. Protein analyses of HE using the *pVC_HEGles2/90* construct.

After showing that the sequence of clone four was good, an attempted expression of HE using the *pVC_HEGles2/90* construct was performed, in a BacNBlue expression system using both *Sf9* and *HighFive* cell lines. The protein expression was again unsuccessful, as it can be seen in the SDS Page analysis of the cell media and pellets (Figure 3.11).

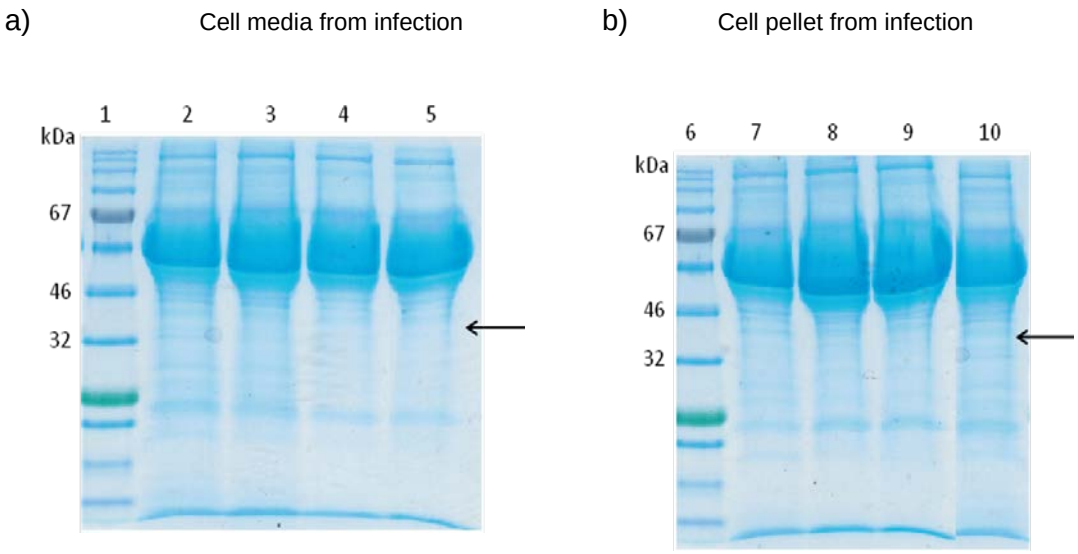

Figure 3.11 SDS-Page of crude extracts from expression using BacNBlue viral DNA. Panel a): cell media and panel b): cell pellet. Crude extracts were taken from the HTV stage. Lane 1: Protein ladder. Lane 2: Expression of HE in *Sf9* cells; Lane 3: Negative control in *Sf9* cells; Lane 4: Expression of HE in HighFive cells; Lane 5: Negative control in HighFive cells; Lane 6: Negative control in *Sf9* cell pellet; Lane 7: Expression of HE in *Sf9* cell pellet; Lane 8: Protein ladder; Lane 9: Negative control in HighFive cell pellet; Lane 10: Expression of HE in HighFive cell pellet. The arrows show where expected position of the HE protein band.

Many attempts were made in order to express the HE protein. I tried the expression repeating the experiment using the same conditions, changing the volume of culture media as described in 3.2, and performed expression trials at 20 °C and 27 °C; however no protein with the expected molecular weight was observed in the protein gels. Since it is a much more sensitive technique than SDS-Page, western blot analysis

was performed on some of the samples that had already been examined by electrophoresis. I did these experiments to corroborate the absence of protein in the crude extracts

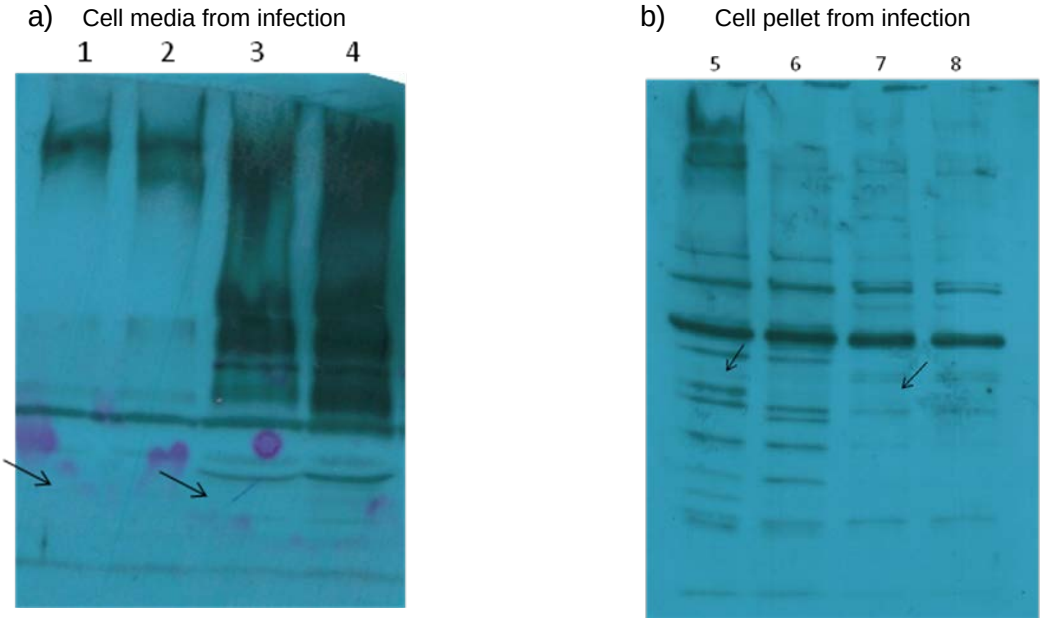

Figure 3.12 Western blots of crude extracts.
Panel a): Blots from HTV stage of expression using *pVC_HEGles2/90* construct with both cell lines. Lane 1: Expression in *Sf9* cells; Lane 2: Negative control in *Sf9* cells; Lane 3: Expression in HighFive cells; Lane 4: Negative control in HighFive cells. Panel b): Blots from cell pellets from both cell lines. Lane 5: Expression of HighFive cells; Lane 6: Negative control in HighFive cells; Lane 7: Expression in *Sf9* cells; Lane 8: Negative control in *Sf9* cells. Arrows indicate where the protein band should be. Anti-HE peptide and goat anti-rabbit horse radish peroxidase (HRP) were used as primary and secondary antibodies respectively.

Based on the lack of protein expression that I observed in the BaculoGold and BacNBlue baculovirus systems, I decided to attempt a different approach for the production of recombinant HE. Although the expression of HE in mammalian cell has not been reported, the cDNA that we were sent from Norway was ligated into an expression vector for mammalian cells, therefore, I decided to do a trial expression in Chinese Hamster Ovary (CHO) cells.

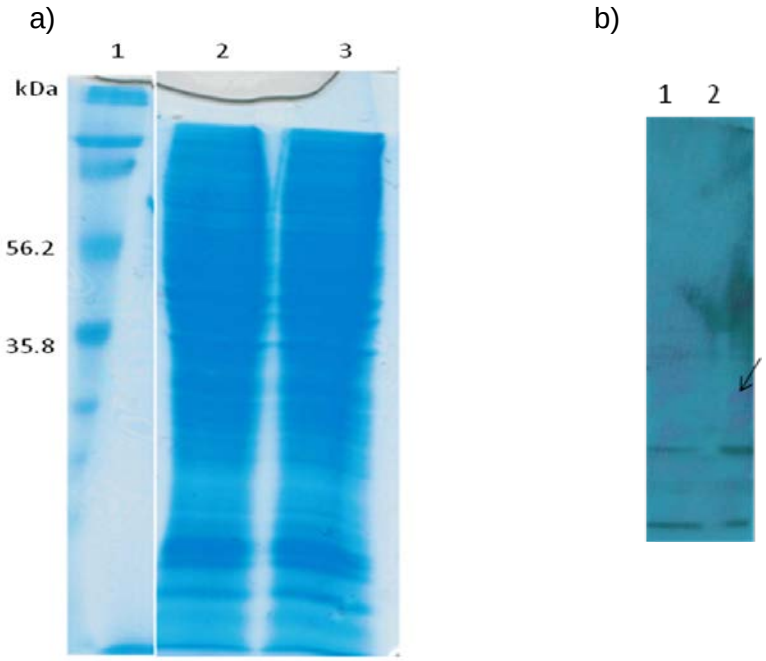

Figure 3.13 Protein analysis of HE by SDS-page and western blot.
Panel a): SDS-Page of expression of HE using *pEGFP_HE_Gles2/90* in CHO cells. Crude extracts show a great quantity of proteins, however not the one with the expected size. Lane 1: Protein ladder; Lane 2: Negative control of CHO cells; Lane 3: Expression in CHO cells. Panel b): Western blot of the expression with CHO cells. Lane 1: Negative control; Lane 2: Expression in CHO cells. The arrow indicates where the protein band should be.

Figure 3.13 shows the electrophoretic analysis of the cell media from CHO cell expression. As mentioned before, there are no reports for the recombinant expression of this protein in mammalian cell lines; however, expression has been reported using a pEGFP-N1 vector in Salmon Head Kidney (SHK) cells [109]. The culturing of these type of cells are carried out at around 20 °C, a lower temperature than that used for culturing of mammalian cells (37 °C). These conditions, together with the fact that there is no evidence of the protein being expressed at 37 °C, the optimal temperature for growing CHO cells, may suggest that the expression of HE using these mammalian cells would not be successful.

3.5. Activity assays from constructs *pVC_HEGles2/90* and *pEGFP_HE_Gles2/90*

Even though a successful expression was not observed in the electrophoresis and in the western blots analyses, I decided to perform several enzymatic activity assays in order to further confirm the non-existence of the HE protein in the samples.

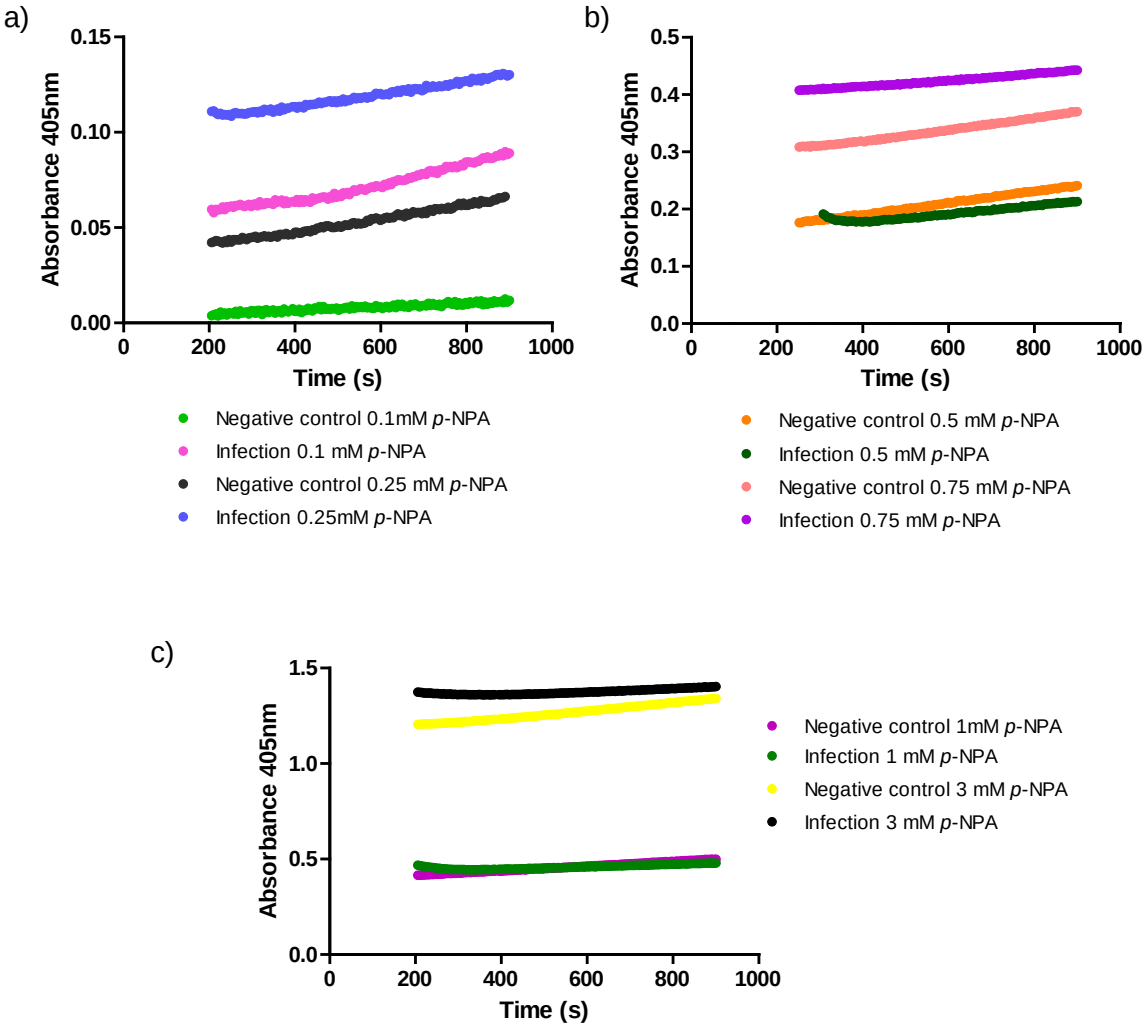

Figure 3.14 Activity assays from expression with HighFive cells using *pVC_HEGles2/90* construct.
Activities were monitored measuring the change in absorbance at 405nm versus time (in seconds) using different concentrations of the substrate *p*-NPA, pH 8.0 at 20°C. **a)** Activity measurements with 0.1 and 0.25 mM of *p*-NPA. **b)** Activity measurements with 0.5 and 0.75 mM of *p*-NPA. **c)** Activity measurements with 1 and 3 mM of *p*-NPA. Around 200 initial data points were removed in order to obtain clear slopes.

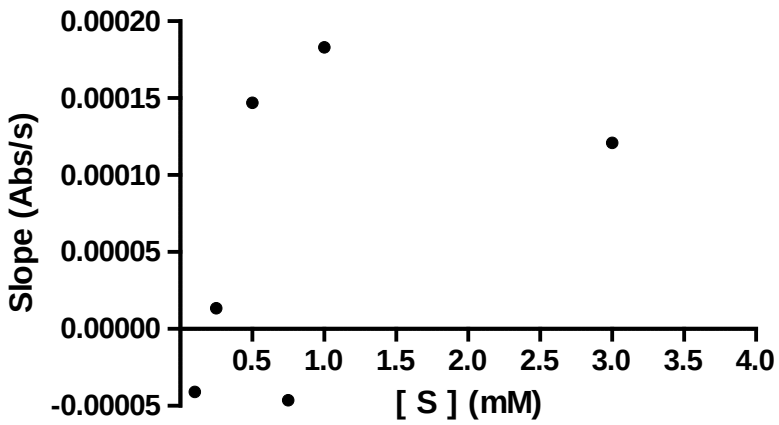

Figure 3.15 Plot of the difference in slopes between the negative controls and infection versus concentrations of *p*-NPA.
The concentrations from figure 3.14 were plotted by subtracting the slopes from negative controls and the infection reaction. Two data points have negative values for 0.1 and 0.25 mM of *p*-NPA.

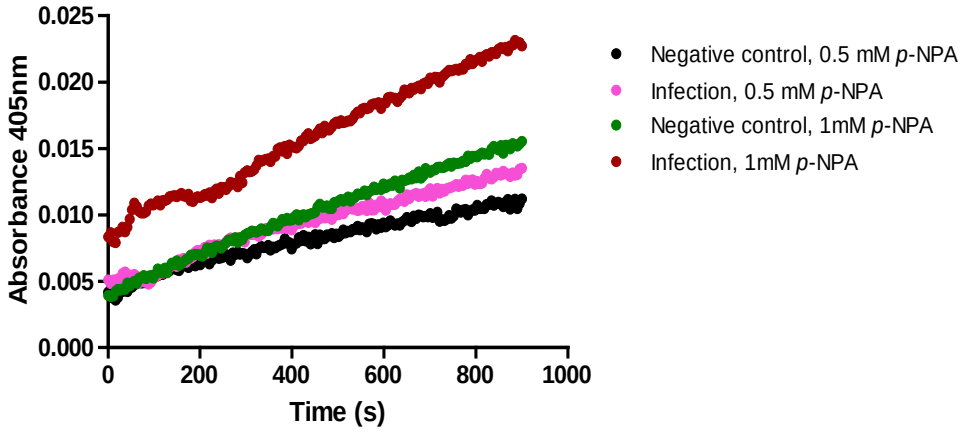

Figure 3.16 Activity assay of expression in *Sf9* cells using the *pVC_HEGles2/90* construct.
Activities were monitored measuring the change in absorbance at 405nm versus time (in seconds) using two concentrations of the substrate *p*-NPA. The change in absorbance from start to end of the reaction is minute.

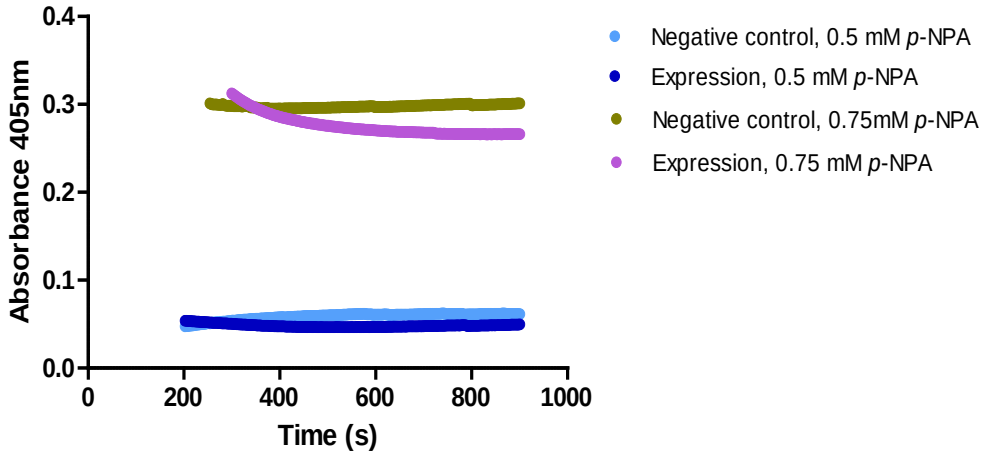

Figure 3.17 Activity assay of expression in CHO cells using the *pEGFP_HE_Gles2/90* construct.
Activities were monitored measuring the change in absorbance at 405nm versus time (in seconds) using two concentrations of the substrate *p*-NPA. The slopes have no significant difference between the negative controls and the expression assays. Around 200 initial data points were removed in order to obtain clear slopes.

As expected, the assays of the infected cell media from the different expression systems and constructs showed no significant esterase activity. The activity assays showed here correspond to infections with the *pVC_HEGles2/90* and *pEGFP_HE_Gles2/90* constructs. The first two hundred data points were deleted from graphs in Figures 3.14 and 3.17 due to some strange interaction of the mixture, forming oscillations that could possibly be generated by the presence of bubbles in the reaction. It can be seen in Figure 3.14, all graphs show that the absorbance fluctuates throughout the time course of the experiment. It can also be observed that the background absorbance proportionally increases with the concentrations of substrate. In addition, the graphs show that the negative control slopes for all concentrations of substrate are lower than the infection slopes for the same concentrations, with the exception of the slopes from 1mM *p*-NPA (Figure 3.14c); however, we cannot relate these changes to enzyme activity given that the negative control samples also show similar behaviour. This can also be observed in Figure 3.17, where no significant difference can be seen between transfection reaction and negative control from the expression with CHO cells. Figure 3.15 shows the difference in slopes between negative control and the cell media from infection versus the concentration of substrate.

Some of the differences are negative indicating greater activity in the negative controls reactions, which presumably is an artefact. In Figure 3.16, the activity assays from expression in *Sf9* cells is shown. Again, the same behaviour between negative control and infected cell media can be observed. The lack of activity in this experiments is again consistent with no active HE being produced during expression.

It is of interest to speculate why the expression of ISAv HE protein was not successful, particularly since several studies report the successful expression of active protein. Indeed, these studies used the same or similar baculovirus expression systems that we employed in this thesis [16, 43, 109, 110]. It is important to point out, however, that the majority of the HE protein studies have expressed this protein and/or carried out several experiments using fish cell lines such as SHK and CHSE [17, 109, 111] by infecting them with the whole active ISA virus, whereas only a couple used the Baculovirus system [43, 73]. This suggests that this system is relatively difficult to make work with this protein; otherwise it would have been used extensively.

A possibility for this unsuccessful expression could be an improper folding of the protein inside the insect cells. The ER is the organelle responsible of manufacturing virtually all proteins in the eukaryotic cells and of the correct folding of the proteins before they are secreted and transported to the cell surface via a series of pathways and signalling. When a protein does not fold correctly, it is recognized by a the ER chaperone protein calnexin, which retains and transports the misfolded protein to the proteosome for degradation [112] by a process called ER-associated degradation or ERAD in which the misfolded protein is sent to the cytosol by retrotranslocation [113].

The folding of proteins in the ER is a slow and sometimes an inefficient process, with the results that large polypeptides can fail to reach their native states [114]. In the folding process, *N*-glycosylation plays an important role especially for our protein of interest, since it needs to have this post-translational modification in order to be active. As previously discussed, *N*-glycosylation is basically an enzymatic addition and trimming pathway of removal and additions of saccharides attached to the polypeptide chain via an Asn residue, forming a branch-like of sugars linked to the protein. At some point of this process, one specific mannose is removed from the saccharide branch

leaving exposed a 1,6-mannose residue, which is a key signal for degradation [114-116]. Although it is very unlikely, this could have been one of the reasons why our protein was not expressed.

Of note, we managed to obtain the cDNA of the Norwegian isolate of ISAv HE protein that the authors used in their publication [73], however we requested for either the high titer virus or the construct they used to carry out their electrophoretic and enzymatic analysis. Unfortunately, they were not able to locate neither of the requested items but instead they sent us the construct mentioned in Chapter 2, a GFP-tagged HE, which is a construct that was intended to use in mammalian expression systems. This tells us that the researches that expressed the HE protein using baculovirus have not been working with this system in a considerable amount of time, which suggests that they are no longer using it as a protein expression tool. Nevertheless, a second part of this interesting project was performed, which is the assessment of esterase activity using a specific esterase assay and substrate namely, 4-*O*Ac MUNANA.

3.6. Esterase activity assay using a specific substrate and assay.

Several reports have shown the expression of the hemagglutinin-esterase (HE) from ISAv and esterase activity analysis of the protein. The majority of these studies perform their activity assays by measuring change in absorbance using *p*-nitrophenyl acetate (*p*-NPA) as the substrate (Fig 3.18a), either when the protein is expressed using baculovirus system or infecting SHK cells with live ISAv [10, 16, 43, 73]. The hydrolysis of *p*-nitrophenyl acetate releases *p*-nitrophenol, and the change in absorbance may be monitored at 405 nm.

Even though this substrate is widely used in enzymatic activity analyses, it has several disadvantages: i) it readily hydrolyzes in basic media, for instance at a pH of 9.0 *p*-NPA has a half-life of ~1.66 hrs [117]; ii) it is a non-specific substrate, which means that any kind of esterase can hydrolyze it; and (iii) it readily undergoes catalyzed hydrolysis by many nitrogen bases [118]. Because of these issues a need exists to find

a substitute substrate that is much more specific than *p*-NPA for this particular esterase from ISAv.

As it was stated previously, the specific receptor for the ISAv HE protein is a 4-*O*Ac-α-sialoside. Dr. Fahimeh Shidmoossavee from the Bennet laboratory at Simon Fraser University synthesized such a substrate that was needed for this specific esterase assay. This consisted in a 4-*O*Ac sialic acid linked to a 4-methylumbelliferone (MU) group at position C2 of the sugar. This compound was named 4-*O*Ac MUNANA (Fig 3.18b).

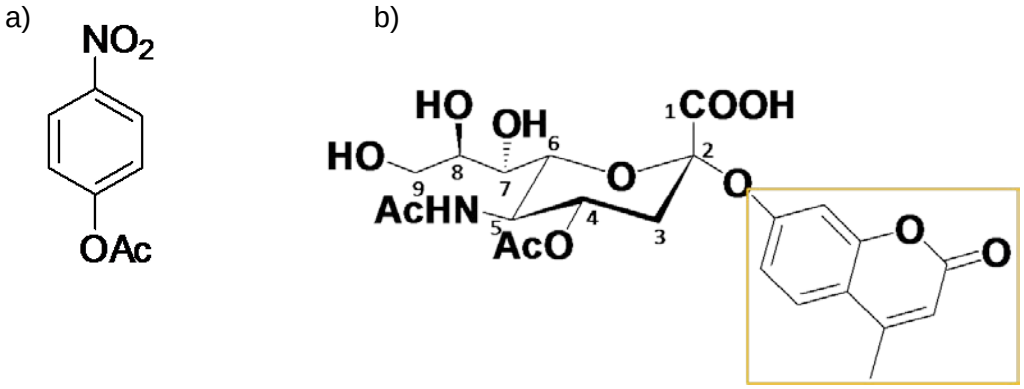

Figure 3.18 Substrates for esterase activity analyses.
Panel a: Structure of *p*-nitrophenyl acetate, the commonly used substrate for esterase activity assays. Panel b: Structure of 4-*O*Ac MUNANA, a specific substrate generated by the Bennet laboratory. This substrate was created to assess esterase activity from ISAv HE protein. The yellow square highlights the 4-methylumbelliferone (MU) group, linked to C2 of the sialoside. The numbers show the carbon positions on each structure.

3.6.1. Specific esterase activity assay using 4-*O*Ac MUNANA as substrate

With our specific substrate in hand it was necessary to develop an assay for the esterase activity of the HE protein from ISAv. This assay can be monitored by fluorescence intensity changes due to the presence of the 4-MU group in the substrate, which is liberated upon neuraminidase-catalyzed hydrolysis. It is a coupled assay involving two main enzymes: the hemagglutinin-esterase from ISAV and a neuraminidase that would either not hydrolyze 4-*O*Ac MUNANA or have a greatly reduced activity against this compound, as this carbohydrate moiety is not the natural substrate for this enzyme. The proposed scheme for the esterase activity assay is shown in Figure 3.19.

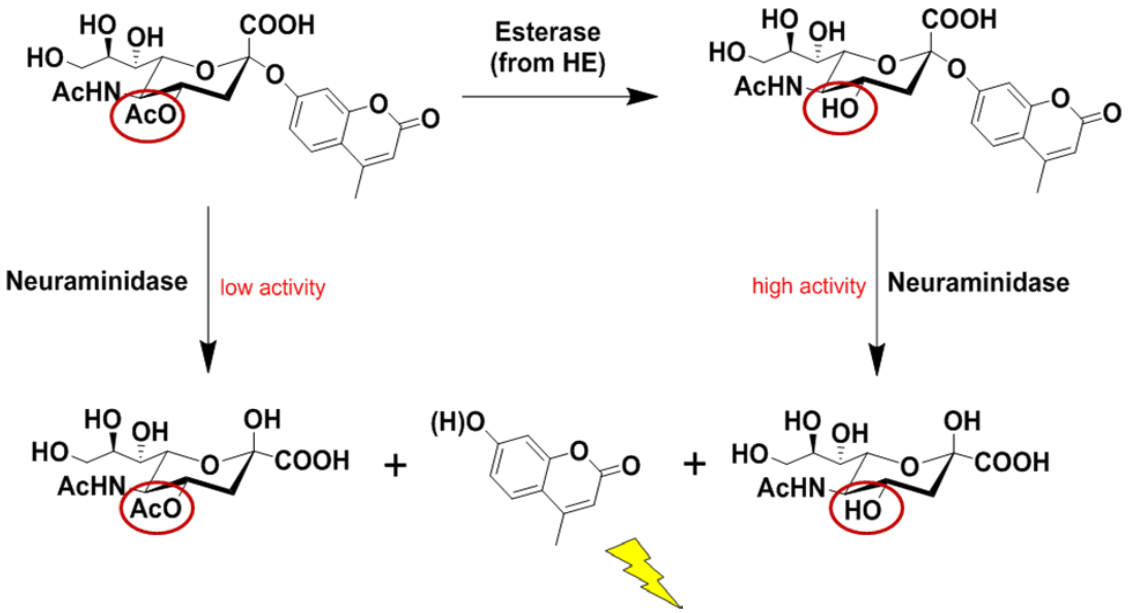

Figure 3.19 Coupled esterase assay using 4-*O*Ac MUNANA as substrate
In this assay, 4-*O*Ac MUNANA should not be hydrolyzed by the neuraminidase enzyme, thus representing our negative control. The neuraminidase present in the reaction should hydrolyze MUNANA, the de-acetylated form of 4-*O*Ac MUNANA, de-attaching the 4-methylumbelliferone (MU) group from the sugar creating fluorescence. The de-acetylation of 4-*O*Ac MUNANA should be performed by the esterase from the HE protein.

4-*O*Ac MUNANA in the presence of a neuraminidase will slowly be hydrolyzed by the enzyme; and this reaction constitutes the negative control. Since this substrate is specific for the HE protein, the esterase will recognize the acetyl group at C4 of the sugar and will cleave it off, transforming this substrate into MUNANA, which is the parent sialic acid linked to a 4-MU group. Once acetyl hydrolysis has taken place, the neuraminidase present in the reaction will recognize MUNANA as its specific substrate, and break the glycosidic bond at C2, releasing the MU into the environment as an equilibrium mixture of the phenol and phenoxide forms, and this results in an increase in fluorescence.

3.6.2. Determination of the appropriate neuraminidase for the esterase assay

Neuraminidases are surface proteins that are found in the envelope of the Influenza viruses, but they can also be found in bacteria as well [119, 120]. Their function is to cleave off the α(2,6) or α(2,3) linkages between the sialic acid present in the host cell and a galactose on glycoconjugates, in order to release viral progeny [121]. As it was described previously, the most common substrate for this enzyme is sialic acid or neuraminic acid.

In order to determine which neuraminidase will work more efficiently in this assay, that is, which will have the lowest activity in the presence of 4-*O*Ac MUNANA and a high activity against MUNANA, three different enzymes were tested. Neuraminidases from: *Micromonospora viridifaciens* (*Mv*NA), *Clostridium perfringens* (*Cp*NA), and a mutant neuraminidase from *M. viridifaciens* (Y370G *Mv*NA) [98] were monitored against these substrates and fluorescence intensity was measured (Fig 3.20).

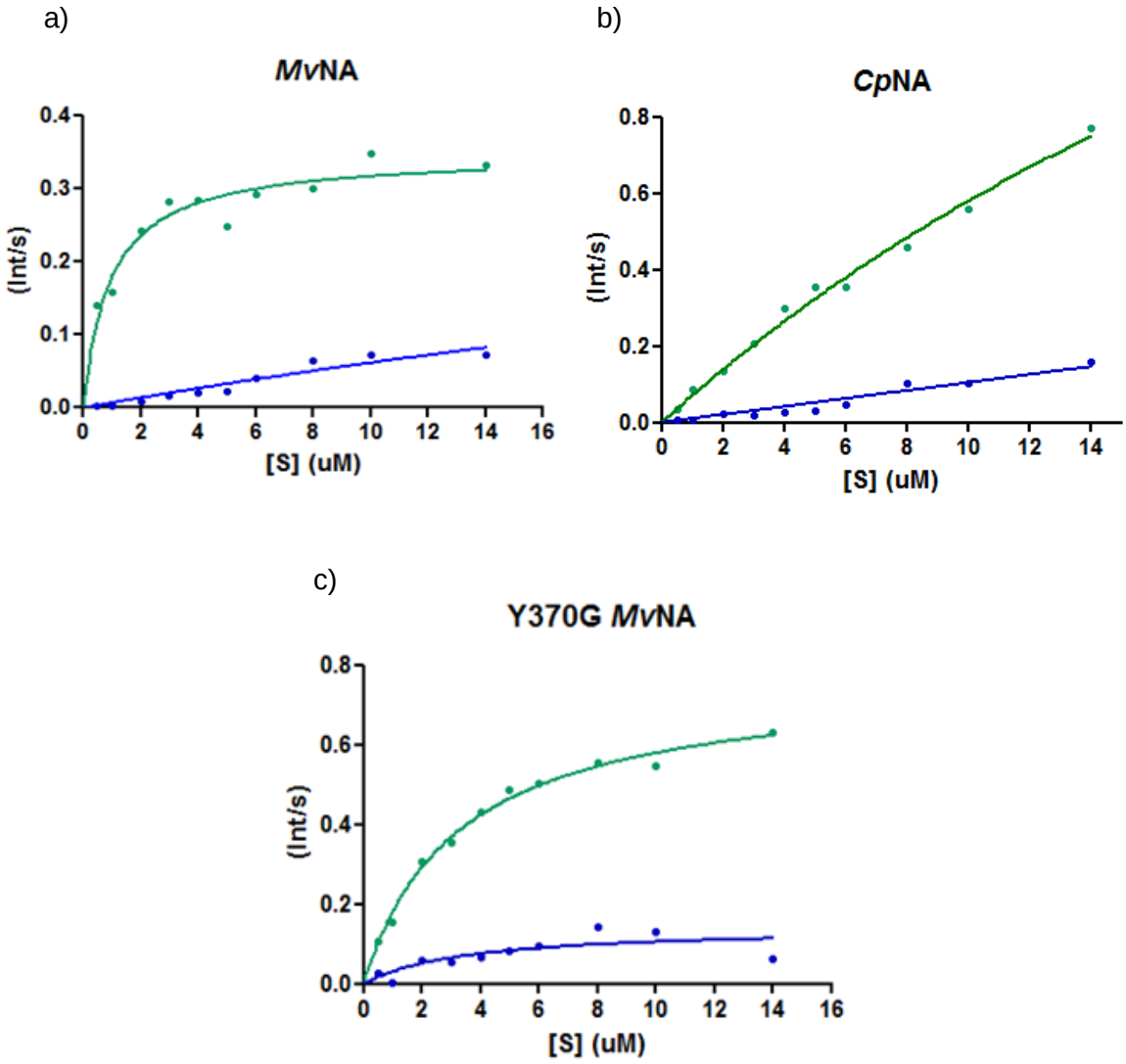

Figure 3.20 Michaelis-Menten plots of various neuraminidases against 4-*O*Ac MUNANA and MUNANA as substrates.
Neuraminidase activity from: **a)** *Micromonospora viridifaciens* (*Mv*NA) *b) Clostridium perfringens* (*Cp*NA) and **c)** mutant neuraminidase from *M. viridifaciens* (Y370G *Mv*NA). It can be seen the hydrolysis rate of different concentrations of MUNANA (green) and 4-*O*Ac MUNANA (blue) as intensity per second (Int/s).

Figure 3.20 shows the rates of hydrolysis for the substrates 4-*O*Ac MUNANA and MUNANA using three different neuraminidases. Figure 3.20a displays the kinetic data for *Mv*NA and this shows that this enzyme has a high affinity for MUNANA (low K_m) and that low concentrations of this substrate are rapidly hydrolyzed (high k_{cat}/K_m). In contrast, hydrolysis of 4-*O*Ac MUANANA is much slower, the apparent binding constant (K_m) to *Mv*NA is much higher and the catalytic efficiency (k_{cat}/K_m) is much lower. In other words, the difference in rates of hydrolysis between these two substrates using *Mv*NA is high, which is the requisite condition needed for a sensitive HE activity assay. These data are in contrast to that shown in figures 3.20b and 3.20c, where *Cp*NA and Y370G *Mv*NA exhibit greatly reduced differences between the rates for hydrolysis of MUNANA and 4-*O*Ac MUNANA. Therefore, it was decided that *Mv*NA is the best enzyme to perform the HE esterase assay.

The proposed esterase activity from HE would be a discontinuous assay. A schematic representation of how this assay would have been performed is shown in figure 3.21.

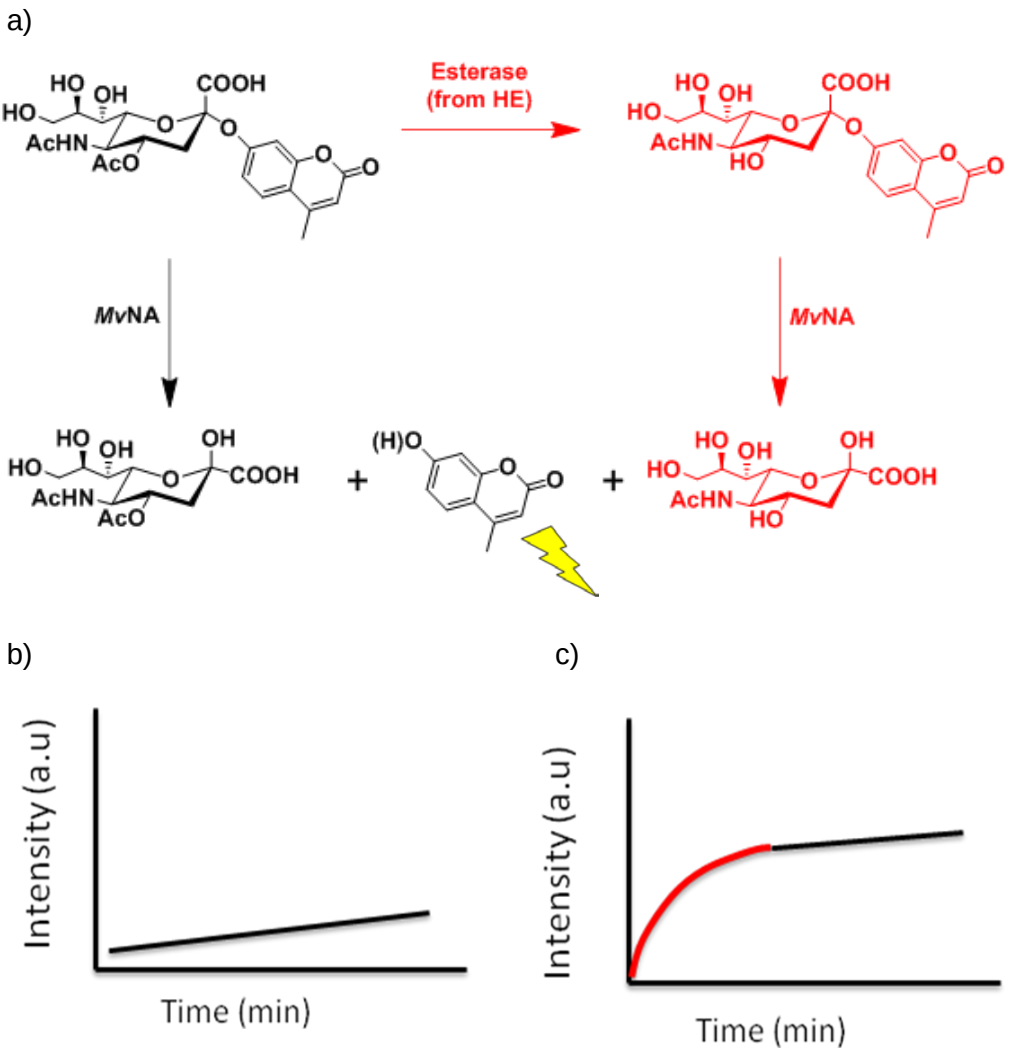

Figure 3.21 Discontinuous assay for esterase activity test from hemagglutinin-esterase Panel a) shows the esterase activity assay using *Mv*NA. The expected result for the negative control reaction is shown in **b)**, where a slow increase in fluorescence versus time is predicted. The esterase activity from HE transforming 4-*O*Ac MUNANA into MUNANA followed by its hydrolysis by *Mv*NA is represented in red on panel a. The expected result from these reactions are shown in **c)**, where a rapid hydrolysis rate slope of MUNANA is predicted followed by a slower rate slope coming from hydrolysis of the remaining 4-*O*Ac MUNANA in the presence of *Mv*NA.

As it was predicted before, *Mv*NA will slowly hydrolyze 4-*O*Ac MUNANA, creating a slow increase in intensity, as it is shown in figure 3.21b. This low rate of hydrolysis tells us there is no MUNANA in the mixture and therefore no esterase activity. Hence, this reaction will constitute our negative control. On the other hand, if 4-*O*Ac MUNANA is incubated for a period of time with HE, it will slowly be converted into MUNANA, then upon addition of *Mv*NA to the reaction mixture, it will rapidly hydrolyzed MUNANA by cleaving the glycosidic bond between the sugar and 4-MU, thus generating a rapid increase in fluorescence. This prediction is shown in figure 3.21c. The rapid hydrolysis of MUNANA by *Mv*NA is predicted as a steep red slope. Once all of the MUNANA formed by the esterase is hydrolyzed, the remaining 4-*O*Ac MUNANA will continue to be hydrolyzed by *Mv*NA at the same rate as the negative control. As a consequence, it will be possible to calculate the esterase activity by measuring the amount of de-acetylation by creating a calibration curve using different concentrations of 4-MU.

REFERENCES

1. Asche, F. and K.H. Roll, *Determinants of inefficiency in Norwegian salmon aquaculture.* Aquaculture Economics & Management, 2013. **17**(3): p. 300-321.

2. Vassdal, T. and H.M.S. Holst, *Technical Progress and Regress in Norwegian Salmon Farming: A Malmquist Index Approach.* Marine Resource Economics, 2011. **26**(4): p. 329-341.

3. Tveterås, S., *Norwegian Salmon Aquaculture and Sustainability: The Relationship Between Environmental Quality and Industry Growth.* Marine Resource Economics, 2002. **17**(2): p. 121-132.

4. Burridge, L., et al., *Chemical use in salmon aquaculture: A review of current practices and possible environmental effects.* Aquaculture, 2010. **306**(1–4): p. 7-23.

5. Krkošek, M., et al., *Declining Wild Salmon Populations in Relation to Parasites from Farm Salmon.* Science, 2007. **318**(5857): p. 1772-1775.

6. Johnsen, B.O. and A.J. Jensen, *The spread of furunculosis in salmonids in Norwegian rivers.* Journal of Fish Biology, 1994. **45**(1): p. 47-55.

7. Toranzo, A.E., B. Magariños, and J.L. Romalde, *A review of the main bacterial fish diseases in mariculture systems.* Aquaculture, 2005. **246**(1–4): p. 37-61.

8. Ellis, A.E., *Innate host defense mechanisms of fish against viruses and bacteria.* Developmental & Comparative Immunology, 2001. **25**(8–9): p. 827-839.

9. Raynard, R.S., A.G. Murray, and A. Gregory, *Infectious salmon anaemia virus in wild fish from Scotland.* Diseases of Aquatic Organisms, 2001. **46**(2): p. 93-100.

10. Falk, K., et al., *Identification and Characterization of Viral Structural Proteins of Infectious Salmon Anemia Virus.* Journal of Virology, 2004. **78**(6): p. 3063-3071.

11. Evensen, O., K.E. Thorud, and Y.A. Olsen, *A morphological study of the gross and light microscopic lesions of infectious anaemia in Atlantic salmon (Salmo salar).* Research in Veterinary Science, 1991. **51**(2): p. 215-222.

12. Cottet, L., et al., *Infectious salmon anemia virus—Genetics and pathogenesis.* Virus Research, 2011. **155**(1): p. 10-19.

13. Thorud, K. and H. Djupvik, *Infectious anaemia in Atlantic salmon (Salmo salar L.).* Bull Eur Assoc Fish Pathol, 1988. **8**(5): p. 109-111.

14. Kibenge, B.F.S., et al., *Isolation and identification of infectious salmon anaemia virus (ISAV) from Coho salmon in Chile.* Dis.Aquat.Organ., 2001. **45**(1): p. 9-18.

15. Mullins, E.J., B.D. Groman, and D. Wadowska, *Infectious salmon anaemia in salt water Atlantic salmon (Salmo salar L.) in New Brunswick, Canada.* 1998. **18**(4): p. 110-114.

16. Falk, K., et al., *Characterization of infectious salmon anemia virus, an orthomyxo-like virus isolated from Atlantic salmon (Salmo salar L.).* Journal of Virology, 1997. **71**(12): p. 9016-23.

17. Dannevig, B., K. Falk, and C. Press, *Propagation of infectious salmon anaemia (ISA) virus in cell culture.* Veterinary research, 1995. **26**(5-6): p. 438-442.

18. Aspehaug, V., et al., *Characterization of the Infectious Salmon Anemia Virus Fusion Protein.* Journal of Virology, 2005. **79**(19): p. 12544-12553.

19. Hellebø, A., et al., *Infectious Salmon Anemia Virus Specifically Binds to and Hydrolyzes 4-O-Acetylated Sialic Acids.* Journal of Virology, 2004. **78**(6): p. 3055-3062.

20. Trilla, A., G. Trilla, and C. Daer, *The 1918 "Spanish Flu" in Spain.* Clinical Infectious Diseases, 2008. **47**(5): p. 668-673.

21. Rossman, J.S. and R.A. Lamb, *Influenza virus assembly and budding.* Virology, 2011. **411**(2): p. 229-236.

22. Klenk, E., *Neuraminsäure, das Spaltprodukt eines neuen Gehirnlipoids*, in *Hoppe-Seyler´s Zeitschrift für physiologische Chemie*. 1941. p. 50.

23. Blix, F.G., A. Gottschalk, and E. Klenk, *Proposed Nomenclature in the Field of Neuraminic and Sialic Acids.* Nature, 1957. **179**(4569): p. 1088-1088.

24. Rutishauser, U. and L. Landmesser, *Polysialic acid in the vertebrate nervous system: a promoter of plasticity in cell-cell interactions.* Trends in Neurosciences, 1996. **19**(10): p. 422-427.

25. Varki, A., *Sialic acids as ligands in recognition phenomena.* The FASEB Journal, 1997. **11**(4): p. 248-55.

26. Wiley, D.C. and J.J. Skehel, *The Structure and Function of the Hemagglutinin Membrane Glycoprotein of Influenza Virus.* Annual Review of Biochemistry, 1987. **56**(1): p. 365-394.

27. Whittle, J.R.R., et al., *Broadly neutralizing human antibody that recognizes the receptor-binding pocket of influenza virus hemagglutinin.* Proceedings of the National Academy of Sciences, 2011. **108**(34): p. 14216-14221.

28. Carr, C.M. and P.S. Kim, *A spring-loaded mechanism for the conformational change of influenza hemagglutinin.* Cell, 1993. **73**(4): p. 823-832.

29. Suzuki, Y., et al., *Sialic Acid Species as a Determinant of the Host Range of Influenza A Viruses.* Journal of Virology, 2000. **74**(24): p. 11825-11831.

30. Angata, T. and A. Varki, *Chemical Diversity in the Sialic Acids and Related α-Keto Acids: An Evolutionary Perspective.* Chemical Reviews, 2002. **102**(2): p. 439-470.

31. Ofek, I. and R. Doyle, *Bacterial Lectins as Adhesins*, in *Bacterial Adhesion to Cells and Tissues*. 1994, Springer US. p. 94-135.

32. Yuyama, Y., et al., *Postnatal Change of Pig Intestinal Ganglioside Bound by Escherichia coli with k99 Fimbriae.* Journal of Biochemistry, 1993. **113**(4): p. 488-492.

33. Jaramillo, L., et al., *Purification and characterization of an adhesin from Pasteurella haemolytica.* Glycobiology, 2000. **10**(1): p. 31-37.

34. Crocker, P.R. and A. Varki, *Siglecs, sialic acids and innate immunity.* Trends in Immunology, 2001. **22**(6): p. 337-342.

35. Crocker, P.R., J.C. Paulson, and A. Varki, *Siglecs and their roles in the immune system.* Nat Rev Immunol, 2007. **7**(4): p. 255-266.

36. Shi, W.-X., et al., *Sialic Acid 9-O-Acetylation on Murine Erythroleukemia Cells Affects Complement Activation, Binding to I-type Lectins, and Tissue Homing.* Journal of Biological Chemistry, 1996. **271**(49): p. 31526-31532.

37. Bevilacqua, M., et al., *Selectins: a family of adhesion receptors.* Cell, 1991. **67**(2): p. 233.

38. Läubli, H., et al., *L-Selectin Facilitation of Metastasis Involves Temporal Induction of Fut7-Dependent Ligands at Sites of Tumor Cell Arrest.* Cancer Research, 2006. **66**(3): p. 1536-1542.

39. Lovely, J.E., et al., *First identification of infectious salmon anaemia virus in North America with haemorrhagic kidney syndrome.* Diseases of aquatic organisms, 1999. **35**(2): p. 145-148.

40. Rimstad, E. and S. Mjaaland, *Infectious salmon anaemia virus.* APMIS, 2002. **110**(4): p. 273-282.

41. Mjaaland, S., et al., *Genomic characterization of the virus causing infectious salmon anemia in Atlantic salmon (Salmo salar L.): an orthomyxo-like virus in a teleost.* Journal of Virology, 1997. **71**(10): p. 7681-6.

42. Krossøy, B., et al., *Cloning and identification of the infectious salmon anaemia virus haemagglutinin.* Journal of General Virology, 2001. **82**(7): p. 1757-1765.

43. Müller, A., et al., *Structural and functional analysis of the hemagglutinin-esterase of infectious salmon anaemia virus.* Virus Research, 2010. **151**(2): p. 131-141.

44. Tetsuka, M., et al., *The Basic Residues in the Membrane-Proximal C-Terminal Tail of the Rat Melanin-Concentrating Hormone Receptor 1 Are Required for Receptor Function.* Endocrinology, 2004. **145**(8): p. 3712-3723.

45. Okamoto, Y., J.D. Bernstein, and S. Shikano, *Role of C-terminal Membrane-proximal Basic Residues in Cell Surface Trafficking of HIV Coreceptor GPR15 Protein.* Journal of Biological Chemistry, 2013. **288**(13): p. 9189-9199.

46. Rimstad, E., et al., *Characterization of the Infectious Salmon Anemia Virus Genomic Segment That Encodes the Putative Hemagglutinin.* Journal of Virology, 2001. **75**(11): p. 5352-5356.

47. Kristiansen, M., et al., *Characterization of the receptor-destroying enzyme activity from infectious salmon anaemia virus.* Journal of General Virology, 2002. **83**(11): p. 2693-2697.

48. Hovland, T., et al., *Observation of infectious salmon anaemia virus in Atlantic salmon, Salmo salar L.* Journal of Fish Diseases, 1994. **17**(3): p. 291-296.

49. Nylund, A., et al., *Replication of the infectious salmon anaemia virus (ISAV) in rainbow trout, Oncorhynchus mykiss (Walbaum).* Journal of Fish Diseases, 1997. **20**(4): p. 275-279.

50. Mann, M. and O.N. Jensen, *Proteomic analysis of post-translational modifications.* Nat Biotech, 2003. **21**(3): p. 255-261.

51. Parekh, R.B. and C. Rohlff, *Post-translational modification of proteins and the discovery of new medicine.* Current Opinion in Biotechnology, 1997. **8**(6): p. 718-723.

52. Jenkins, N. and E.M.A. Curling, *Glycosylation of recombinant proteins: Problems and prospects.* Enzyme and Microbial Technology, 1994. **16**(5): p. 354-364.

53. Wells, L., K. Vosseller, and G.W. Hart, *Glycosylation of Nucleocytoplasmic Proteins: Signal Transduction and O-GlcNAc.* Science, 2001. **291**(5512): p. 2376-2378.

54. Munro, S., *An investigation of the role of transmembrane domains in Golgi protein retention.* The EMBO Journal, 1995. **14**(19): p. 4695-4704.

55. in *Essentials of Glycobiology*, A. Varki, et al., Editors. 2009: Cold Spring Harbor NY.

56. Shi, X.D., J., *Protein N-Glycosylation in the Baculovirus-Insect Cells System.* Current Drug Targets, 2007. **8**(10): p. 1116-1125.

57. Schäffer, C., M. Graninger, and P. Messner, *Prokaryotic glycosylation.* PROTEOMICS, 2001. **1**(2): p. 248-261.

58. Moens, S. and J. Vanderleyden, *Glycoproteins in prokaryotes.* Archives of Microbiology, 1997. **168**(3): p. 169-175.

59. Brimer, C.D. and T.C. Montie, *Cloning and Comparison of fliC Genes and Identification of Glycosylation in the Flagellin ofPseudomonas aeruginosa a-Type Strains.* Journal of Bacteriology, 1998. **180**(12): p. 3209-3217.

60. Leclerc, G., S.P. Wang, and B. Ely, *A new class of Caulobacter crescentus flagellar genes.* Journal of Bacteriology, 1998. **180**(19): p. 5010-5019.

61. Doig, P., et al., *Characterization of a post-translational modification of Campylobacter flagellin: identification of a sero-specific glycosyl moiety.* Molecular Microbiology, 1996. **19**(2): p. 379-387.

62. Szymanski, C.M., et al., *Evidence for a system of general protein glycosylation in Campylobacter jejuni.* Molecular Microbiology, 1999. **32**(5): p. 1022-1030.

63. Young, N.M., et al., *Structure of the N-Linked Glycan Present on Multiple Glycoproteins in the Gram-negative Bacterium, Campylobacter jejuni.* Journal of Biological Chemistry, 2002. **277**(45): p. 42530-42539.

64. Hsieh, P. and P.W. Robbins, *Regulation of asparagine-linked oligosaccharide processing. Oligosaccharide processing in Aedes albopictus mosquito cells.* Journal of Biological Chemistry, 1984. **259**(4): p. 2375-82.

65. Stollar, V., et al., *Sialic acid contents of Sindbis virus from vertebrate and mosquito cells: Equivalence of biological and immunological viral properties.* Virology, 1976. **69**(1): p. 104-115.

66. Jarvis, D.L., *Developing baculovirus-insect cell expression systems for humanized recombinant glycoprotein production.* Virology, 2003. **310**(1): p. 1-7.

67. Burda, P. and M. Aebi, *The dolichol pathway of N-linked glycosylation.* Biochimica et Biophysica Acta (BBA) - General Subjects, 1999. **1426**(2): p. 239-257.

68. Herscovics, A., *Structure and function of Class I α1,2-mannosidases involved in glycoprotein synthesis and endoplasmic reticulum quality control.* Biochimie, 2001. **83**(8): p. 757-762.

69. Wilson, I.B.H., et al., *Core α1,3-fucose is a key part of the epitope recognized by antibodies reacting against plant N-linked oligosaccharides and is present in a wide variety of plant extracts.* Glycobiology, 1998. **8**(7): p. 651-661.

70. Tretter, V., et al., *Fucose α1,3-Linked to the Core Region of Glycoprotein N-Glycans Creates an Important Epitope for IgE from Honeybee Venom Allergic Individuals.* International Archives of Allergy and Immunology, 1993. **102**(3): p. 259-266.

71. Parker, M.D., D. Yoo, and L.A. Babiuk, *Expression and secretion of the bovine coronavirus hemagglutinin-esterase glycoprotein by insect cells infected with recombinant baculoviruses.* Journal of Virology, 1990. **64**(4): p. 1625-1629.

72. Kazi, L., et al., *Expression of Hemagglutinin Esterase Protein from Recombinant Mouse Hepatitis Virus Enhances Neurovirulence.* Journal of Virology, 2005. **79**(24): p. 15064-15073.

73. Müller, A., et al., *Heterologous expression and purification of the infectious salmon anemia virus hemagglutinin esterase.* Protein Expression and Purification, 2008. **62**(2): p. 206-215.

74. Haase, S., et al., *Genetic Engineering of Baculoviruses*. Current Issues in Molecular Virology - Viral Genetics and Biotechnological Applications. 2013.

75. Rohrmann, G., *Introduction to the baculoviruses, their taxonomy, and evolution.* Baculovirus Molecular Biology : Second Edition, 2011.

76. Rohrmann, G.F., *Polyhedrin Structure.* Journal of General Virology, 1986. **67**(8): p. 1499-1513.

77. Moscardi, F., *Assessment of the application of baculoviruses for control of lepidoptera.* Annual Review of Entomology, 1999. **44**(1): p. 257-289.

78. Lacey, L.U., T., *Biological control of codling moth (Cydia pomonella, lepidoptera: Tortricidae) and its role in integrated pest management, with emphasis on entomopathogens* Vedalia, 2005. **12**(1): p. 33-60.

79. Sun, X.-l. and H.-y. Peng, *Recent advances in biological control of pest insects by using viruses in China.* Virologica Sinica, 2007. **22**(2): p. 158-162.

80. Taticek, R.A. and M.L. Shuler, *Effect of elevated oxygen and glutamine levels on foreign protein production at high cell densities using the insect cell-*

baculovirus expression system. Biotechnology and Bioengineering, 1997. **54**(2): p. 142-152.

81. Hu, Y.-c., *Baculovirus as a highly efficient expression vector in insect and mammalian cells.* Acta Pharmacol Sin, 2005. **26**(4): p. 405-416.

82. Radner, S., et al., *Transient transfection coupled to baculovirus infection for rapid protein expression screening in insect cells.* Journal of Structural Biology, 2012. **179**(1): p. 46-55.

83. Xu, J., et al., *Soaking RNAi-mediated modification of Sf9 cells for baculovirus expression system by ectopic expression of Caenorhabditis elegans SID-1.* Applied Microbiology and Biotechnology, 2013. **97**(13): p. 5921-5931.

84. Chen, Y.-R., et al., *The Transcriptome of the Baculovirus Autographa californica Multiple Nucleopolyhedrovirus in Trichoplusia ni Cells.* Journal of Virology, 2013. **87**(11): p. 6391-6405.

85. Krammer, F., et al., *Trichoplusia ni cells (High FiveTM) are highly efficient for the production of influenza A virus-like particles: a comparison of two insect cell lines as production platforms for influenza vaccines.* Molecular Biotechnology, 2010. **45**(3): p. 226-234.

86. R. Hollister, J. and D.L. Jarvis, *Engineering lepidopteran insect cells for sialoglycoprotein production by genetic transformation with mammalian β1,4-galactosyltransferase and α2,6-sialyltransferase genes.* Glycobiology, 2001. **11**(1): p. 1-9.

87. Seo, N.-S., J.R. Hollister, and D.L. Jarvis, *Mammalian Glycosyltransferase Expression Allows Sialoglycoprotein Production by Baculovirus-Infected Insect Cells.* Protein Expression and Purification, 2001. **22**(2): p. 234-241.

88. Passarelli, A.L. and L.A. Guarino, *Baculovirus late and very late gene regulation.* Current Drug Targets, 2007. **8**(10): p. 1103-1115.

89. Guarino, L.A., et al., *A Virus-Encoded RNA Polymerase Purified from Baculovirus-Infected Cells.* Journal of Virology, 1998. **72**(10): p. 7985-7991.

90. Todd, J.W., A.L. Passarelli, and L.K. Miller, *Eighteen baculovirus genes, including lef-11, p35, 39K, and p47, support late gene expression.* Journal of Virology, 1995. **69**(2): p. 968-74.

91. Carpentier, D.J. and L. King, *The long road to understanding the baculovirus P10 protein.* Virologica Sinica, 2009. **24**(4): p. 227-242.

92. Luckow, V.A. and M.D. Summers, *Trends in the Development of Baculovirus Expression Vectors.* Nat Biotech, 1988. **6**(1): p. 47-55.

93. Fraser, M.J., *Expression of eucaryotic genes in insect cell cultures.* In Vitro Cellular & Developmental Biology, 1989. **25**(3): p. 225-235.

94. Kitts, P.A., M.D. Ayres, and R.D. Possee, *Linearization of baculovirus DNA enhances the recovery of recombinant virus expression vectors.* Nucleic Acids Research, 1990. **18**(19): p. 5667-5672.

95. Kost, T.A., J.P. Condreay, and D.L. Jarvis, *Baculovirus as versatile vectors for protein expression in insect and mammalian cells.* Nat Biotech, 2005. **23**(5): p. 567-575.

96. Kitts, P.A. and R.D. Possee, *A method for producing recombinant baculovirus expression vectors at high frequency.* BioTechniques, 1993. **14**(5): p. 810-817.

97. Gearing, K.L. and R.D. Possee, *Functional analysis of a 603 nucleotide open reading frame upstream of the polyhedrin gene of Autographa californica nuclear polyhedrosis virus.* Journal of General Virology, 1990. **71**(2): p. 251-262.

98. Watson, J.N., et al., *Mutagenesis of the Conserved Active-Site Tyrosine Changes a Retaining Sialidase into an Inverting Sialidase†.* Biochemistry, 2003. **42**(43): p. 12682-12690.

99. DePristo, M.A., M.M. Zilversmit, and D.L. Hartl, *On the abundance, amino acid composition, and evolutionary dynamics of low-complexity regions in proteins.* Gene, 2006. **378**(0): p. 19-30.

100. Xue, H.Y. and D.R. Forsdyke, *Low-complexity segments in Plasmodium falciparum proteins are primarily nucleic acid level adaptations.* Molecular and Biochemical Parasitology, 2003. **128**(1): p. 21-32.

101. Karlin, S., et al., *Amino acid runs in eukaryotic proteomes and disease associations.* Proceedings of the National Academy of Sciences, 2002. **99**(1): p. 333-338.

102. Cortés, A., *A chimeric Plasmodium falciparum Pfnbp2b/Pfnbp2a gene originated during asexual growth.* International Journal for Parasitology, 2005. **35**(2): p. 125-130.

103. Huntley, M. and G.B. Golding, *Evolution of Simple Sequence in Proteins.* Journal of Molecular Evolution, 2000. **51**(2): p. 131-140.

104. Ramasarma, T., et al., *Transmembrane Domains*, in *eLS*. 2001, John Wiley & Sons, Ltd.

105. Kibenge, F.S.B., et al., *Growth of infectious salmon anaemia virus in CHSE-214 cells and evidence for phenotypic differences between virus strains.* Journal of General Virology, 2000. **81**(1): p. 143-150.

106. Blake, S., et al., *Genomic relationships of the North American isolate of infectious salmon anemia virus (ISAV) to the Norwegian strain of ISAV.* Diseases of Aquatic Organisms, 1999. **35**(2): p. 139-144.

107. Kibenge, F., et al., *Infectious salmon anaemia virus (ISAV) isolated from the ISA disease outbreaks in Chile diverged from ISAV isolates from Norway around 1996 and was disseminated around 2005, based on surface glycoprotein gene sequences.* Virology Journal, 2009. **6**(1): p. 88.

108. Clouthier, S.C., et al., *Genomic organization of infectious salmon anaemia virus.* Journal of General Virology, 2002. **83**(2): p. 421-428.

109. Wolf, A., et al., *A hemagglutinin-esterase-expressing salmonid alphavirus replicon protects Atlantic salmon (Salmo salar) against infectious salmon anemia (ISA).* Vaccine, 2013. **31**(4): p. 661-669.

110. Fourrier, M., et al., *Deletions in the highly polymorphic region (HPR) of infectious salmon anaemia virus HPR0 haemagglutinin–esterase enhance viral fusion and influence the interaction with the fusion protein.* Journal of General Virology, 2014. **95**(Pt 5): p. 1015-1024.

111. Aamelfot, M., et al., *Characterisation of a monoclonal antibody detecting Atlantic salmon endothelial and red blood cells, and its association with the infectious salmon anaemia virus cell receptor.* Journal of Anatomy, 2013. **222**(5): p. 547-557.

112. Schrag, J.D., et al., *The Structure of Calnexin, an ER Chaperone Involved in Quality Control of Protein Folding.* Molecular Cell, 2001. **8**(3): p. 633-644.

113. Walter, P. and D. Ron, *The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation.* Science, 2011. **334**(6059): p. 1081-1086.

114. Smith, M.H., H.L. Ploegh, and J.S. Weissman, *Road to Ruin: Targeting Proteins for Degradation in the Endoplasmic Reticulum.* Science, 2011. **334**(6059): p. 1086-1090.

115. Gauss, R., et al., *A Complex of Pdi1p and the Mannosidase Htm1p Initiates Clearance of Unfolded Glycoproteins from the Endoplasmic Reticulum.* Molecular Cell, 2011. **42**(6): p. 782-793.

116. Clerc, S., et al., *Htm1 protein generates the N-glycan signal for glycoprotein degradation in the endoplasmic reticulum.* The Journal of Cell Biology, 2009. **184**(1): p. 159-172.

117. Knowles, J.R. and C.A. Parsons, *Proximity Effect in Catalysed Systems: a Dramatic Effect on Ester Hydrolysis.* Nature, 1969. **221**(5175): p. 53-54.

118. Butler, A.R. and I.H. Robertson, *Nucleophilic and general base catalysis by pyridine and methylpyridines in the hydrolysis of aryl acetates.* Journal of the Chemical Society, Perkin Transactions 2, 1975(6): p. 660-663.

119. Drzeniek, R., *Viral and Bacterial Neuraminidases*, in *Current Topics in Microbiology and Immunology*, W. Arber, et al., Editors. 1972, Springer Berlin Heidelberg. p. 35-74.

120. Amano, Y. and Q. Cheng, *Detection of influenza virus: traditional approaches and development of biosensors.* Analytical and Bioanalytical Chemistry, 2005. **381**(1): p. 156-164.

121. Woods, J.M., et al., *4-Guanidino-2,4-dideoxy-2,3-dehydro-N-acetylneuraminic acid is a highly effective inhibitor both of the sialidase (neuraminidase) and of growth of a wide range of influenza A and B viruses in vitro.* Antimicrobial Agents and Chemotherapy, 1993. **37**(7): p. 1473-1479.