Christine Sachse

Tetrahymena thermophila: An expression platform for the production of viral antigens.

-2010-

Biology

Tetrahymena thermophila: An expression platform for the production of viral antigens.

Thesis submitted in partial fulfilment

of the requirements of the degree

Doctor rer. nat. of the

Faculty for Mathematics and Natural Sciences

Chair in Biology and Biology Didactics - Zoology

Bergische Universität Wuppertal, Germany

submitted by

Christine Sachse

Münster, Germany

-2010-

Diese Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20101207-123231-9
[http://nbn-resolving.de/urn/resolver.pl?urn=urn:nbn:de:hbz:468-20101207-123231-9]

Dean of Faculty:	Prof. Dr. P. Wiesen
First reviewer:	Prof. Dr. G. Preisfeld
Second reviewer:	Dr. M. W. Hartmann

for my parents and Frank

Acknowledgements

Though the following dissertation is an individual work, I could never have reached the heights or explored the depths without the help, support guidance and efforts of a lot of people. Firstly, I want to express my special gratitude to my supervisor Prof. Dr. Gela Preisfeld. She not only gave me the opportunity to write this dissertation, she was also a constructive critic in the best sense.

This thesis would not have been written, and I would not have spent three beautiful years at the Cilian AG, without the cooperation with Sanofi Pasteur (France), that funded all the projects. Above all, I want to thank the two persons, Dr. Marcus W. W. Hartmann and Remco P. Brandt, who hired me for the project, and then stood beside me during the whole period. They always had a lot of genuine interest, and understanding – when I needed it for my work. They encouraged me to not only grow as an experimentalist and a biologist but also as an instructor and an independant thinker. I am not sure many graduant students are given the opportunity to work with such independance.

I also gratefully acknowledge the support of the Cilian AG that I have received while working on this project. I want to thank in general all fellows, members of the Cilian AG, for creating such a friendly and stimulating context that widened my scientific understanding of the biotechnology of ciliates in many aspects. Especially, I want to thank Kathrin Schildmann and Jan Roßdorf for supporting me so much.

I believe I gained more ciliate knowledgement through oral communication than from all the papers and books I read. Special thanks in this respect go to Dr. Ulrike Bockau, who invested a lot of energy in reading my thesis. For this, and especially for a plenty of enlightening discussions. Furthermore, I thank Dr. Ulrike Bockau for giving me her friendship and also for providing for some much needed humor and entertainment in what could have otherwise been a somewhat stressfull office environment. Also I thank all my friends, that supported me during the whole phase of writing the thesis.

Finally, and most importantly, my very special thanks to the two persons whom I owe everything I am today, my parents, Ute and Reiner Sachse. Their unwavering faith and confidence in my abilities and in me is what has shaped me to the person I am today. Thank you for everything. My thanks also go to my grandparents, Lisa Zeroba, Hella und Willi Sachse, that supported me during my whole life, even if my grandfather left us in 2008. Also I thank Rudi for showing me so much joy of living. I would also like to thank

my significant other, Frank van Loh, for his love, support and indefatigability to sustain my most overblown sentimentality.

Parts of the results of this thesis were applied for a patent:

Hartmann, M. W. W., Sachse C., Apelt J., Bockau, U.

System for the heterologous expression of a viral protein in a ciliate host cell.

Filed: 17. June 2009

Patent application number: GB0910357.3

Contents

List of abbreviations

aa	amino acids
APS	Ammonium persulfate
ARP	autonomously replicating pieces
β-hex	β-hexasominidase
β-NADP	β-Nicotinamide adenine dinucleotide phosphate
bp	base pair
bsdR	blasticidin resistance
BTU2	beta tubulin gene 2
C	cells
°C	degree centigrade
Cbs	chromosome breakage sequence
CET	conjugant electrotransformation
cDNA	complementary DNA
cGMP	current good manufacturing practice
CHT	ceramic hydroxy apatite
CmR	chloramphenicol resistance
Con A	Concanavalin A
DCG	dense core granule
DHFR-TS	bifunctional dihydrofolate reductase and thymidilate synthase
DNA	deoxyribonucleic acid
DNase I	Deoxyribonuclease I
dNTP	deoxynucleotide
DMSO	Dimethylsulfoxide
E. coli	*Escherichia coli*
EDTA	Ethylenediamine-tetraacetic acid disodium salt
e.g.	*exempli gratia* (for example)
ER	Endoplasmatic Reticulum
ESTs	expressed sequence tags
FL HA	Florida hemagglutinin
g	gram
x g	centrifugational force
Glc	Glucose
GlcNAc	N-Acetylglucosamin
GRAS	generally regarded as safe
Grl(p)	granule lattice (protein)
H4-1	Histon Gene 4-1
HA	hemagglutinin
HRP	Horse Raddish Peroxydase
ICDH	Isocitrate Dehydrogenase
IES	internally eliminated sequence
kDa	kilo dalton
L	liter
LB-medium	Luria-Bertani-medium
m	meter
MAC	Macronucleus
Man	Mannose
Mb	megabase
MDCK cells	Madin-Darby canine kidney cells
mg	milligram
MIC	Micronucleus
min	minute
mL	milliliter
mm	millimeter
mM	millimolar
mRNA	messenger RNA
MTs	metallothioneins
MTT1	Metallothioneine Gene T1
NA	neuraminidase
NaOH	Sodium hydroxide
NC HA	New Caledonia hemagglutinin
NC531 HA	aa1-531 of New Caledonia hemagglutinin
NC344 HA	aa1-344 of New Caledonia hemagglutinin
neo	neomycin
NIH	National Institutes of Health
p.A.	per analysis
P. pastoris	*Pichia pastoris*
PBS(-T)	Phosphate buffered saline (-Tween)
PCR	Polymerase chain reaction

pH	negative logarithm (based ten) of the molar concentration of dissolved hydronium ions
PLA_1	Phospholipase A_1(AS 1-319)
PNGase F	Peptide-N Glycosidase F
rDNA	ribosomal DNA
RNA	Ribonucleic acid
RNAi	RNA interference
rRNA	ribosomal RNA
rpm	rounds per minute
S. cerevisiae	*Saccharomyces cerevisiae*
SDS	sodium dodecylsulfate
SDS-PAGE	sodium dodecylsulfate polyacrylamide gel electrophoresis
sec	second
SN	supernatant
SPP	supplemented proteose peptone
T. thermophila	*Tetrahymena thermophila*
TAE	Tris-acetate-EDTA
TEMED	Tetramethylenediamine
TMD	transmembrane domain
Tris	Trishydroxymethylaminomethane
tRNA	transfer RNA
UR HA	Uruguay hemagglutinin
UV	ultraviolet
µg	microgram
µl	microliter
µm	micrometer
V	volt
VERO cells	African green monkey kidney cells
VLP	virus like particle
v/v	volume per volume
w/v	weight per volume

Symbols for amino acids

Alanine	Ala	A	Leucine	Leu	L
Arginine	Arg	R	Lysine	Lys	K
Asparagine	Asn	N	Methionine	Met	M
Aspartic acid	Asp	D	Phenylalanine	Phe	F
Cysteine	Cys	C	Proline	Pro	P
Glutamine	Gln	Q	Serine	Ser	S
Glutamic acid	Glu	E	Threonine	Thr	T
Glycine	Gly	G	Tryptophan	Trp	W
Histidine	His	H	Tyrosine	Tyr	Y
Isoleucine	Ile	I	Valine	Va	V

1. Introduction

Influenza is an infectious disease. The most common symptoms are chills, fever, sore throat, muscle pains, severe headache, coughing, fatigue and general discomfort. In serious cases, influenza causes pneumonia, which can be fatal, particularly for the young and elderly. Typically, influenza is transmitted through the air by coughs or sneezes, creating aerosols containing the virus. Furthermore, influenza can be transmitted by direct contact with bird droppings or nasal secretions, or through contact with contaminated surfaces. Influenza spreads around the world in seasonal pandemics, resulting in deaths of between 250,000 and 500,000 people every year, up to millions in some pandemic years[1]. The most effective way to prevent the disease or severe outcomes from the illness is vaccination.

1.1 Expression systems for the production of recombinant vaccines

The success of a vaccine is governed by three critical factors, namely safety, cost and efficacy. Ideally, vaccines should be free of sight effects, lack residual pathogenicity and should have no potential for reversion to virulence. With respect to cost, all expenditures related to research, licensing and production must be taken into account and balanced against market demand. In terms of efficacy, they should offer protection from disease, which depends in large part on their ability to trigger immune response appropriate to the pathogen targeted. With the advent of molecular cloning techniques in the 1970s, large-scale production of recombinant subunit antigens became possible using heterologous protein expression systems, such as *Escherichia coli* (*E. coli*) and yeast. The ability to produce large amounts of purified antigen in non-pathogenic, single-celled organisms had advantages with respect to cost and safety, and led quickly to the development of successful vaccines for humans and animals[2,3].

By the mid-1980s, genetic engineering techniques made an alternative approach possible, namely, alteration of genes for virulence determinants to produce rationally attenuated pathogens. These altered organisms could be "marked" in such a way that they could be distinguished from wild type pathogens, thus making them useful in eradication programs. Shortly thereafter, genetic engineering techniques facilitated the application of viral and bacterial vectors as delivery systems for candidate vaccine antigens. This approach led e.g. to the development of successful vaccines for rabies using vaccinia and canarypox as life vectors for the delivery of the rabies virus G-glycoprotein[4]. Finally, in the 1990s, plasmid expression

vectors (so called "DNA vaccines") encoding microbial antigens offered another promising approach toward vaccine development[5]. Referred to broadly as genetic immunization, this method has been found to be highly effective against a wide range of pathogens, is low in cost and at least theoretically safe[3].

When compared with live agents and DNA vaccines, recombinant subunit antigens have clear advantages from a licensing and regulatory standpoint. Additionally, once expressed, recombinant antigens can be rapidly purified using for instance affinity tags and administered to animals at high concentration levels. Furthermore, the supplementation with new generation vaccine adjuvants, have enhanced the immunity in vertebrate hosts[3]. When using especially the *E. coli* system for the expression of recombinant antigens, missfolding and an incorrect processed protein that lack conformational epitopes required for the production of neutralizing antibodies depicts a problem[6,7]. Similarly, the formation of protein aggregates (inclusion bodies) commonly seen following overexpression in bacterial systems is detrimental to native 3-dimensional structure of the vaccine candidate[6]. While more complex eukaryotic systems (insect and mammalian tissue culture cells) are more accurate from the standpoint of protein folding and processing, they are more cost intensive[3].

Despite its obvious strength and clear success with some antigens, *E. coli* remains problematic, particular for viral and metazoan membrane proteins. Established alternatives to *E. coli* include yeast based expression systems, and cultures of insect cells where overexpression is performed applying baculovirus vectors. Among the fungi, Baker´s yeast (*Saccharomyces cerevisiae*), and methylotropic strains of yeast such as *Pichia pastoris* (*P. pastoris*) are most efficient with respect to yield. Furthermore, yeasts are capable of secreting proteins into the extracellular space, and are clearly superior to bacteria with respect to protein folding. Nevertheless, certain post-translational modifications, in particular N- and O-linked glycosylation patterns, vary with those seen in higher cells, with potential implications for antigenicity and immunogenicity. Additionally, studies with the G-glycoprotein from rabies virus have clearly demonstrated that a majority of the protein expressed in *S. cerevisiae* is insoluble and non-immunogenic[8]. Insect cells on the other hand, provide a more natural environment for the expression of recombinant proteins produced in higher eukaryotes including viral antigens[3], so that the G-glycoprotein expressed in Sf9 insect cells did provide protection[9,10]. Additionally, the expression of hemagglutinin what is applicable for a vaccine, from the influenza A virus was successfully reached using the insect cell/baculovirus production technology[11-14]. More recently, plants have emerged as another

promising alternative since they are the largest and most economical biomass generators. They possess an eukaryotic posttranslational modification machinery similar to that of mammals, and they do not harbor human pathogens[15]. The hemagglutinin produced in *Nicotiana benthamiana* plants applying an avian influenza virus hemagglutinin showed antigenicity *in vitro* and immunogenicity *in vivo*[16,17].

Additionally, virus propagated in mammalian cell-derived tissue culture has been reported to be representative of the natural virus[18]. Hence, mammalian cell-derived influenza vaccine has been shown to provide equivalent or better protection in animal models than egg-derived vaccine[18-20], and in clinical trials has been shown to be both safe and efficacious[21,22]. Three continuous cell lines are currently in use, Madin-Darby canine kidney (MDCK) cells[21-25], African green monkey kidney (VERO) cells[20,25-27], and the human retina PER-C6 cells[28]. MDCK cells are favored because of higher yields[25], but this adherent cell line has to be grown in cost intensive bioreactor infrastructure like in a microcarrier culture[24]. Another disadvantage, using MDCK cells, is the tumorigenicity when injected into immunocompromized mice[29].

None of the expression systems, currently available for recombinant protein production, is ideal. The variety of systems being in development and the large number of basic approaches for the improvement of existing expression platforms shows the necessity of optimization in a broad range and documents the requirement of recombinant vaccines. The number and complexity of the proteins will require an equally portfolio of production methods. For this purpose, a large interest in new production systems exists, which could be used for a multitude of proteins as well as for a unique protein.

1.1.1 Production of influenza vaccines

Influenza A viruses represent a continuous pandemic threat and are responsible for a highly contagious, acute viral respiratory disease, which causes significant morbidity and mortality worldwide each year[30-36]. Influenza disease impacts every year up to 10 % of the world´s population, that is up to 500 millions of people[37]. Vaccination has been the most effective way to prevent the disease. Although, classical egg-derived influenza vaccines can be sufficiently produced in quantities that are needed in an inter-pandemic seasonal influenza situation, they may be insufficient in case of pandemic influenza. Current influenza vaccines include a single isolate of three antigenetically distinct influenza virus subtypes, two influenza A virus strains

(H1N1 and H3N2), and a single influenza B virus isolate. Annual reformulation in an attempt to antigenetically match the circulating strains for each of the three subtypes is a time-consuming process[37]. Furthermore, various studies have shown that the passage of virus in hen´s eggs can alter the composition of the hemagglutinin making it different from the wild type virus[38-40], so that immune responses do not correspond exactly with the structure of the circulating clinical virus making the vaccine less efficacious[19]. This increases the need to produce a recombinant vaccine.

Influenza A virus is a negative-sense single-stranded RNA virus belonging to the *Orthomyxoviridae* family. The virion consists of a host cell-derived envelope, which is lined on the inner side by the matrix protein M1 (Figure 1). On the surface there are multiple copies of hemagglutinin (HA), which are responsible for attachment to the host cell surface receptors, and neuraminidase (NA), which causes the release of new virus particles from the host cell[41]. HA is a trimeric glycoprotein that consists of two subunits, HA1 and HA2. While HA1 is responsible for receptor recognition, HA2 triggers the membrane fusion. The fusion domain or fusion peptide of HA consists of a hydrophobic sequence located at the N-terminus of the HA2 subunit. Under resting conditions at neutral pH, the fusion domains are buried in hydrophobic pockets at the interfaces between monomers in the HA trimer[42]. Under fusion conditions at low pH values the fusion peptide becomes exposed and inserts into the target membrane[43]. The HA2 chain contains a transmembrane domain near its C-terminus that anchors HA in the viral envelope. The transmembrane domain and the fusion peptide are the only segments of HA that interact direct with the lipid bilayers of the host membranes[44]. Nevertheless, both full-length as well as transmembrane domain-truncated HA were tested to induce antibody responses *in vivo*. Although, for both versions, antibody responses were detected, but depending on the influenza strain, there were higher responses for the untruncated version[45].

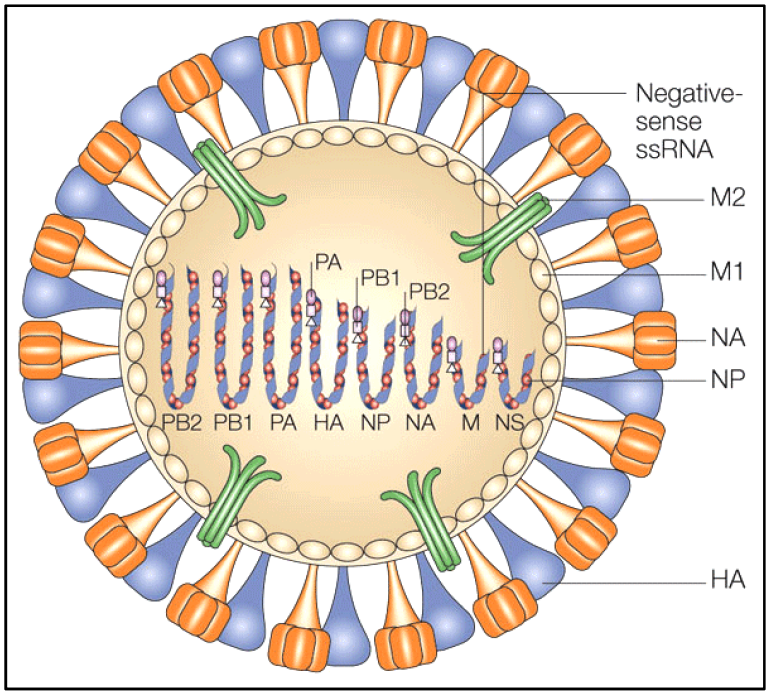

Figure 1: Schematic diagram of an influenza A virus virion.
Two surface glycoproteins, hemagglutinin (HA) and neuraminidase (NA), and the M2 ion-channel protein are embedded in the viral envelope, which is derived from the host´s plasma membrane. The ribonucleoprotein complex comprises a viral RNA segment associated with the nucleoprotein (NP) and three polymerases (PA, PB1 and PB2). The matrix (M1) protein is associated with both ribonucleoprotein and the viral envelope and is coded by the M-gene. This gene also encodes the M2 ion channel The NS gene encodes two different non-structural proteins[46].

Currently, influenza vaccines are generated in embryonated hen's eggs. Virions are harvested from the egg allantoic fluid, chemically inactivated and treated with detergent. Either a whole virion preparation is generated, or the HA and NA proteins are partially purified to produce split-product, subvirion, or subunit vaccines[47]. This system served well for over 50 years, but there are some disadvantages as the need of millions of high quality fertile eggs. Alternatively, cell culture based methods have been established for faster and higher yield production. However, biosafety issues and difficulties with growth and yield of certain isolates may cause constraints during the cell culture based production. Besides whole virus vaccines, subunit vaccines comprising HA and virus like particles (VLPs) are considered as alternative strategies and have been shown successfully to induce neutralizing immune response[48-57]. Recently, the three hemagglutinin proteins needed for the vaccine were produced in a non-transformed, non-tumorigenic continuous cell line (*expres*SF$^{+®}$ insect cells) grown in serum-free medium. Each of the three recombinant HAs were expressed in this insect cell line using a viral vector from baculovirus. The individual HAs are extracted from the cells with buffer and detergent and further purified by using a combination of filtration and column chromatography[11]. This method shows a new recombinant production process.

Nevertheless, high costs due to the patent situation of the system represents a disadvantage of the baculovirus based expression platform.

Recently, the expression system based on the common pond-water ciliate *Tetrahymena thermophila* (*T. thermophila*) has shown considerable promise for the development of recombinant subunit vaccines[58,59]. With hundreds of cilia, dense core granules, and alveolar sacs, its metabolism is geared to the production of membrane and secretory proteins. This feature is unique to ciliates, and distinguishes *Tetrahymena* from all other systems currently used for heterologous protein expression. For all these reasons, *T. thermophila* is predestinated for the production of an influenza vaccine.

1.2 *Tetrahymena thermophila*

Tetrahymena belongs to the phylum Ciliophora and is placed in the class Oligohymenophorea, subclass Hymenostomatia. The genus *Tetrahymena* was named in 1940, and includes several morphologically distinct species but also includes a large group of morphologically indistinguishable sibling species, originally named *T. geleii* and later named *T. pyriformis*. The *T. pyriformis* sibling-species swarm includes some species that lack micronuclei and other, sexual, species with micronuclei. In 1976 the original name of the entire assemblage, *T. pyriformis*, was assigned to one of its amicronucleate member, and the other species were given Latin binomial names like *Tetrahymena thermophila*[60].

T. thermophila is a large and complex cell yet a rather small and simple ciliate. Its normal dimension is about 50 µm in length and 20 µm in maximum width. What is distinctive about ciliates, including *T. thermophila*, is nuclear dualism and the arrangement of structures in the surface region (cortex) of the cell. *T. thermophila* possesses a diploid micronucleus (MIC) possessing two complete haploid genome sets capable of both mitosis and meiosis, plus a macronucleus (MAC) that is made up of multiple copies of a rearranged subset of the complete MIC genome (Figure 2). The germline MIC is mostly inert transcriptionally, whereas the somatic MAC is transcriptionally active[60]. The internal (endoplasmic) structure is unremarkable, except for a hypertrophy of the phagosomal compartment (food vacuoles) and the presence of an osmoregulatory system, the contractile vacuole that also is involved in ionic balance. A complex structural order is found in the cell surface region[61,62]. Underneath the plasma membrane, eight distinct structural systems are presented in an ordered array. From the outside in, these systems are: flattened membrane-enclosed cortical alveoli,

trademarks of the alveolate evolutionary lineage, 18 to 21 longitudinal microtubule bands, a continuous membrane skeleton (epiplasm), and ciliary units (basal bodies) with associated microtubular and fibrillar elements, longitudinal aligned in 18 to 21 rows[60]. Continuing, there are longitudinal arrays of dense-core granules (mucocysts), arrays of mitochondria lined up parallel to the rows of basal bodies and secretory granules[63], flattened lamellae of the endoplasmic reticulum[62], and arrays of small Golgi elements[64], both underneath the mitochondria. Finally, three unique cell-surface landmarks are superimposed on the repeated cortical array. A complex oral apparatus, made up of four compound ciliary elements (hence *Tetra-hymena*) embedded within an elaborate fibrillar substructure, is located near the anterior end of the cell. An elongate cell-anus or cytoproct is situated near the posterior end of the cell along the same ciliary row as the oral apparatus. Two (rarely one or three) contractile vacuole pores lie side by side near the posterior end of the cell, to the cell´s right of the oral-cytoproct meridian and external to the contractile vacuole[60].

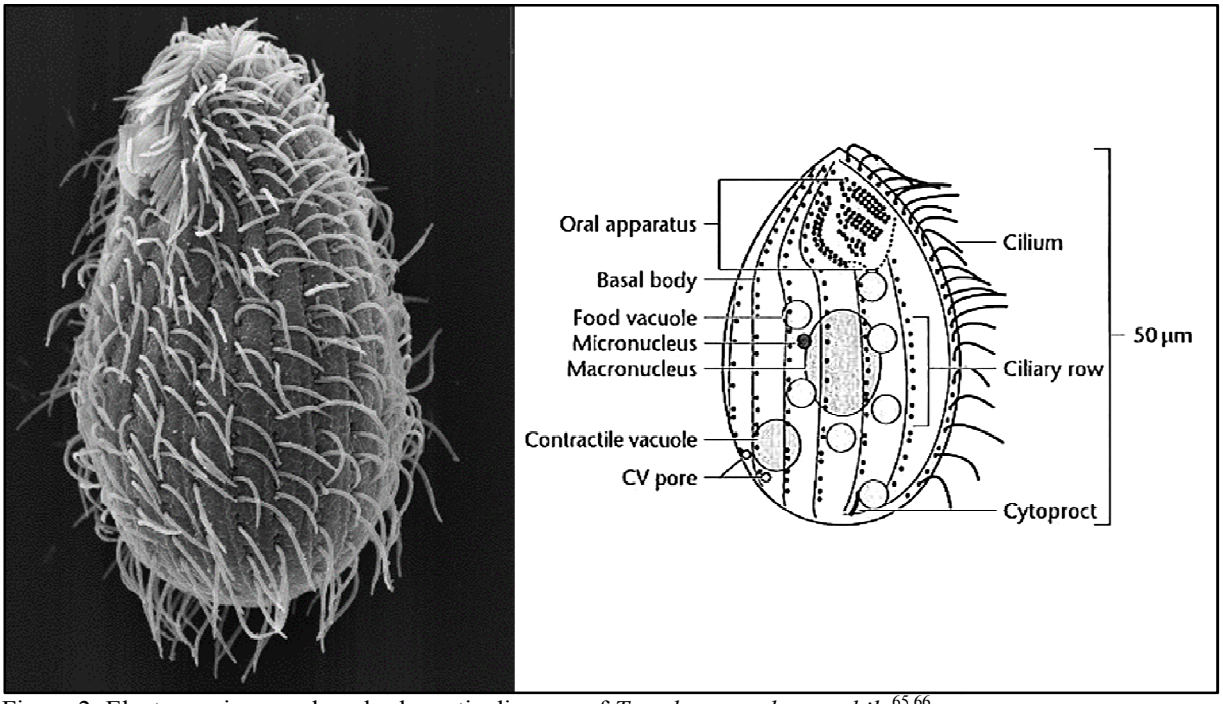

Figure 2: Electron micrograph and schematic diagram of *Tetrahymena thermophila*[65,66].
The schematic diagram shows the anterior end of the cell that is oriented upwards, and the ventral (oral) surface of the cell that faces the viewer. Seven of the total of 18-21 ciliary rows are seen, with basal bodies shown as dots next to longitudinally oriented microtubule bands. Cilia are drawn emerging from the basal bodies of one of the ciliary rows and omitted from the other rows.

The plasma membrane of *Tetrahymena* is continuous with the ciliary membrane but not with the outer alveolar membrane and the inner alveolar membrane which line the cortical alveoli[60]. The plasma and ciliary membranes appear as a typical double layer in cross section,

with the outer layer denser[62]. An external polysaccharide surface coat is indicated by ruthenium red staining[67].

The membrane skeleton of *Tetrahymena*, called epiplasm, is a continuous and apparently amorphous dense layer that underlies the inner membrane of the cortical alveoli. Exposure to detergent at high ionic strength results in extraction of all microtubular structures, leaving a residue consisting primarily of intact membrane skeleton[68]. This cell "ghost" is a closed envelope that still retains the shape of the cell, indicating that the membrane skeletal layer has considerable rigidity, although it must be sufficiently elastic to allow *Tetrahymena* cells to alter their shape dramatically while moving through media of high viscosity[69]. The membrane-skeleton is separated by the circumciliary rings from the terminal plates that penetrate the distal end of the basal bodies. The three together form a structurally continuous layer[70,71].

1.2.1 Inward and outward traffic

Tetrahymena, a ciliated organism, is an attractive model for trafficking studies because it possesses at least three secretory pathways, namely regulated phagocytosis, constitutive secretion of lysosomal enzymes, and regulated exocytosis via dense core granules (mucocysts)[60]. *Tetrahymena* forms large food vacuoles (phagosomes) at the inner end of a complex oral apparatus and defecates their remaining contents at a specialized cytoproct[72]. Lysosomes were characterized long ago in *T. thermophila* as a heterogeneous class of compartments containing lysosomal enzymes, the acid hydrolases, many of which are constitutively secreted into the culture medium[73-75]. In later studies, lysosomal enzymes were identified at two sites – food vacuoles (phagolysosomes) – and bound to the exterior surface of the plasma and ciliary membranes; they were interfered at a third site, believed to be small lysosomal vesicles[76].

Lysosomal enzymes, synthesized at the endoplasmic reticulum (ER), pass through the Golgi apparatus, and are sorted by so far unknown signals to the lysosomes. Vesicles of a pool may fuse with incoming phagosomes to form phagolysosomes, which are intracellularly processed, and whose residual contents finally are egested. Alternatively, vesicles from this pool may fuse with the plasma membrane at so far unknown sites, releasing soluble enzyme forms into the culture medium. By this pathway also membrane-bound forms of lysosomal enzymes may be incorporated into the plasma membrane[76].

The plasma membrane of *Tetrahymena* overlays the alveolar membrane, where it includes invaginations that penetrate the cortical alveoli and the tightly adherent membrane skeleton. These invaginations are traditionally named "parasomal sacs" and are found just anterior to basal bodies in the ciliary rows[77,78]. They are also found near basal bodies of the membranelles and undulating membrane of the oral apparatus[79]. The parasomal sacs have a fuzzy coat on the cytoplasmic side, therefore called coated pits. These coated pits are the only places apart from the cytoproct where the plasma membrane is in regular contact with the internal cytoplasm[80]. The cortical alveoli and the underlying membrane skeleton are interposed[62]. Hence, these coated pits have been postulated to be sites both of exocytosis and of endocytosis[61,81].

Numerous ciliates and some flagellates possess cortically located membrane-bounded organelles that discharge their contents upon stimulation[82]. These *Tetrahymena* dense core granules (DCG; mucocysts) are ellipsoidal membrane-bounded structures, about 1 µm long, with a dense crystalline substructure[83]. Their tips commonly penetrate between cortical alveoli to abut against the plasma membrane, probably with membrane-skeleton interposed[84]. About 4500 of these structures are arrayed in approximately 40 longitudinal files, half of them along the ciliary rows and the other half along "secondary meridians" situated halfway between the ciliary rows at alveolar junctions[61,62]. Extrusion of the contents of *Tetrahymena* DCGs begins with the fusion of DCGs membrane with the plasma membrane at a site marked rosette of intramembraneous particles in the plasma membrane. The contents are expelled after membrane fusion, producing a latticework of filaments connected in a three-dimensional network (mucin)[82]. True mucins are highly glycosylated proteins bearing hydrated carbohydrate chains, whereas *Tetrahymena* mucins seem to contain little or no carbohydrates. Secretion by DCGs exhibits dependence on Ca^{2+} at two distinct stages. The first stage is the triggering of the membrane fusion, where a Ca^{2+} ionophore enter the cell an liberate Ca^{2+} from intracellular stores such as the cortical alveoli[85]. The second stage, the activation of the rearrangement of the proteins of the DCGs, requires extracellular Ca^{2+} in millimolar concentrations. The rearrangement results from Ca^{2+} binding to these proteins subsequent to exocytosis[86]. When a large number of DCGs discharge simultaneously following stimulation, the secretion product merges to produce a thick layer that forms a capsule surrounding the cell[87]. The function of this secretion might be protection against predators. Chemically, the matrix of the DCG of *T. thermophila* is made up of no more than a dozen major proteins, which were named granule lattice proteins (Grlp)[88,89]. Five members were initially identified starting with isolation of the most abundant DCG content[89]. Biochemical analysis of Grl1p

indicated that conformational changes could underlie the crystalline expansion, whereas gene disruption suggested that Grl1p itself might be present throughout the entire core, because no visible ordered structure was formed in its absence[86]. Granules formed in such ΔGRL1 cells did not undergo rapid expansion upon exocytosis[89]. Screening for exocytosis-defective mutants among cells transformed with an antisense library, a novel family member, Grl8, was identified[90,91]. Granules lacking Grl8p, like those lacking Grl1p, had no discernable organized structure. Recent results showed that six GRL genes (GRL1, 3, 4, 5, 7 and 8) each play a non-redundant role in core formation[92]. But only the ΔGRL3, ΔGRL4, and ΔGRL7 lines were each completely defective in Alcian blue-stimulated capsule formation[92].

1.2.2 Genetics of *Tetrahymena*

Ciliates characteristically possess two nuclei, a somatic macronucleus (MAC) and a germline micronucleus (MIC)[93]. *Tetrahymena* reproduces asexually by transverse cell division. During this asexual growth, all gene expression occurs in the MAC, which is thus considered the somatic nucleus. The MAC is polyploid and divides nonmitotically, with apparently random distribution of chromosome copies that behave as if acentromeric. In contrast, the MIC is diploid and contains five pairs of chromosomes, that are metacentric and are distributed mitotically[94-96]. In the course of cell division, phenotypic or allelic assortment takes place. The model for allelic assortment assumes, that daughter cells consists of about 45 copies of each MAC and there is no control on the specific allele distribution to the daughter MAC[97,98]. For any allelic pair in a heterozygote, the end point of this process is a MAC containing only one of the two alleles, what may result in a change in phenotype[99]. The MAC genome derives from that of the MIC, but the two genomes are quite distinct. During MAC differentiation, several types of developmentally programmed DNA rearrangements occur[100]. One rearrangement is the deletion of segments of the MIC genome known as internally eliminated sequences (IESs). Approximately 6,000 IESs are removed, resulting in the MAC genome being an estimated 10 % to 20 % smaller than that of the MIC[101]. A second programmed DNA rearrangement is the site-specific fragmentation at each location of the 15-base pair (bp) chromosome breakage sequence (Cbs)[102-104]. During fragmentation, sections of the MIC genome containing each Cbs, as well as up to 30 bp on either side, are deleted[105]. Telomeres are then added to each new end[106], generating some 180 to 250 MAC chromosomes[107,108]. One process that occurs during MAC differentiation is the amplification of the number of copies of the MAC chromosomes. The minichromosome that contains two mirrorimage

(palindromic) copies of the rRNA genes of the MIC, which encodes the 5.8S, 17S, and 26S rRNAs, is maintained at an average of 9,000 copies per MAC[109]. Six other chromosomes that have been examined are each maintained at an average of 45 copies per MAC[110]. The MIC first moves out to the cell periphery and undergoes mitosis.

The sexual reproduction process is called conjugation. This process depends on three conditions. Firstly, cells of two different mating types must be present. Secondly, cells of both mating types must have attained sexual maturity. Thirdly, cells have to be in a state of incipient starvation, what means that the cells have been shifted into a nutritionally inadequate medium[111]. In mating pairs, each MIC undergoes meiosis, followed by the exchange of haploid products across mating junction (Figure 3). Fusion of the gametes in each cell generates a diploid zygotic nucleus that then undergoes mitosis. One daughter retains the diploid chromosome configuration and functions as the new MIC, and the other daughter differentiates to become the new somatic MAC. The old MAC is apoptotically destroyed. The two conjugating cells then separate, becoming the exconjugants, and begin to divide vegetatively. Using specialized strains or experimental manipulations, the normal developmental program for zygote formation can be perturbed. These alternative pathways have been harnessed to develop true heterokaryons in which the MIC and the MAC are genetically unrelated. Recessive lethal mutations can be stably propagated in the homozygous state in the transcriptionally silent MIC of heterokaryons. Mass mating between mutant heterokaryons generates homogenous populations of progeny that are homozygous for the mutant allele[112].

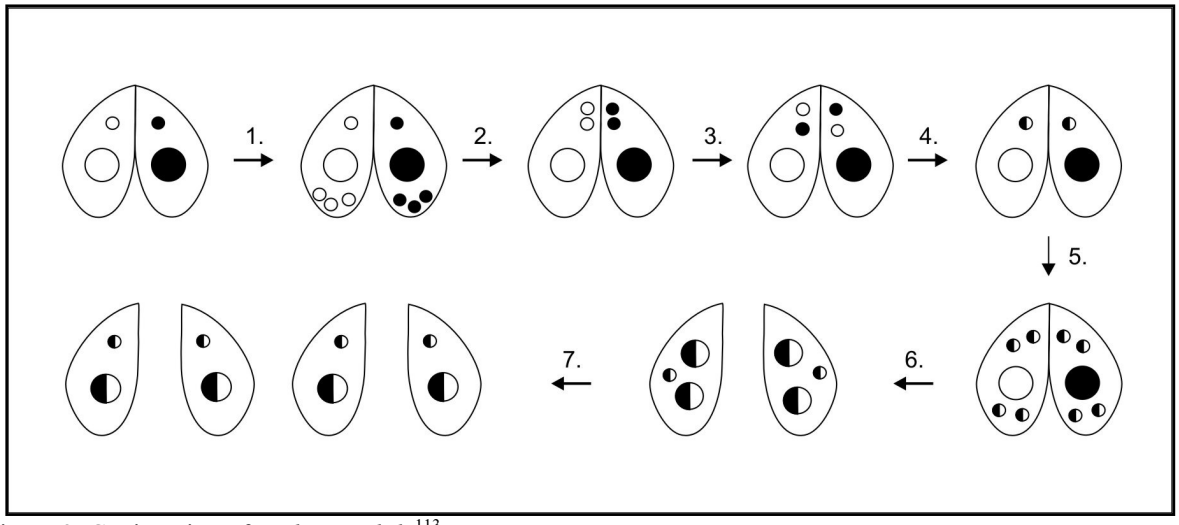

Figure 3: Conjugation of *T. thermophila*[113]
Conjugation starts with two meiotic divisions of the micronucleus (MIC) (1). Three of the haploid nuclei migrate to the posterior section of the cell and are disintegrated. The fourth nucleus undergoes postmeiotic mitosis (2). This leads to the development of two haploid daughter nuclei, followed by a substitution of one nucleus each (3) and the fusion with the other nucleus. A zygotous nucleus (synkaryon) arises (4). The synkaryon divides two times mitotically to four diploid products (5). One of the posterior located nuclei disintegrates, the second becomes the new MIC. The other two nuclei develop to the new MAC/*anlagen* (6). At this point of time the mating pairs divide and the old MAC is eliminated. The first division of these exconjugants leads to a division of the MIC. Four genomic identic daughter cells are formed, with each a MIC and a MAC (karyonides).

Tetrahymena was the first member of the phylum Ciliophora to have its complete somatic (MAC) genome sequenced[107,108]. The amplified MAC genome consists of about 104 Mb, that contain more than 27,000 protein coding genes. Purified MAC genomic DNA of the *Tetrahymena* strain SB210, an inbred strain B derivate that has been extensively used for genetic mapping and for the isolation of mutants, was sequenced by the whole genome shotgun method to 9 x coverage and assembled into 2,955 contigs and 1971 scaffolds that appear to represent a highly accurate and complete draft genome sequence[107,108]. A performed gene prediction resulted in 27,424 putative protein-coding genes, over four times more than the most commonly studied unicellular eukaryotic model organism, *Saccharomyces cerevisiae*, and even more than many metazoans[107,108,114]. A first analysis of expressed sequence tags (ESTs) provided direct evidence that over 40 % of the predicted genes are indeed transcribed into mRNA and has allowed extensive refinement of the gene models[107]. About 15,000 genes match genes of other organisms. In addition to that, the genome analysis also elucidated that a huge number of genes are based on gene duplication mechanisms. This is especially true for genes that play a role in structural complexity, sensing, response to environmental conditions and using of different resources. The sequenced genome analysis of *T. thermophila* once more illustrates the complexity of this single cell eukaryotic microorganism[107,108,115]. Since the initial shotgun assembly, finishing efforts have succeeded in closing numerous sequencing and physical gaps. In addition,

MIC/MAC comparative genomic hybridization (CGH) has identified 763 small scaffolds as probable MIC DNA contaminants. These results in a reduced number of MAC contigs and scaffolds to 1,826 and 1,117, respectively[107]. These efforts also confirmed, that the MAC genome shows a low repetitiveness and the absence of sequences highly related to invasive DNA elements[107,108].

Tetrahymena, like many ciliates, uses an alternative genetic code, in which UGA is the only stop codon and UAA and UAG encode glutamine, resulting in longer potential open reading frames in the genomic sequence[116,117]. *T. thermophila* expresses a tRNA predicted to decode this only *Tetrahymena* stop codon into Selenocystein, the so called 21st amino acid[107,108,118]. Although *Tetrahymena* can use all 64 codons, it does not use all. The most significant aspect of the codon usage in this species is that the AT-rich codons tend to be used more frequently than others[119]. In fact the AT pull is so strong in coding regions that amino acid composition of proteins is shifted toward those coded by codons with high AT codons, as seen in other species with extreme AT bias[120]. Also non-coding regions show high AT contents. The bias in codon usage is larger in highly expressed genes and these genes are less AT rich. The nucleotide positions before start codons is A-rich, and C is significantly avoided after stop codons[121]. Alternative splicing appeared to be rare in *Tetrahymena*, as also reported for *Paramecium*[122].

1.2.3 *Tetrahymena* as a heterologous expression system

The ciliate *Tetrahymena thermophila* has been successfully used as a model system in molecular and cell biology. Fundamental discoveries such as ribozymes, telomeric repeats, telomerases or the function of scan RNA were first studied in this eukaryotic microorganism[100,107,108,123-129]. *T. thermophila* displays many characteristics that predestinates it to create safe vaccines. As a complex eukaryote, it has clear advantages over bacteria in terms of protein folding, processing and targeting. Besides, the apathogenicity as well as the high biological safety of this eukaryotic organism (GRAS-status, generally regarded as safe) are important points. There is no evidence that *Tetrahymena* contain any viruses, endoparasites or DNA sequences with oncogenetic potential. Therefore, the low levels of regulatory cGMP guidelines are, because of minor requirements for technical and personnel equipment, economical important[59]. The differences between humans and ciliates during evolution resulted in significant differences concerning the codon bias and the whole genome organization, indicating that it is implausible that ciliates could be a host for human

pathogens. In addition, nothing has been reported about toxins. *T. thermophila* can be used easily in lab-scale. A method for freezing *Tetrahymena* cells was developed, what is an important item for long term storage[130]. In addition, simple bioreactor infrastructure can be used and several foreign proteins have been expressed, suggesting that *T. thermophila* has the potential to become an excellent expression host[59,131,132]. It can grow very well in a variety of inexpensive axenic media, including skimmed milk, with generation times in the 2 to 3 hours range and optimally 1.4 hours[131]. Typically, cultures remain in rapid-exponential growth phase up to 2 x 10^5 cells per milliliter and can reach cell densities of 2 x 10^6 cells per milliliter or more[133]. Up-scalability is one of the most important problems in producing vaccines, because millions of vials are necessary to start and continue vaccination programs, as for example has already been done for measles, influenza or the hepatitis B virus.

High-level expression of i-antigens from *Ichthyophthirius multifiliis* has recently been achieved in the *Tetrahymena* system using a strong inducible promoter[134]. These proteins are targeted to the plasma membrane in their correct 3-dimensional conformation[59]. Later on, it could be demonstrated that the GPI-anchored circumsporozoite protein of the malaria parasite *Plasmodium falciparum*, was expressed and targeted to the surface of *T. thermophila*[58]. The expression of a functional human protein was shown by the Cilian AG. It was successfully demonstrated that the expression and secretion of a human enzyme, the DNaseI, is possible. This recombinant human enzyme, expressed in *Tetrahymena*, was active, indicating that also disulfide bridges were correctly formed. Furthermore, a detailed N-glycan structure of the recombinant enzyme was presented, illustrating a very consistent glycosylation pattern[113,135]. This glycosylation pattern is more conform, than the pattern of higher eukaryotic cells, than the yeast glycosylation pattern (Figure 4).

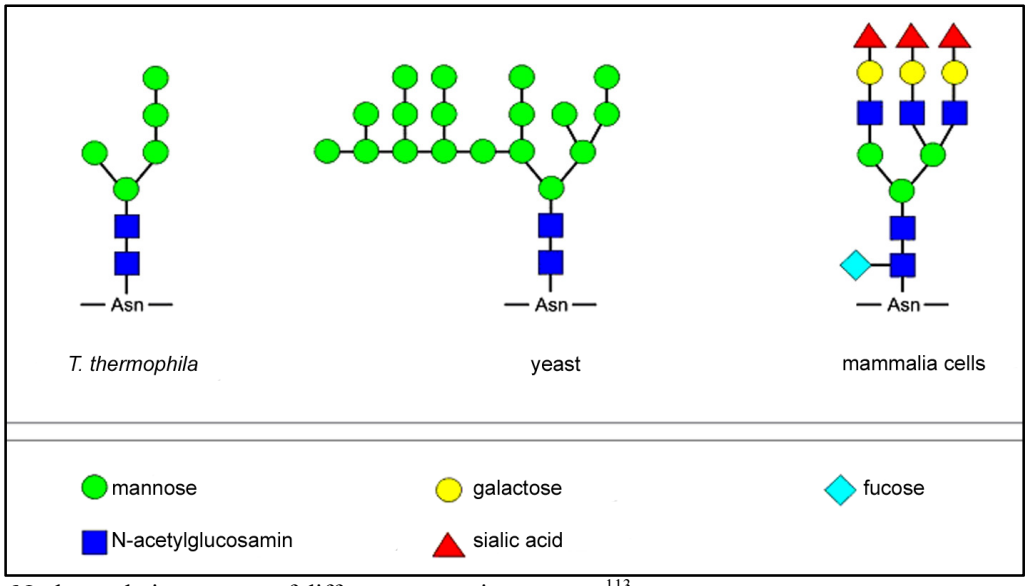

Figure 4: N-glycosylation pattern of different expression systems[113].
The glycosylation pattern of a recombinant protein is expression host specific. While *T. thermophila* produces a consistent mannose terminated oligo-glycosylation pattern, the yeast based expression system shows N-glycans of the high mannose type. Recombinant proteins produced using mammalian cell culture based systems, contain a complex N-glycan feature with terminal galactose, N-acteylglucosamine or sialic acid residues.

The nuclear dimorphism of the ciliates offers different possibilities of manipulating the organisms properties[136]. Altering the phenotype ultimately needs direct or indirect genetic engineering of the vegetative MAC. A basic requirement for the heterologous protein expression, is the technique to transform the expression host. A variety of DNA transformation techniques have been successfully applied to a broad range of organisms. For *Tetrahymena*, microinjection has been used to transform both the polyploid MAC of vegetative *Tetrahymena* and the newly formed MAC *anlagen* of conjugating cells[137]. Electroporation of vegetatively growing cells yielded a lower level of macronuclear transformation[138,139]. However, high frequency macronuclear transformation was obtained when transformation vectors were introduced directly into macronuclear *anlagen* of conjugating cells[140]. Conjugant electrotransformation (CET) yields efficient macronuclear transformation using high copy number vectors containing selectable markers and has been used to develop new vectors[141]. Later on, the efficient method for the mass transformation of the MAC and the MIC in *T. thermophila* using biolistic bombardment was developed[136]. By bombarding mating cells of *Tetrahymena*, it was possible to generate macronuclear transformants at a rate equivalent to the transformation frequencies obtained using CET. Also genetically stable micronuclear transformants and stable macronuclear transformants after allelic assortment were generated[136,142]. The first transformation approaches were based on the use of plasmids that take advantage of the vast amplification of the rDNA gene during *anlagen*/MAC development[143]. However, episomal presence of the plasmid depends on the

presence of antibiotics in the culture medium and the plasmids often recombinate homologously and non-directionally into the endogenous rDNA[113,144]. The stable integration of expression or knock out cassettes into the diploid MIC provides a second method to manipulate the ciliates genome, because after conjugation of two different mating types the old MACs disintegrate and new ones form that carry the new information derived from the recombinant MIC. The advantage is that one obtains stable clones that can be crossed *via* classical Mendelian genetics to combine various properties of different *T. thermophila* strains[145]. But this approach is elaborative and time consuming. Furthermore, it has been shown that scan RNAs derived from the old MAC epigenetically control the genome rearrangement of the new developing MAC[100,129,146]. Thus this RNAi like mechanism may cause problems due to partial deletion of foreign expression cassettes in the developing MAC. Recently, it was developed a knock out/ knock in system (pKOI) that is based on the stable integration into the endogenous gene locus of the bifunctional dihydrofolate reductase and thymidylate synthase (DHFR-TS)[142,147]. This system is based on the MAC manipulation and an allelic or phenotypic assortment process, based on the randomized distribution of the MAC chromosomes units (ARPs) during mitosis[147]. The main advantage of this system is that a stable knock out can be monitored by the complete loss of the DHFR-TS activity, resulting in an auxotrophy for thymidine[142,147]. Thus, this marker system allows the propagation of recombinant *T. thermophila* cells without the rDNA sequences, and transformants after the allelic assortment are stable, and do not need the addition of antibiotics[142,147,148].

Consequently, today there are episomal high-copy vectors including a rDNA origin or integrative vectors available[143,144,149-151]. So far, ciliate expression vectors rely on large double rDNA origin stretches to ensure a stable propagation in *T. thermophila* cells or on large flanking integration sites of non coding regions that are necessary for a proper and efficient homologous recombination into the gene loci of the MIC or MAC. In both cases the AT-richness of these functional DNA sequences cause problems in handling and cloning[142]. It was shown for the vector system of ciliates that a Cre-recombinase system is applicable. It allows the independent construction of expression cassettes using a small plasmid (donor vector) on the one hand side and the preparation of different *T. thermophila* vectors with integration sites (acceptor vector) on the other hand side. As a second step, expression cassettes can be easily shifted from the donor plasmid into various acceptor backbone constructs[142].

For the selection of transformed *Tetrahymena*, the vector backbones code for a paromomycin resistance. Recently, a new neomycin resistance cassette has been developed that was optimized for the codon usage of *T. thermophila*. Using this optimized neo gene, a new resistance marker cassette has been constructed, that results when used for transformation in a ten times more drug resistant transformants[152]. Furthermore, a beta-tubulin marker, that uses the resistance to the mitotic toxin taxol, was established, and a novel blasticidin resistance gene was applied for the selection of transformants[58,142]. The selection of transformants without antibiotics was initialized by the development of thymidine auxotrophic strains. Furthermore, regulation of the expression of a recombinant protein can be regulated using constitutive as well as inducible promoters[59,134,144,153,154]. It was demonstrated that a region approximately 0.9 kb of the cadmium inducible MTT1 promoter was capable of driving high-level gene expression in *T. thermophila*[113,134,142]. Metallothioneins (MTs) are ubiquitous proteins with the capacity to bind heavy metal ions, mainly cadmium, zinc or copper. The MTT5 promoter region for instance, also binds other metal ions efficiently, like lead or zinc and is mainly expressed during conjugation induced under starvation conditions[155]. The MTT2 and MTT4 mainly bind copper[156,157].

To further enhance the expression yield a protease deficient strain was created. Additionally, hypersecretory strains for the acidic hydrolase Phospholipase A_1 were developed[158,159]. The use of proprietary signal and pro peptides enables the targeting[135,147]. The recombinant proteins are led into the cytosol (intracellular expression), into the medium (expression and secretion) or to the surface (surface display)[58,59,113,135]. All these tools together, the ciliate-expression system is a promising molecular and cellular biological system to analyze putative vaccines that can be developed to an excellent platform for production of influenza vaccines.

1.3 Scope of this thesis

In this work, *T. thermophila* should be applied to develop a system for the production of an influenza vaccine.

For this purpose, expression cassettes for three different hemagglutinins of two influenza A and one influenza B virus strain should be designed and assembled, including a gene optimization for the three sequences that adopt the endogenous codons to the codons preferred for the expression in *Tetrahymena*.

Following, these three different hemagglutinins should be produced and characterized using the *T. thermophila* expression platform.

Afterwards, an up-scaling process should be applied for one of these three hemagglutinins including an optimization process for this special protein. Finally, a purification process should be evaluated for one of these hemagglutinins.

2. Materials

2.1 Organisms

2.1.1 *Escherichia coli* (*E. coli*) strains

ElectroMax™ DH10B™ Cells

Genotype: F-*mcr*A Δ (*mrr-hsd*RMS-*mcr*BC) Φ 80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A 1*ara* D139Δ(*ara,leu*)7697 *gal*U *gal*Kλ¯*rps*L *nup*G

The *E. coli* strain was purchased from Invitrogen GmbH, Karlsruhe, Germany.

2.1.2 *Tetrahymena thermophila*-strains

T. thermophila B 1868/4	Genotype: Wild type; mating type: IV
T. thermophila B 1868/7	Genotype: Wild type; mating type: VII
T. thermophila B 2086/1	Genotype: Wild type; mating type: II

The strains B 1868/4 and B 1868/7 of *T. thermophila* were purchased from ATCC, Manassas, VA, USA. The strain B 2086/1 was kindly provided by P. J. Bruns (1998).

2.2 Media for *E. coli*

LB medium	10 g/L peptone 5 g/L yeast extract 5 g/L NaC1 in distilled water, pH 7.5
LB agar plates	15 g/L agar in LB medium

SOC medium	Invitrogen GmbH, Karlsruhe, Germany

2.3 Media for *T. thermophila*

Proteose peptone (PP 1 %)	10 g/L in distilled water
Skimmed milk medium	20 g/L skimmed milk powder 5 g/L yeast extract 1 mL/L Ferrous sulphate/chelate solution (Sigma F0518) 10 g/L glucose monohydrate in distilled water, pH 7.0

The glucose was autoclaved separately and after cooling down it was added to the medium aseptically.

Supplemented proteose peptone medium (SPP)

5 g/L proteose peptone
5 g/L yeast extract
1 mL/L Ferrous sulphate/chelate solution
2 g/L glucose monohydrate
in distilled water, pH 7.0

The glucose was autoclaved separately and after cooling down it was added aseptically to the medium. The proteose peptone was substituted with wheat peptone (5 g/L) for the supplemented wheat peptone medium.

2.4 Buffers and solutions

Main buffers

Blot buffer I	300 mM Tris, pH 10.8 20 % (v/v) methanol p. A. in distilled water

Blot buffer II	25 mM Tris, pH 10 20 % (v/v) methanol p. A. in distilled water
Blot buffer III	25 mM Tris, pH 10 40 mM 6-aminocaproic acid 20 % (v/v) methanol p. A. in distilled water
5-10 % blocking solution	5 or 10 % (w/v) skimmed milk powder or 5 % BSA in PBS-T
Dryls-Solution[160]	1.5 mM sodium citrate 1 mM NaH_2PO_4 1 mM Na_2HPO_4 1.5 mM $CaCl_2$ in distilled water

The $CaCl_2$ solution was autoclaved separately and after cooling down it was added aseptically to the medium.

Chemiluminescent solution	2.5 mM Luminol in DMSO 0.405 mM p-coumaric acid in DMSO 100 mM Tris-HCl, pH 8.5 0.011 % H_2O_2 made up to 20 mL with distilled water
ICDH buffer	0.1 M tri-ethanolamine hydrochloride 4.6 mM isocitric acid 25 mM sodium carbonate 51.3 mM sodium chloride

ICDH solution	9.1 mM β-NADP 0.124 M $MnSO_4$
Phosphate buffered saline (PBS)	2.7 mM KCl 1.5 mM KH_2PO_4 137 mM NaCl 8 mM Na_2HPO_4 in distilled water
PBS-T	0.05 % (v/v) Tween®20 in PBS
RIPA buffer	10 mM Tris-HCl, pH 7.2 150 mM NaCl 5 mM EDTA 7 µM E-64
Stacking gel buffer (4 x)	0.5 M Tris-HCl, pH 6.8 in distilled water
Resolving gel buffer (4 x)	1.5 M Tris-HCl, pH 8.8 in distilled water
SDS running buffer	25 mM Tris-HCl, pH 8.6 192 mM glycine 0.1 % (w/v) SDS in distilled water

SDS sample buffer (4x)	200 mM Tris-HCl, pH 6.8 16 % (w/v) SDS 48 % (v/v) glycerol 8 % (v/v) β-mercaptoethanol 0.04 % (w/v) Coomassie Brilliant Blue R-250 in distilled water
TAE buffer (50 x)	40 mM Tris 0,114 % (v/v) acetic acid p. A. 1 mM EDTA in distilled water, pH 8.9

Buffers for protein chromatography

Phosphate buffer (10 mM)	15.5 mM Na_2HPO_4 4.06 mM $NaH_2PO_4*H_2O$ in distilled water, pH 7.4
Phosphate buffer (150 mM)	136 mM Na_2HPO_4 14 mM $NaH_2PO_4H_2O$ in distilled water, pH 7.5
Running buffer 1	10 mM phosphate buffer 5 % glycerol, pH 7.4
Elution buffer 1	Running buffer 1 plus 150 mM NaCl
Elution buffer 2	Running buffer 1 plus 1 M NaCl

Loading buffer	Running buffer 1 plus 3 % Tween®20
Running buffer 2	10 mM phosphate buffer 5 % glycerol, pH 7.5
Elution buffer 3	150 mM phosphate buffer 5 % glycerol, pH 7.5
Running buffer 3	PBS
Elution buffer 4	PBS 0.5 M methyl α-D-mannopyranoside, pH 7.4

2.5 Stock solutions

Ampicillin stock solution	100 mg/mL ampicillin in distilled water working concentration: 100 µg/mL
Cadmium chloride stock solution	5 mg/mL cadmium chloride in distilled water working concentration: 5-40 µg/mL
Chloramphenicol stock solution	30 mg/mL chloramphenicol in 100 % ethanol working concentration: 30 µg/mL
Chymostatin	10 mM chymostatin in DMSO working concentration: 10-100 µM

E-64	2.5 mg/mL E-64 in distilled water working concentration: 7-175 µM
Kanamycin stock solution	75 mg/mL kanamycin in distilled water working concentration: 75 µg/mL
Leupeptin	10 mg/mL leupeptin in distilled water 50 µg/mL (120 µM)
Paromomycin stock solution	130 mg/mL paromomycin sulfate in distilled water working concentration: 150-1000 µg/mL
Penicillin/Streptomycin stock solution	10 mg/mL penicillin in distilled water 2.5 mg/mL streptomycin in distilled water working concentration: 250 µg/mL
Zinc sulfate stock solution	1 M zinc sulfate in distilled water working concentration: 2-4 mM

2.6 Suppliers of antibiotic reagents

Ampicillin	Carl-Roth GmbH & Co KG, Karlsruhe, Germany
Chloramphenicol	Carl-Roth GmbH & Co KG, Karlsruhe, Germany
Kanamycin	Carl-Roth GmbH & Co KG, Karlsruhe, Germany

Paromomycin sulfate	Sigma-Aldrich Laborchemikalien GmbH, Schnelldorf, Germany
Penicillin G	Sigma-Aldrich Laborchemikalien GmbH, Schnelldorf, Germany
Streptomycin sulfate	Sigma-Aldrich Laborchemikalien GmbH, Schnelldorf, Germany

2.7 Suppliers of chemical reagents

Acrylamide 4K-Lösung	Carl-Roth GmbH & Co KG, Karlsruhe, Germany
β-NADP	Sigma-Aldrich Laborchemikalien GmbH, Schnelldorf, Germany
Cadmium chloride	Sigma-Aldrich Laborchemikalien GmbH, Schnelldorf, Germany
Chymostatin	PeptaNova GmbH, Sandhausen, Germany
Coomassie Brilliant Blue	Biomol GmbH, Frankfurt, Germany
Dimethylsulfoxide (DMSO)	Carl-Roth GmbH & Co KG, Karlsruhe, Germany
dNTP Mix (10 mM each)	Fermentas GmbH, St. Leon-Rot, Germany
E-64	PeptaNova GmbH, Sandhausen, Germany
Ferrous sulphate/chelate solution (100 x)	Sigma-Aldrich Laborchemikalien GmbH, Schnelldorf, Germany
Isocitric acid	ICN Biomedicals GmbH, Frankfurt, Germany
Leupeptin	PeptaNova GmbH, Sandhausen, Germany
Luminol	Carl-Roth GmbH & Co KG, Karlsruhe, Germany

Manganese sulfate	Merck KGaA, Darmstadt, Germany
Peptone from casein	Carl-Roth GmbH & Co KG, Karlsruhe, Germany
Proteose peptone	BD, Franklin Lakes, NJ, USA
Skimmed milk powder	Nordmilch eG, Bremen, Germany
Silfar®	Wacker Chemie AG, Nunchritz, Germany
SuperSignal West Pico chemiluminescent substrate	Thermo, Waltham, MA, USA
Triton X-100	Sigma-Aldrich Laborchemikalien GmbH, Schnelldorf, Germany
Tween®20	Sigma-Aldrich Laborchemikalien GmbH, Schnelldorf, Germany
Wheat peptone	OrganoTechnie, La Courneuve, France
Yeast extract	BD Biosciences, Heidelberg, Germany
Zinc chloride	Carl-Roth GmbH & Co KG, Karlsruhe, Germany
2-mercaptoethanol	Carl-Roth GmbH & Co KG, Karlsruhe, Germany

Other chemicals that are not listed here were purchased from Carl-Roth GmbH & Co KG, Karlsruhe, Germany or from Sigma-Aldrich Laborchemikalien GmbH, Schnelldorf, Germany.

2.8 Suppliers of enzymes

Cre-Recombinase	New England Biolabs GmbH, Frankfurt, Germany
Herkulase® Enhanced DNA Polymerase	Stratagene, La Jolla, Canada, USA

ICDH	Sigma-Aldrich Laborchemikalien GmbH, Schnelldorf, Germany
PNGase F	New England Biolabs GmbH, Frankfurt, Germany
Restriction endonucleases	New England Biolabs GmbH, Frankfurt, Germany
T4 DNA Ligase	New England Biolabs GmbH, Frankfurt, Germany

2.9 Suppliers of antibodies and antigens

Influenza anti A/New Caledonia/ 20/99 HA serum (sheep)	NIBSC, Hertfordshire, United Kingdom
Influenza anti A/New Caledonia/ 20/99 HA serum (guinea pig)	Sanofi Pasteur, France
Influenza Antigen A/NewCaledonia/20/99	NIBSC, Hertfordshire, United Kingdom
Influenza anti A/Brisbane/ 10/2007-like HA serum	NIBSC, Hertfordshire, United Kingdom
Influenza Antigen A/ Brisbane/ 10/2007-like (prepared from A/ Uruguay/716/2007)	NIBSC, Hertfordshire, United Kingdom
Influenza anti B/Florida/4/2006	NIBSC, Hertfordshire, United Kingdom
Influenza Antigen B/Florida/4/2006	NIBSC, Hertfordshire, United Kingdom
Anti-sheep IgG-Peroxidase antibody produced in donkey	Sigma-Aldrich Laborchemikalien GmbH, Schnelldorf, Germany
Anti guinea pig IgG-Peroxidase antibody produced in goat	Sigma-Aldrich Laborchemikalien GmbH, Schnelldorf, Germany

2.10 Suppliers of standards

2.10.1 Protein standard

PageRuler™ Plus Prestained Protein Ladder	Fermentas GmbH, St. Leon-Rot, Germany

2.10.2 DNA standards

GeneRuler™ 100 bp Ladder plus	Fermentas GmbH, St. Leon-Rot, Germany
GeneRuler™ 1 kb DNA Ladder	Fermentas GmbH, St. Leon-Rot, Germany

2.11 Suppliers of kits

DNeasy blood and tissue Kit	Qiagen GmbH, Hilden, Germany
FastPlasmid® Mini Kit	5 Prime GmbH, Hamburg, Germany
HiSpeed™ Midi Kit	Qiagen GmbH, Hilden, Germany
Imperial™ Protein Stain Kit	Thermo, Waltham, MA, USA
Montage™ PCR Centrifugal Filter Devices	Millipore GmbH, Schwalbach, Germany
QIAquick Gel Extraction Kit	Qiagen GmbH, Hilden, Germany
RNeasy Plus Mini Kit	Qiagen GmbH, Hilden, Germany
RevertAid™ First Strand cDNA Synthesis Kit	Fermentas, St. Leon-Rot, Germany

2.12 Suppliers of consumables

Gold particles (0.6 μm)	Seashell Technology, La Jolla, CA, USA
Macrocarrier	Bio-Rad Laboratories GmbH, München, Germany
Nitrocellulose membrane	Carl-Roth GmbH & Co KG, Karlsruhe, Germany

Rupture discs (900 psi)	Bio-Rad Laboratories GmbH, München, Germany
Stopping screen	Bio-Rad Laboratories GmbH, München, Germany
Tris-Glycin-Gele (8 & 12 %)	Anamed Elektrophorese GmbH, Groß-Bieberau, Germany
Whatman paper	Carl-Roth GmbH & Co KG, Karlsruhe, Germany
X-ray film, Fuji Super RX	BW plus Röntgen, Kamp-Lintfort, Germany
24 well plates	TPP AG, Trasadingen, Switzerland
96 well plates	TPP AG, Trasadingen, Switzerland

Other consumables were purchased from Omnilab Laborzentrum GmbH & Co KG, Bremen, Germany, or Carl-Roth GmbH & Co KG, Karlsruhe, Germany.

2.13 Laboratory Equipment

Binokular	SZX-ILLD 200	(9)
Biolistic gun	PDS1000/Helium	(3)
Blot Apparatus	Semi Dry Blotter	(2)
Chemiluminescence Detection system	Diana III	(9)
Concentrator	Concentrator 5301	(6)
Electrophoresis Apparatus (DNA)	Agagel Mini	(2)
Electrophoresis Apparatus (Protein)	Mini Twin	(2)
Electroporator (for *E. coli*)	Electroporator 2510	(6)
Fermentor (0.5 L)	Sixfors®	(5)

Fermentor (50 L)	Biostat UD 50	(7)
Gel documentation	AIDA	(9)
Incubation shaker	Innova 4230	(11)
Incubation shaker	Multitron AJ	(5)
Microscope	CK-40	(10)
PCR cycler	mastercycler® gradient	(6)
Photometer	PowerWaveX Select plate reader	(4)
Purification device	Kronlab	(8)
Sonification bath	Sonorex	(1)
Spectral photometer	Libra S12	(12)
Thermomixer	Comfort	(6)
UV-Transilluminator	IDA	(9)

Manufacturer:

(1) Bandelin electronic GmbH & Co KG, Berlin, Germany; (2) Biometra GmbH, Göttingen, Germany; (3) Bio-Rad Laboratories GmbH, München, Germany; (4) Bio-Tek Instruments Inc., Winooski, Vermont, USA; (5) Infors AG, Bottmingen, Switzerland; (6) Eppendorf AG, Hamburg, Germany; (7) Sartorius, Goettingen, Germany; (8) YMC Europe GmbH (Kronlab), Dinslaken, Germany; (9) Raytest GmbH, Straubenhardt, Germany; (10) Olympus GmbH, Hamburg, Germany; (11) New Brunswick Scientific, Edison, NJ, USA; (12) Biochrom AG, Berlin, Germany.

2.13.1 Centrifuges

Centrifuge, 5810R with Swing-bucket-Rotor A-4-62	Eppendorf AG, Hamburg
Centrifuge, 5415D with Rotor F45-24-11	Eppendorf AG, Hamburg
Centrifuge, 5417R with Rotor F45-30-11	Eppendorf AG, Hamburg
Sorvall evolution RC	Thermo, Waltham, MA, USA

2.14 Databanks and Computer programs

2.14.1 Databanks

Tetrahymena Genome Databank (TGD): http://www.ciliate.org

The Institute for Genomic Research: http://www.tigr.org

PubMed of National Centers of Biotechnology (NCBI):

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi

2.14.2 Computer software for Sequence analysis

AIDA Image Analyzer Software

Clone manager suite, Version 6.00

2.14.3 Computer software for the calculation of theoretical molecular weights

Compute pI/MW tool (www.expasy.ch)

ProtParam (www.expasy.ch)

Translate tool (www.expasy.ch)

2.14.4 Computer software for data acquisition

IRIS V 5.0 (32 bit software) (Sixfors®)

Prep Con 5 (Purification device)

2.14.5 Computer software for Literature index

Reference Manager Professional Edition Version 11

2.14.6 Computer software for conversion of figures

Adobe Photoshop, Version 6

CorelDraw, Version 10.427; 2000 Corel Corporation

2.14.7 Computer software for Calculation and Texts

MS Office 2007

2.15 References for DNA Sequences

Beta-tubulin-terminator (BTU2)[161]	NCBI Accession#: L01416
Histon-promoter (H4-1)[162]	NCBI Accession#: X00417
Metallothionein promoter (MTT1)[134]	NCBI Accession#: AY061892
Metallothionein promoter (MTT5)[163]	NCBI Accession#: AY884209

Codon optimized sequences of the antigen genes and the Grl3-locus are listed in the appendix.

2.16 Source of oligonucleotides

Oligonucleotides were purchased from Invitrogen GmbH, Karlsruhe, Germany or Eurofins MWG Operon, Ebersberg, Germany. Oligos were diluted to a final concentration of 1 µg/µl with distilled water. The forward primer and the reverse primer are identified by labeling F or R, respectively. Where present, restriction sites are marked in lower case letters.

2.16.1 Primers for cloning

<u>Primer for the amplification of the truncated New Caledonia derived HA (NC531 and NC344)</u>

NC Sac AA F	5`-TTTgagctcAAGATACTATTTGTATTGGTTATCACG-3` 36 bp
NC531 R	5`-TAAagatctTCAAATTTAGTAAACTC-3` 26 bp

NC344 F	5`-ATCAAAGCTAAGTTATTAGTTTTA-3` 24 bp
NC344 R	5`-GTCTagatctTCATCTAGATTAG-3` 23 bp

The primers above were used for truncation of the HA of the New Caledonia influenza virus strain, which was codon optimized and synthesized by GeneArt AG, Regensburg, Germany. The forward primer contains a SacI restriction site, the reverse primer a BglII restriction site. Furthermore, a stop codon was included for the reverse primers.

Primers for the assembly of a Grl3 knock out vector

Grl3 5´ F	5`-AAGGTGccgcggTGGGAGAGAAAATATG-3` 28 bp
Grl3 5´ R	5`-AAAggatccAAAACGATAGCTAAATTC-3` 27 bp
Grl3 3´ F	5`-CACCAACGAAgtcgacTAATCCAGTC-3` 26 bp
Grl3 3´ R	5`-CAAggtaccCAAAAAAAATTAAAATAC-3` 27 bp

2.16.2 Primers for Sequencing

Primer for sequencing of the different HAs, to verify the success of the cre-reactions and to analyze genomic DNA

NC Seq 1 F	5`-CTGGTGTTTCTGCTTCTTGTTCTC-3` 24 bp
NC Seq 2 R	5`-TAATTACCAATATTAGGAGGGTGG-3` 24 bp

NC Seq 3 F 5`-GAGGGAGGATGGACAGGTATGG-3`
22 bp

UR Seq 1 F 5`-CTGGTAATGATAATTCTACTGCTA-3`
24 bp

UR Seq 2 R 5`-ATCTAGGGCAAGCACCGTAAGT-3`
22 bp

FL Seq 1 F 5`-GACTGCTACTTAAGGAGAAGTTA-3´
23 bp

FL Seq 2 R 5`-ACCAGTGTAGTAAGGCTTAGAC-3`
22 bp

MTT Seq 1 F
5`ATATCTCGAGGATAAGTAATATATTTAGTGCACAATGTTTGAATG-3`
45 bp

MTT Seq 2 F 5`-CACGATTTATGCAATGATCC-3`
20 bp

MTT Seq 3 R 5`-AAATGATAGCCTAAATAATGAG-3`
22 bp

2.16.3 Primers for cDNA analysis

Grl3 1 F 5`-GGTGCTGGATAAAGTGCTGGTAG-3`
23 bp

Grl3 2 R 5`-GGCAAGGTGCTTTTGGAGGAGAG-3`
23 bp

betaHex 1 F 5`-ATGCAAAAGATACTTTTAATTACTTTC-3`
26 bp

betaHex 2 R	5`-TATATTTTAGGAATGTTGTAATC-3`
	23 bp

2.17 Vectors

The used cloning plasmid and the episomal expression plasmid were kindly provided by Cilian AG. All constructed vectors are listed below.

Cloning plasmid pDL-325

The vector was derived from the pCR4TOPO vector where the ampicillin resistance cassette has been removed. The vector contains two loxP sites to enable a cre-dependant homologous recombination. The expression cassette consists of the MTT1 promoter (metallothionein gene of *T. thermophila*), a blasticidin resistance cassette (bsdR from the pCBM-BSD-vector) and a beta tubulin 2 terminator (btu2). Furthermore, a sacB gene was inserted (levansucrase of *Bacillus subtilis*), a chloramphenicol resistance cassette (CmR of the vector pDNR-LIB, Clontech) and a kanamycin resistance cassette (KanR of the vector pCR4TOPO) for propagation in *E. coli*[142].

Expression plasmid pAX

The expression plasmid pAX was published by Weide *et al*.[142] and used for the expression of all HAs.

Mucocyst knock out expression plasmid pGX

An integrative mucocyst knock out vector was developed during this work that combined the knock out of a Grl3 gene with the integration of an expression cassette for the recombinant production of NC531. This vector is based on the published pKOIX vector[142]. The 5´ and 3´ DHFR-TS integration flanks were replaced by parts of the 5´ and 3´ coding and non-coding regions of the Grl3 gene of *T. thermophila*. The sequence is demonstrated in appendix.

pDL_MTT1_NC

The hemagglutinin (HA) of the influenza A virus strain New Caledonia/20/99 (H1N1; Accession number: CY033622.1) was codon optimized and synthesized by an external provider (GeneArt AG, Regensburg, Germany). The synthetic gene of this hemagglutinin was named NC HA and transferred by the external provider to the working plasmid pDL-325 by replacing the bsdR open reading frame.

pDL_MTT1_NC531

The antigen of the New Caledonia strain was truncated by removal of the coding sequence of the transmembrane domain (TMD) of the NC HA .The truncated version was generated by PCR and cloned into the working plasmid pDL-325 by replacing the bsdR open reading frame, and named NC531 HA.

pDL_MTT5_NC531

The pDL_MTT1_NC531 was modified by exchanging the MTT1 promoter to the MTT5 promoter region of *T. thermophila*.

pDL_MTT1_UR

The HA of the influenza A virus strain Uruguay/716/2007 (H3N2; Accession number: CY035022) was codon optimized and synthesized by an external provider (GeneArt AG, Regensburg, Germany) and transferred to the working plasmid pDL-325 by replacing the bsdR open reading frame, and was named UR HA.

pDL_MTT1_FL

The HA of the influenza B virus strain Florida/4/2006 (Accession number: EU515937) was codon optimized and synthesized by an external provider (GeneArt AG, Regensburg, Germany) and transferred to the working plasmid pDL-325. This was performed by replacing the bsdR open reading frame, and named FL HA.

pAX_MTT1_NC

The expression cassette including the MTT1 promoter, the NC antigen, the BTU2 terminator and the chloramphenicol resistance (CmR) were transferred via cre-dependant homologous recombination into the episomal expression plasmid pAX[142].

pAX_MTT1_NC531

The expression cassette with the MTT1 promoter, the truncated NC531 antigen, the BTU2 terminator and the CmR were transferred by a cre-dependant homologous recombination into the pAX vector[142].

pGX_MTT1_NC531

The expression cassette including the MTT1 promoter, the NC531 coding region, the BTU2 terminator and the CmR were transferred by the homologous cre-dependant recombination to the integrative mucocyst knock out vector pGX.

pAX_MTT5_NC531

The expression cassette comprising the coding region of the NC531 antigen including the MTT5 promoter, the BTU2 terminator and the CmR were transferred from the working plasmid to the expression plasmid pAX by cre-dependant homologous recombination[142].

pAX_MTT1_UR

The UR expression cassette from the pDL working plasmid including the MTT1 promoter, BTU2 terminator and the CmR were transferred via cre-dependant homologous recombination into the episomal expression plasmid pAX[142].

pAX_MTT1_FL

The FL expression cassette including the MTT1 promoter, the BTU2 terminator and the CmR were transferred to the expression plasmid pAX by cre-dependant homologous recombination[142].

3. Methods

3.1 Cultivation and storage of bacteria and ciliates

3.1.1 Strain maintenance of *Escherichia coli*

Cultivation and selection of *E. coli* was performed using LB-agar plates supplemented with an appropriate antibiotic. The plates were incubated at 37 °C and afterwards stored at 4 °C. For long term storage, 500 µl of the *E. coli* culture was mixed with 500 µl (v/v) Glycerol and stored at -80 °C.

3.1.2 Strain maintenance of *Tetrahymena thermophila*

A 1.5 mL culture of *T. thermophila* was regenerated once a week. The cells were cultivated in 24 well plates in 1.5 mL SPP medium under selection pressure (400 µg/mL paromomycin) and stored at 30 °C in an incubator.

3.1.3 Long-term storage in a culture

For long-term storage of transformed *T. thermophila* cultures, autoclaved water with a chick pea in a glass tube was used. The inoculation was performed by addition of 10 µl of *Tetrahymena* culture to the glass tube. The cultures were renewed biannually. Alternatively, long term storage was achieved by freezing transformed *Tetrahymena* cells as described previously[130]. In modification, SPP medium was used for initial cultivation of the target cells.

3.1.4 Cultivation of *E. coli*

E. coli submers culture were inoculated with a single colony from an agar plate. Cultivations were performed using 15 mL tubes or 100 mL flasks. Cultures were incubated at 37 °C and 180 rpm in a shaking incubator. Selection was applied using an appropriate antibiotic. Cultures were harvested by centrifugation at 4,000 x g and 4 °C for 30 minutes.

3.1.5 Cultivation of *T. thermophila*

Recombinant *T. thermophila* cells were cultured using a 5 mL starter culture of SPP medium supplemented with 400 µg/mL paromomycin. Cultures were incubated at 80 rpm and 30 °C for three days. Larger cultures for subsequent experiments were inoculated with 1 mL of starter culture in 10 mL SPP supplemented with 400 µg/mL paromomycin. The expression of recombinant protein was induced using cadmium chloride or zinc. Cell pellets were harvested by centrifugation at 1,500 x g at 10 °C for 3 minutes. Supernatants were also collected for analysis.

3.1.6 Determination of the cell titer of *T. thermophila*

The cell titer was determined using a *Fuchs-Rosenthal* counting chamber. The suspension of cells was diluted using 2.5 % (v/v) glutaraldehyde. Four big squares were counted and the titer was calculated using a mathematic formula:

cell titer (cells/mL) = (M/64) x 8 x D x 10.000

M = mean of the counted squares

D = factor of dilution

3.2 Transformation of bacteria and ciliates

3.2.1 Transformation of *E. coli* electroporation

Transformation of electro competent *E. coli* cells was performed using an Eppendorf 2510 electroporator. One aliquot (40 µl) of electro competent cells (2.1.1.) was taken for each transformation, mixed with 20 µl of purified DNA, and transferred into an ice-cold cuvette (gap 1 mm, Eppendorf AG, Hamburg, Germany). The electroporation was done at 1,800 V. The regeneration of the cells was performed in 500 µl SOC medium for 1 h at 37 °C and 450 rpm. Cells were spread after the transformation on a prewarmed selective plate and grown at 37 °C overnight.

3.2.2 Transformation of *T. thermophila*

Preparation of DNA

Plasmid DNA for transformation of *T. thermophila* was isolated via MIDI preparation (3.2.4) and concentrated to ~1 µg/µl.

Conjugation and preparation of the cells

Prior to conjugation of *T. thermophila* cells, cultures were starved. *T. thermophila* cultures of two different mating types growing in exponential phase were pelleted by centrifugation at 21 °C and 900 x g for 3 minutes, washed in Dryls buffer (pH 7.4) and an additional centrifugational step was performed (21 °C, 900 x g, 3 min). A cell titer of 200,000 cells/mL in 100 mL Dryls buffer was adjusted. The resulting culture of two mating types was incubated for approximately 20 hours at 30 °C, shaking at 180 rpm to prevent conjugation. Subsequently, cells begin to conjugate.

After 4-5 hours, conjugation rate should be high, approximately 70-80 %. Following, after 10 hours of incubation, cells are competent, meaning the new MAC is going to be built, and the MAC transformation can be performed. 20-40 mL of the conjugated cells were pelleted (21 °C, 900 x g, 2 min) and resuspended in 1 mL of Dryls buffer. This suspension was taken for the transformation.

Preparation of gold particles

The preparation of the gold particles (diameter 0.6 µm) was done according to a protocol of Seashell Technology. DNadeI™ gold particles (S550d) were stored in binding buffer at a concentration of 50 mg/mL. Firstly, the gold was diluted by the addition of binding buffer to a concentration of 30 mg/mL and plasmid DNA was added to give a ratio of 2-5 µg per mg gold. After mixing, the same volume of precipitation buffer was added and the reaction was mixed again. Samples were mixed for 3 minutes at room temperature then centrifuged for 10 sec at 9.3 x g. The supernatant was discarded and 500 µl ice-cold ethanol (100 %) was added to the pellet. After mixing, a centrifugational step was performed (room temperature, 10 sec, 9.3 x g), the supernatant was discarded and the pellet was resuspended in 20 µl of

100 % ethanol. The gold was sonicated for 1-2 min and placed on a Macrocarrier. This Macrocarrier was dried on a filter paper on $CaCl_2$.

Transformation by biolistic bombardment

Transformation of *T. thermophila* cells was done using a PDS-1000/He Biolistic® Delivery system where an applied helium pressure accelerates the gold particles. In preparation for the transformation, a rupture disk (900 psi) was put into the rupture disk retaining cap and the fully assembled microcarrier launch assembly (with cover lid) was inserted inside the bombardment chamber. The microcarrier launch assembly holds the DNA/microcarrier preparation on a Macrocarrier sheet over the Stopping Screen in the path of the helium shock wave. The target shelf holds the *T. thermophila* cells in a Petri plate in the path of the accelerated DNA/microcarrier preparation. The door of the chamber was closed and the chamber was evacuated, the vacuum was held at the desired level of ~27 inches of mercury (Hg vacuum) and the fire button was continuously depressed until the rupture disk burst and the helium pressure gauge dropped to zero. The vacuum was released from the chamber and the *T. thermophila* cells were removed from the chamber. The transformed cells were regenerated in proteose peptone medium (PP 1 %) overnight at 30 °C. The detection of transformed cells was done afterwards under selection pressure (as described in 3.2.3).

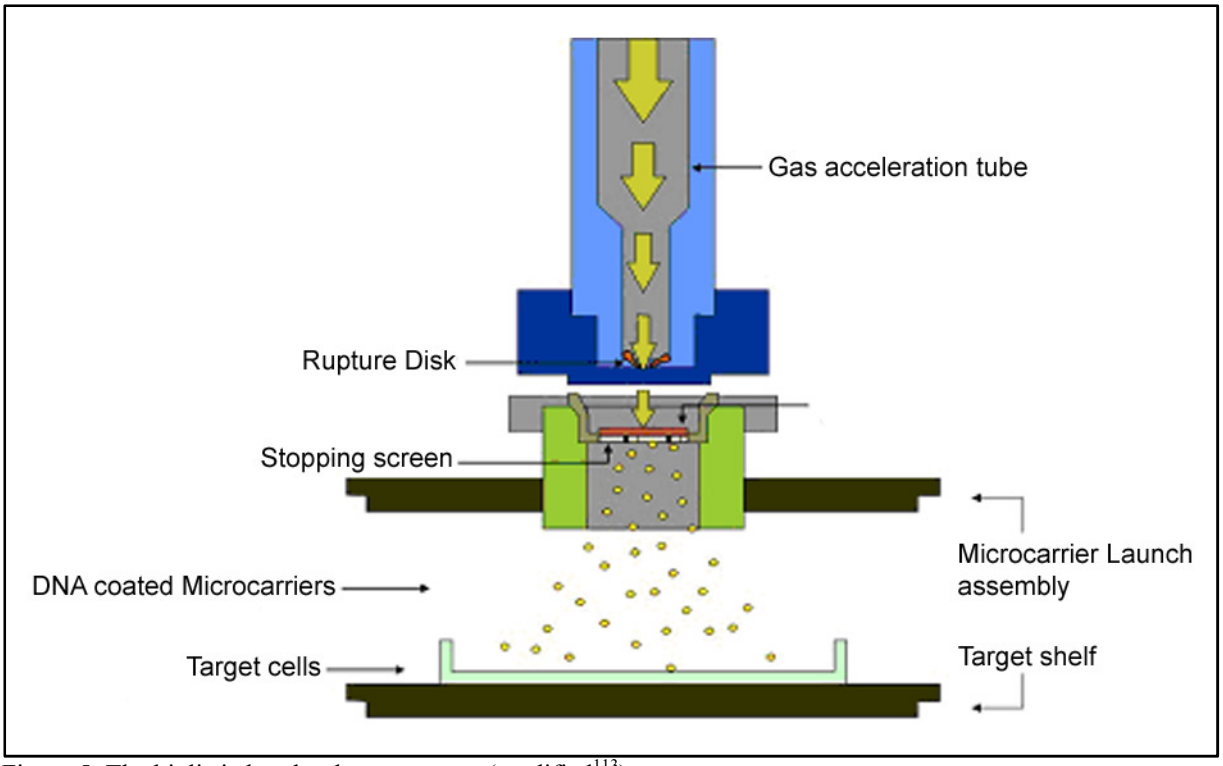

Figure 5: The biolistic bombardment process (modified[113]).
At helium pressure of 900 psi the Rupture Disk bursts and the Macrocarrier with the DNA loaded gold particles is accelerated. Because of the Stopping Screen, the Macrocarrier is hold. The gold particles are transported to the target cells and penetration and transformation takes place.

3.2.3 Selection of transformed *T. thermophila*

For selection of positive transformants, tranformed cells were harvested by a centrifugational step (3 min, 2100 rpm, 21 °C), resuspended in 50 mL of skimmed milk medium supplemented with 150 µg/mL paromomycin. Cells were spread in 96-well plates and after an incubation period of 2-3 days investigated by a light microscope. Growing cultures indicated resistance to the added antibiotic and therefore correct transformation. Putative transformants were cultivated in skimmed milk medium under selection pressure (300 µg/mL paromomycin). After approximately 7 days, the selection pressure was raised to 400-1000 µg/mL in skimmed milk medium. The strain maintenance was performed using SPP medium and the supplementation of 400 µg/mL of paromomycin.

3.3 Cloning of DNA

3.3.1 Purification of DNA after agarose gel electrophoresis

QIAquick Gel extraction kit was used for the extraction of DNA. Gel slices of DNA fragments obtained from an agarose gel electrophoresis step (3.5) were put into a sterile

1.5 mL reaction tube. The extraction was performed according to the manufacturer's manual. The DNA was eluted by addition of 30 µl of sterile and distilled water and stored at -20 °C.

3.3.2 Purification of DNA using Montage™ PCR Centrifugal Devices

Products of a PCR, a ligation or a restriction were purified from contaminating salts or oligonucleotides using Montage™ PCR Centrifugal Devices. The procedure was done in accordance with the manufacturer's recommended protocol. The obtained DNA solution (20 µl) was stored at -20 °C.

3.3.3 Isolation of total DNA from *T. thermophila* cells

For the isolation of total DNA from a transformed or a wild type *T. thermophila* culture, the DNeasy blood and tissue Kit was used. The preparation was performed with slight modifications from the manufacturer's protocol (Purification of Total DNA from Animal Blood or Cells). The harvest was done using a *T. thermophila* culture (~4 x 10^6 cells) and a centrifugation step for 3 minutes and 1500 x g at 10 °C. The lysis of the cells with Buffer AL was prolonged to 30 minutes and the elution was performed using 200 µl of sterile and distilled water instead of Buffer AE. The samples were immediately used for the analysis of DNA via PCR or stored at -20 °C.

3.3.4 Isolation of high purity plasmid DNA with the FastPlasmid™ Mini Kit

The isolation of plasmids was done using the FastPlasmid™ Mini Kit. A single isolated colony from a freshly spread *E. coli* strain was inoculated into 5 mL of selective culture medium. The growth of bacterial cultures took place overnight at 37 °C and shaking at 180 rpm. 1.5 mL of the cultures were harvested and the cell pellet was treated as described in the manufacturer's manual. The elution was done using 50 µl of sterile distilled water and the isolated DNA was stored at 4 °C. Restriction analyses were performed using this DNA.

3.3.5 Isolation of high purity plasmid DNA with the HiSpeed® Plasmid Midi Kit

A Midi preparation was done using the HiSpeed® Plasmid Midi kit. The protocol is based on alkaline lysis followed by anion exchange chromatography. For the preparation, 50 mL of an

E. coli culture, grown overnight, was harvested. The preparation was performed according to the manufacturer's manual. The obtained DNA was concentrated to 1 µg/µl using a concentrator and stored at 4 °C until the transformation of *T. thermophila* cells.

3.3.6 Restriction analysis

The restriction of DNA was done using endonucleases with the appropriate buffer. DNA fragments with blunt ends or cohesive ends were formed, depending on the enzyme used. The samples were incubated for 1 to 3 hours at the manufactuer's recommended temperature and separated by agarose gel electrophoresis (3.5).

3.3.7 Ligation

The ligation of a plasmid with insert was performed by applying the T4 DNA Ligase in the appropriate buffer. To aid the efficiency of the ligation, the insert DNA should be present in a high molar ratio. DNA concentration was estimated by agarose gel electrophoresis of the DNA (3.5). DNA fragments for ligation were purified using the QIAquick Gel extraction kit (3.3.1). The ligation was performed according to the manufacturer's recommended protocol and incubated overnight at 16 °C.

3.3.8 Cre dependant homologous recombination

The expression cassette of the pDL cloning plasmids was transferred to the expression plasmid pAX as described previously[142].

3.3.9 PCR method

PCR is an *in vitro* technique capable of amplifying trace amounts of DNA of defined length and sequence from heterogenous nucleic acid mixtures. During the reaction, the sequence between specific primers is repeatedly replicated by a thermostable DNA polymerase upon every thermal cycling step. PCR reactions were performed using a PCR mastercycler®. All programs were set up and cycled according to the details below.

PCR program:		Component	
1) 95 °C	3 min	template-DNA (0.1 µg/µl)	1 µl
2) 95 °C	1 min	Reaction buffer (10 x)	5 µl
3) 55 °C	1 min	dNTP mix (10 mM each)	1 µl
4) 65 °C	1-2 min, 39 x	primer 1 (1 µg/µl)	1 µl
5) 65 °C	5-8 min	primer 2 (1 µg/µl)	1 µl
6) 4°C	∞	Polymerase	2.5 U
		Distilled water	ad 50 µl

3.4 Sequencing of DNA

Sequencing analysis was performed by MWG (Ebersberg, Germany). Clone manager suite (2.14) was used for the evaluation of sequence data.

3.5 DNA agarose gel electrophoresis

Agarose gel electrophoresis is a method to separate DNA achieved by moving negatively charged nucleic acid molecules through an agarose matrix within an electric field. Shorter molecules move faster and migrate further than longer ones. Conformation is also a factor. To avoid this problem linear molecules are usually separated, usually DNA fragments obtained from a restriction digest, or linear PCR products. A DNA plasmid that has not been cut with a restriction enzyme will move with different speed corresponding to the conformational forms of the DNA. Nicked or open circular, linearised, or supercoiled plasmid forms will separate at different rates, listed from slowest to fastest. Increasing the agarose concentration of a gel reduces the migration speed, and enables the separation of smaller DNA molecules. For the determination of the size of DNA fragments a standard was used (2.10).

The electrophoresis was performed in a horizontal apparatus in TAE buffer. Depending on the expected size of the DNA fragments, an agarose concentration of 1-1.5 % (w/v) in TAE buffer was used. Samples were supplemented with loading dye (6 x, Fermentas, St. Leon-Rot, Germany) and applied to the gel. A voltage of 3-10 V/cm was applied. For the visualization of DNA bands ethidium bromide was used, that intercalates into DNA molecules and can be visualized with UV light. A picture of the obtained gels was taken.

3.6 Electrophoresis und Visualization of proteins

3.6.1 Preparation of samples

The cell titer of an exponential or stationary phase SPP-culture of a transformed or a non-transformed wild type strain were calculated using a *Fuchs-Rosenthal* counting chamber (3.1.5). 1-10 mL of the culture were harvested by centrifugation at 10 °C, 1,500 x g for 3 minutes. Cells and supernatants were collected and frozen at -20 °C or immediately used for SDS-PAGE. The cells were solubilized by applying RIPA buffer and a following sonification step for 15 minutes. Prior to loading a SDS-gel, the cell extracts and supernatants were mixed with a sample buffer containing SDS and heated to 95 °C for 3 minutes.

3.6.2 Denaturing SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE relies on a discontinuous buffer system. Two ions of differing electrophoretic mobility (glycinate and chloride) form a moving boundary when voltage is applied. Proteins have an intermediate mobility, causing them to concentrate, or stack, into a narrow zone at the beginning of electrophoresis. As the boundary moves through the gel, the sieving effect of the polyacrylamide gel matrix causes different proteins to move at different rates. The stacking effect is responsible for the high resolving power of SDS-PAGE. The sample is loaded in a relatively broad zone, and the moving boundary concentrates the proteins into sharp bands prior to separation.

Protein samples for SDS-PAGE were prepared using SDS and for some samples a thiol reductant, 2-mercaptoethanol was added. SDS forms complexes with proteins giving them a rod like shape and similar charge to mass ratio. The reductant cleaves disulfide bonds between and within proteins allowing complete denaturation and dissociation. Heat treatment in the presence of SDS and reductant effectively eliminates the effects of protein structure and native charge on electrophoretic mobility, so the migration distance depends primarily on molecular weight.

Gels were comprised of polyacrylamide with a bisacrylamide cross linker. Each gel has a 5 % polyacrylamide stacking layer extending approximately 5 mm from the bottom of the loading well to the top of the resolving gel. The resolving gel was of 8 or 12 % polyacrylamide with 10 lanes. Gels were bought from Anamed GmbH (Groß-Bieberau, Germany) or self made. Resolving gels contained a final concentration of 0.375 M Tris-HCl (pH 8.8) and 0.1 % SDS.

The 5 % stacking gel contained a final concentration of 0.125 M Tris-HCl (pH 6.8) and 0.1 % SDS.

Following protocol was used for the preparation of four gels:

8 % resolving gel (25 mL)		5 % stacking gel (12 mL)	
9.7 mL	H_2O	6.74 mL	H_2O
6.3 mL	4x resolving gel buffer	3 mL	4x stacking gel buffer
6.7 mL	Acrylamide solution	2 mL	Acrylamide solution
250 µl	10 % SDS	120 µl	10 % SDS
250 µl	10 % APS	120 µl	10 % APS
10 µl	TEMED	20 µl	TEMED

12 % resolving gel (25 mL)	
8.2 mL	H_2O
6.3 mL	4x resolving gel buffer
10 mL	Acrylamide solution
250 µl	10 % SDS
250 µl	10 % APS
10 µl	TEMED

The agents APS and TEMED were added immediately before the gels were cast, because they initiate the polymerization of the gel matrix.

After the casting of the resolving gel, the gel was protected by the addition of 5 mL of isopropanol to hinder contact with oxygen. After 60 minutes of polymerization, the isopropanol was discarded and a polymerized gel could be seen. The stacking gel was prepared and cast on top of the resolving gel. A comb was used to form pockets for loading the samples. After 15 minutes of polymerization the gels were ready for SDS-PAGE. The separation of the samples was performed at approximately 120 V for 2-3 hours. A protein standard was used for the SDS-PAGE to estimate molecular weights of the separated proteins.

3.6.3 Western blot analysis

Protein blotting, the transfer of proteins to solid-phase membrane supports, has become a powerful and popular technique for the visualization and identification of proteins. Protein blotting involves two major phases: transfer and detection. The transfer step involves moving the proteins from the gel and immobilizing them on a synthetic membrane support (blot). The transfer was done by semi-dry blot to a nitrocellulose membrane. Proteins were detected by

the application of a specific antibody followed by the use of a secondary antibody coupled by the enzyme of the horse radish peroxidase. Antibody localization was visualized by the addition of a chemical substrate and the optical signal was detected using X-ray film or a visualization chamber.

In the electrophoretic transfer, the membrane and protein-containing gel were placed together with filter paper between two electrodes. Proteins migrate to the membrane following a current (I) that is generated by applying a voltage (V) across the electrodes, following Ohm's law (V = I x R) where R is the resistance generated by the materials placed between the electrodes (that is, the transfer buffer, gel, membrane, and filter papers). The electric field strength (V/cm) that is generated between the electrodes is the driving force for electrophoretic transfer. In a semi-dry transfer, the gel and membrane are sandwiched between two stacks of filter paper and are in direct contact with plate electrodes[164,165]. The term "semi-dry" refers to the limited amount of buffer, which is confined to the two stacks of filter paper. The blot was performed by 1 mA/cm^2 gel area for 1-2 hours. Following the membrane was washed by PBS-T and incubated in 10 % skimmed milk in PBS-T overnight at 4 °C to block unspecific protein binding sites. The primary antibody was added in a working dilution in 5 % skimmed milk or BSA in PBS-T for 1 hour at room temperature and shaking. Afterwards, three washing steps followed by PBS-T (each 10 min), and the diluted secondary antibody was added for 1 hour at room temperature and shaking. After four additional washing steps the detection of proteins was done by addition of a chemiluminescent substrate.

3.6.4 Quantitative analysis

To estimate recombinant expression yields a quantitative analysis was performed. To load a SDS-gel a calibration curve by different dilutions of the purchasable antigen and dilutions of the recombinant expressed protein obtained during a 0.5 L or 50 L fermentation process were made. After separation by SDS-PAGE and the transfer of proteins to a nitrocellulose membrane by Western blot protein was detected by addition of a specific primary antibody and the appropriate secondary antibody (as described in 3.6.3.). A luminol substrate was added and the visualization of the quantification was performed with the CCD camera system DIANA III. The analysis was done using the AIDA program and yields were calculated thereby.

3.6.5 Coomassie staining

Separated proteins after a SDS-PAGE were stained by Imperial™ Protein stain Kit as described by the manufacturer's manual.

3.6.6 Isocitrate dehydrogenase assay

Intracellular enzymes can be released into the culture medium due to lysis of cells. The enzyme of the isocitrate dehydrogenase (ICDH) is a mitochondrial enzyme and comprises a high cytosolic activity and stability. Therefore, it represents an ideal marker enzyme to detect leakage of the cells. The enzymatic activity was measured as described by the manufacturers manual (Biochemika, Boehringer Mannheim GmbH, Germany). The increase of the extinction was measured using a photometer at 339 nm. For this test an aliquot of 1 mL of a complete culture and 300 µl of the cell free supernatant were harvested.

3.7 Glycosylation analysis

For the deglycosylation one aliquot of purified antigen of *T. thermophila* was digested by N-glycosidase F (PNGase F). This amidase of the *Flavobacterium meningosepticum* hydrolyzes N-glycans of glycopeptides or glycoproteins between the innermost GlcNAc- and asparagine-residue. The enzymatic digest was done in two steps, the denaturation of the sample and the PNGase F digest.

Denaturation	N-Glycosidase F-digest
10 µl purified NC531	20 µl denatured sample
2 µl Denaturing buffer (10 x)	3 µl G7-Puffer (10 x)
8 µl distilled water	3 µl NP-40
	2 µl N-Glycosidase F
	2 µl distilled water
Incubation: 100°C, 10 min	Incubation: 37°C, 1-24 h

Samples were analyzed via SDS-PAGE with subsequent transfer to a nitrocellulose membrane by Western blot.

3.8 Estimation of protein content (Bradford assay)

The estimation of whole protein content was done by the colorimetric method of Bradford. It is based on the binding of the dye Coomassie Brilliant Blue G-250 in acidic milieu and the

associated shift of the maximum of the absorption from 465 nm to 595 nm. The determination of the protein content was performed according to the manufacturer's manual (Thermo, Waltham, MA, USA). 30 µl of each sample was mixed with 150 µl of Coomassie Protein Assay Reagent in a 96-well plate and after a 5 min of incubation at room temperature the absorption was measured at 595 nm. Protein concentrations were calculated by interpolation on a calibration curve generated using BSA standards.

3.9 Analysis of mucocyst knock out strains

3.9.1 Dibucaine stimulation

A transformed NC531 expressing Grl3 knock out strain and a non-transformed wild type strain were cultivated in a 5 mL cell culture under selection pressure (400 µg/mL paromomycin). After three days of growth the cells were harvested by centrifugation (21 °C, 1,000 x g, 3 min) and the supernatant was discarded. The cell titer was calculated and cells were adjusted to a concentration of 1 x 10^6 cells/mL. 2 mL of this cell suspension was washed with Dryls buffer and the excretion of mucocyst contents was initiated by the addition of dibucaine in a final concentration of 1.4 mM. After an additional centrifugation step (21 °C, 1,000 x g, 3 min) the mucocyst contents were visualized by taking a picture.

3.9.2 RNA analysis

The RNA of a transformed NC531 expressing Grl3 knock out strain and a non-transformed *T. thermophila* wild type strain was isolated using the RNeasy Plus Mini Kit (Qiagen, Hilden, Germany). The mRNA was isolated due to the manufacturer's manual.

3.9.3 cDNA synthesis

cDNA was made by using the RevertAid™ First Strand cDNA Synthesis Kit (Fermentas, St. Leon-Rot, Germany). The procedure was done according to the manufacturer's manual. The verification of the cDNA was done by PCR analysis and a subsequent agarose gel electrophoresis.

3.10 Fermentation procedures

3.10.1 Lab-scale fermentation (0.5 L)

For the fermentation of *T. thermophila* expression strains a 0.5 L fermentation process was done using a Sixfors® multibioreactor (Infors AG, Bottmingen, Switzerland). The fermentor was equipped by a paddle impellor, and a digital control subunit (DCU).

The stirrer was limited to a rotation not exceeding 800 rpm, because faster stirring lead to high cell damage. The concentration of dissolved oxygen was adjusted to be held at 20 % of the air saturation level. Thereby, the oxygen supply was regulated at first priority and the rotation of the stirrer as second priority. The measurement of oxygen saturation was performed using an O_2-electrode. The temperature was maintained at 30 °C. The pH value was regulated during the whole fermentation procedure at 7.0 by acetic acid and NaOH.

The complete mounted fermentor including SPP medium was autoclaved for 30 min at 120 °C. Inoculation of the fermentor was performed sterile by using a syringe connected through a tube to the fermentor. The first phase of the fermentation was performed as a batch procedure. When cells reached the end-logarithmic phase the second phase was initiated by adding a feeding solution all at once. This fed-batch fermentation procedure was found out to best suited for the production of recombinant antigens by *T. thermophila*. The feeding solution was composed of concentrated SPP medium, so that after the feeding step the final concentration of 1 fold was reached within the fermentation broth.

3.10.2 Up scaling fermentation process (50 L)

The 50 L fermentation procedure was performed using a Biostat® UD (Sartorius, Goettingen, Germany) fermentor. The fermentation procedure was performed according the 0.5 L fermentation process, except that the medium was autoclaved within the fermentor itself and inoculation was done through a membrane.

3.10.3 Harvest of the cells

Cells obtained of the 0.5 L fermentation process were transferred into a sterile and ice cold fernbach flask and incubated on ice for 20 min at 80 rpm. Afterwards, cells were harvested by centrifugation at 10 °C and 1,500 x g for 10 min (Sorvall evolution RC, SLC-6000 rotor). The

supernatant was discarded and the cell pellet was washed using ice cold phosphate buffer and resuspended. Following, the cells were pelleted by an additional centrifugational step at 10 °C and 2,400 x g for 8 min (SLA-1500 rotor). The washing and centrifugational step was repeated 3-4 times to remove mucocyst contents. Mucocyst-free cell pellets were resuspended in 90 mL of phosphate buffer (10 mM) and immediately frozen at -20 °C.

The harvest of the cells from the 50 L fermentation process was achieved using a hollow fiber module (0.3 m^2, 3 L/min, tube diameter 11 mm). Applying this module, the cells were separated from the fermentation broth. The washing steps to remove mucocyst contents were performed as described for the 0.5 L fermentation process. The obtained mucocyst-free cell pellet was frozen immediately at -80 °C.

3.11 Purification via column chromatography

3.11.1 Cell lysis

For the purification of recombinant NC531 HA, 80 g of frozen and washed cells obtained from the 50 L fermentation process were resuspended in 400 mL loading buffer and supplemented by the cysteine protease inhibitor E-64 (70 µM). Cells were disrupted using an ultraturrax (IKA UT T25 + S25N-25G) at 10,000 rpm for 5 min on ice. The lysate was filled up to 1.8 L using loading buffer and the solubilization was performed by stirring at 4 °C for 17 hours. This cell lysate was filtrated using a hollow fiber module (0.45 µm with 850 cm^2 surface) and washed three times with 2.3 L of loading buffer. The duration of this procedure was about 2 h at a flow rate of 1 L/min and a tube with a diameter of 8 mm. The pH dropped to 7.1 and was adjusted to 7.4 prior to loading onto the first column.

3.11.2 First column: Capto™ Q

The first column was a Capto™ Q column (GE healthcare Europe GmbH, Munich, Germany) that is a strong quaternary ammonium anion exchange medium. It is designed for capturing proteins from large feed volumes by packed bed chromatography. A 125 mL Capto™ Q column was loaded with the lysate of the cells at a flow rate of 15 mL/min and afterwards washed with running buffer 1. The elution was performed with elution buffer 1 by a 400 mL step elution and with the addition of 150 mM NaCl. To elute all protein that was not eluted during the 150 mM NaCl step the column was washed with 1 M NaCl in elution buffer 2.

Fractions of each 10 mL were taken and E-64 was added to each fraction in a final concentration of 175 µM. Samples were further analyzed by SDS-PAGE, Western blot and Bradford assay. NC531 HA positive fractions were pooled for the next column chromatography step.

3.11.3 Second column: Ceramic Hydroxyapatite (CHT)

Ceramic Hydroxyapatite (CHT) (Bio-Rad, Hercules, CA, USA) was used to capture contaminating proteins. The NC531 HA did not bind to this column but a high percentage of other proteins did. The pooled fractions obtained from the Capto™ Q column were concentrated and a buffer exchange to remove the NaCl was performed using a Labscale TFF module (30 kDa) with running buffer 2. The sample that was loaded onto the column, with a bed volume of 20 mL, had a volume of 150 mL. The pH was adjusted to 7.5 prior to loading onto the column and the column was loaded with a flow rate of 7 mL/min and subsequently washed with running buffer 2. The elution was performed by a step elution by applying elution buffer 3. The flow through of 200 mL was taken to load to the next column. The eluate and the flow through were tested by SDS-PAGE, Western blot analysis and Bradford assay.

3.11.4 Third column: Con A

The third column was a Con A Sepharose 4B (GE healthcare Europe GmbH, Munich, Germany) with a bed volume of 15 mL. The flow through of the CHT column was supplemented with 150 mM NaCl and applied with a flow rate of 5 mL/min to the column and afterwards washed with running buffer 3. The elution was done about 10 column volumes by applying elution buffer 4. Samples were analyzed by SDS-PAGE, Western blot and Bradford assay.

3.12 Safety regulation

Genetic modifcation experiments were performed pursuant the law "*Gesetz zur Regelung von Fragen der Gentechnik*" (*Gentechnikgesetz*, *GenTG*) of 29.07.2009 observing the formulated §7 Abs. 1 *GenTG*. Announcement of the work performed was pursuant the law § 8 Abs. 1 and § 11 of *Gentechnikgesetz*.

4. Results

As described in the introduction, none of the expression systems currently available is ideal for a safe, low-cost and efficient production of a vaccine. New production platforms have to be found and developed. *Tetrahymena thermophila* represents a promising expression system that can be advanced to an excellent production platform of recombinant therapeutics and subunit vaccines. Recently, it was successfully demonstrated that the expression of i-antigens from *Ichthyophthirius multifiliis* in the *Tetrahymena* system using a strong inducible promoter is feasible[134]. These proteins were targeted to the plasma membrane in their correct 3-dimensional conformation[59]. Besides, it could be demonstrated that the GPI-anchored circumsporozoite protein of the malaria parasite *Plasmodium falciparum* was expressed and targeted to the surface of *T. thermophila*[58]. This work addresses to the expression and secretion of viral antigens. A production and purification process for the influenza vaccine component was established to reach a purity which is acceptable for industrial purposes.

4.1 Expression of three different recombinant hemagglutinins

As optimal viral vaccine component, the main antigen of the influenza virus, the hemagglutinin (HA), was chosen. Due to the fact that the common subunit influenza vaccine is composed of three HAs originating from different influenza strains, three HAs were intended to be produced using the *T. thermophila* expression system[37].

4.1.1 Design and construction of hemagglutinin expression cassettes

The genes of three various HAs of the influenza A virus New Caledonia/20/99 (NCBI Accession number CY033622.1, H1N1), of the influenza A virus Uruguay/716/2007 (NCBI Accession number CY035022, H3N2), and of the influenza B virus Florida/4/2006 (NCBI Accession number EU515937) were identified and codon optimized. Therefore, the native codons were adapted to the codon preferences of *T. thermophila*. The sequences were optimized and synthesized by an external provider (GeneArt AG, Regensburg, Germany; see appendix). The obtained synthetic genes encode for 566 amino acids (aa) of the New Caledonia derived HA precursor protein (Figure 6 A, NC HA), for 533 aa of the Uruguay derived HA (Figure 6 C, UR HA) and for 537 aa of the Florida derived HA (Figure 6 D, FL HA), comprising the respective original ER leader sequence. All synthetic

genes were flanked by an EcoRV and a BglII restriction site to transfer the genes to the cloning plasmid pDL that was kindly provided by the Cilian AG. The cloning was performed by an external provider (GeneArt AG, Regensburg, Germany), and the created plasmids were named pDL_MTT1_NC, pDL_MTT1_UR, and pDL_MTT1_FL. The heterologous expression of the recombinant HAs was controlled by the application of the strong and cadmium inducible metallothionein 1 promoter (MTT1) system and the well established beta tubulin 2 (BTU2) terminator sequence[134]. The generated plasmids contained expression cassettes that were composed of the MTT1 promoter, the target gene, the btu2 terminator and a chloramphenicol resistance gene. The whole cassette was flanked by loxP sites on the 5' and 3' ends that enabled the application of cre-dependant transfer of the expression cassette from the pDL cloning plasmid to the *T. thermophila* expression plasmid. A chloramphenicol resistance (CmR) was inserted between the loxP site and the BTU2 terminator to reduce background clones, because this CmR is only translated if a correct site-specific recombination between expression and cloning plasmid has been occurred.

Regarding one target antigen, the HA of the NC strain, two different versions were planned to be tested. The first version comprised the whole coding region of the NC derived HA that included the HA1 chain, the HA2 chain, the fusion peptide and the transmembrane domain (TMD). The second version was a C-terminally truncated version of the NC HA coding region lacking the TMD and the fusion peptide. The additional analysis of the truncated NC protein based on the fact that the TMD anchors the HA in the native viral membrane, and is therefore able to anchor the HA in any *T. thermophila* based membrane that can result in an inhibition of the efficient secretion of the recombinant antigen. Furthermore, it could also hinder the passage through a column chromatography material. For the shortened NC derived HA, a PCR was performed to obtain a coding region without the TMD (aa 1-531), and to add a stop codon directly downstream the amplified sequence. This shortened NC HA was cloned using an EcoRV and a BglII restriction site via the ligation method into the cloning plasmid pDL, and named pDL_MTT1_NC531 (Figure 6 B, NC531 HA).

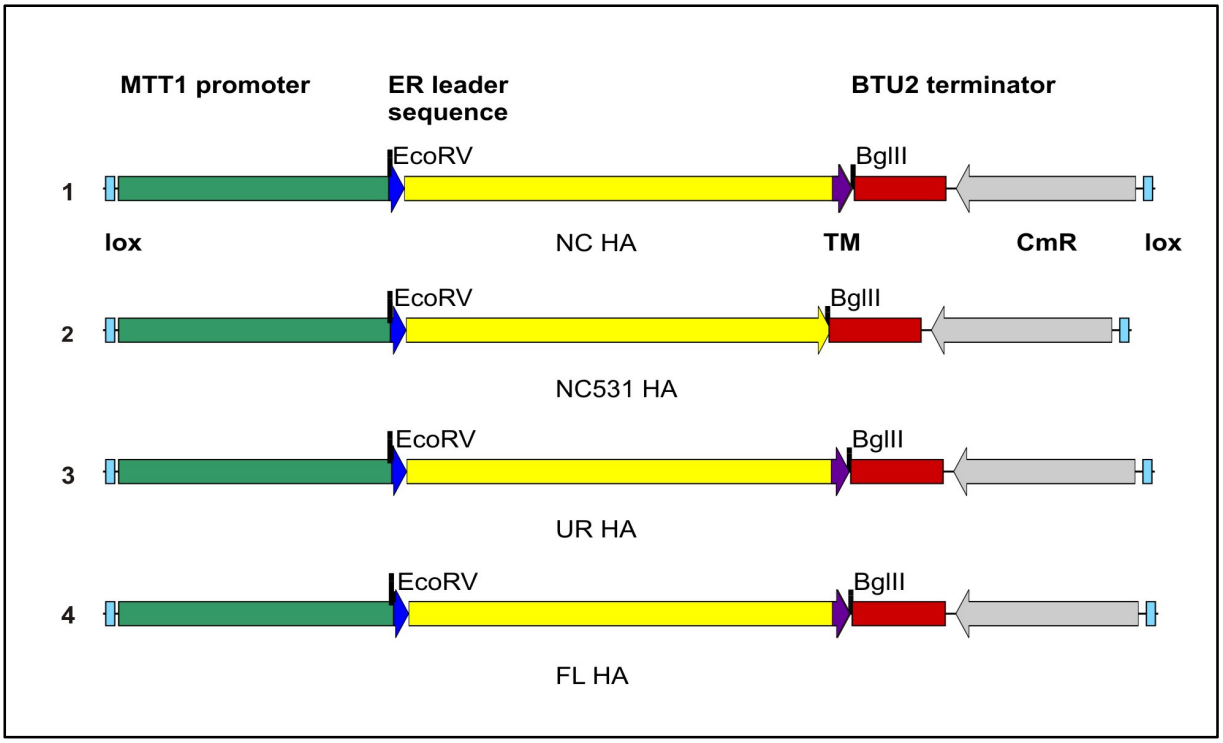

Figure 6: The expression modules.
Scheme of the created HA expression modules (linear structure). The figure demonstrates the structure concept of the expression cassettes. For all expression cassettes the MTT1 promoter (green) and the BTU2 terminator (red) were used. As ER leader sequence the endogenous signal peptide of the different HAs (dark blue) was employed. The New Caledonia derived HA (**A**, NC HA, yellow), the Uruguay derived HA (**C**, UR HA, yellow), and the Florida derived HA (**D**, FL HA, yellow) comprised a transmembrane domain (TMD, purple), which anchors the HA in the viral membrane. A truncated version of the New Caledonia derived HA (**B**, NC531 HA, yellow) excluded the TMD. The expression cassettes were flanked by two loxP sites (lox, light blue) and contained a bacterial selection marker (CmR, chloramphenicol resistance, grey). The synthetic genes of the HAs were cloned in order to generate the expression cassette by EcoRV and BglII sites.

Subsequently, the expression cassettes of the four created working plasmids were transferred to an episomal expression plasmid (pAX) that was kindly provided by the Cilian AG via cre-dependant homologous recombination[142]. This episomal expression plasmid carries the paromomycin resistance cassette (neo2) in order to select for transformed ciliates. The obtained expression plasmids were named with reference to the cloning plasmids as pAX_MTT1_NC (Figure 7 A), pAX_MTT1_NC531 (Figure 7 B), pAX_MTT1_UR (Figure 7 C), and pAX_MTT1_FL (Figure 7 D). In general, the transformation of the cre-reaction yielded in *E. coli* clones that were able to grow on LB-agar plates supplemented with chloramphenicol, ampicillin and sucrose. The purified plasmids of representative clones were analyzed by digestion using specific restriction endonucleases to control the success of the recombinase approach. For the separation of DNA fragments, a 1 % agarose gel electrophoresis was performed. As demonstrated in Figure 7, the restriction analysis regarding the clones obtained from the transformation of the cre-reaction revealed correct DNA

fragments for the different expression plasmids. Furthermore, the sequence of the truncated NC531 was verified via sequencing analysis (data not shown).

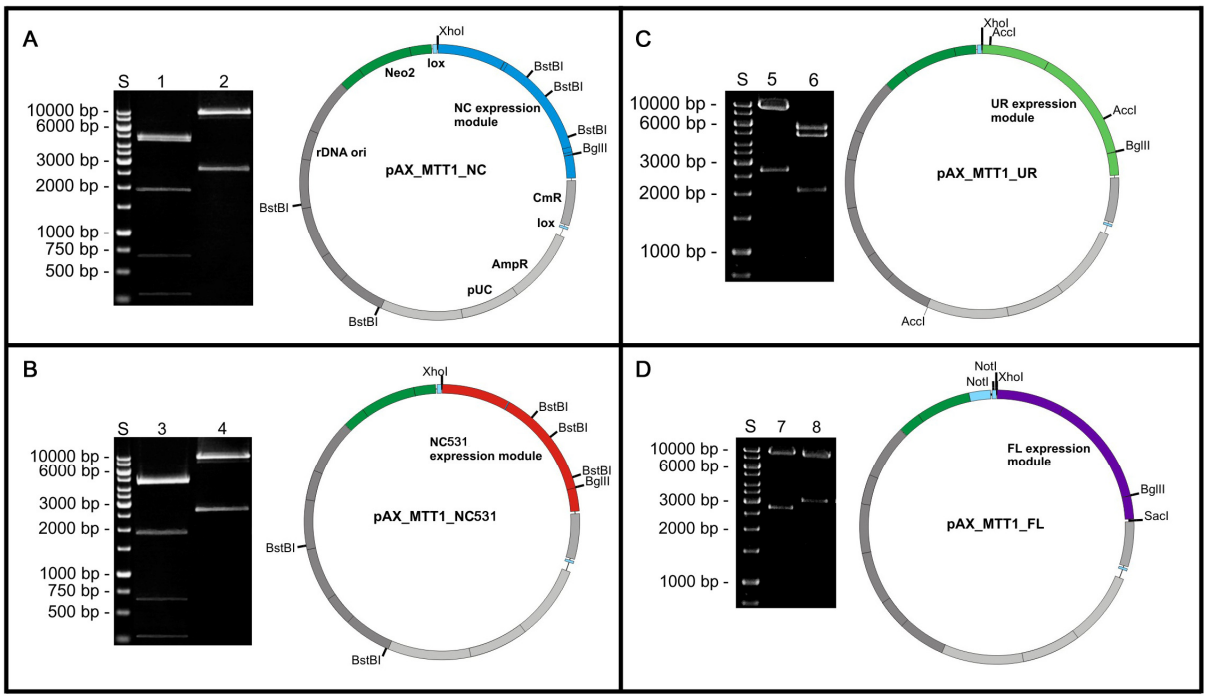

Figure 7: Restriction analyses of the expression plasmids.
Agarose gel electrophoresis of the restriction analyses using pAX_MTT1_NC (**A**), pAX_MTT1_NC531 (**B**), pAX_MTT1_UR (**C**), and pAX_MTT1_FL (**D**) are demonstrated. The restriction analysis of the four expression plasmids confirms successful generation of the expression plasmids.
A: Examination of pAX_MTT1_NC.
1: The restriction analysis with BstBI is shown (4826 bp, 4695 bp, 1929 bp, 669 bp, 381 bp). 2: Analysis due to a XhoI and a BglII site (9761 bp, 2703 bp).
B: Testing of pAX_MTT1_NC531.
3 and 4: The analogous approach as shown in A. Lane 3 (4826 bp, 4551 bp, 1929 bp, 669 bp, 381 bp) and lane 4 (9761 bp, 2595 bp).
C: Analysis of pAX_MTT1_UR.
5: The restriction analysis with XhoI and BglII is shown (9761 bp, 2661 bp). 6: Analysis using an AccI site (5519 bp, 4863 bp, 2040 bp).
D: Investigation of pAX_MTT1_FL.
7: Analysis due to a XhoI and BglII restriction (9761 bp, 2649 bp). 8: A NotI and SacI restriction is shown (9346 bp, 3022 bp, 42 bp).
A 1 kb DNA ladder was used as standard (S).

4.1.2 Generation of HA expressing *T. thermophila*

After the construction and assembly of the expression vectors, the generation of transformed *Tetrahymena* cells was performed. Conjugating *T. thermophila* wild type cells were transformed using biolistic bombardment, applying the four different expression plasmids and the successful transformation was monitored by selection against the antibiotic paromomycin in a final concentration of 150 µg/mL.

After approximately three weeks of growth, paromomycin resistant clones were verified performing a molecular analysis. Therefore, the whole genomic DNA was extracted from

putative transformants and wild type cells, and the DNA was examined via PCR analysis. Figure 8 exhibits the verification of one representative *T. thermophila* culture transformed with pAX_MTT1_NC (A), pAX_MTT1_NC531 (B), pAX_MTT1_UR (C), or pAX_MTT1_FL (D). For a proper transformant, a fragment corresponding to a part of the MTT1 promoter and the HA coding region (PCR 1), a fragment corresponding to the entire HA coding sequence (PCR 2), and a fragment that reflects the MTT1 promoter (PCR 3) was amplified and visualized performing an 1 % agarose gel electrophoresis. For the wild type, a proper signal regarding the endogenous MTT1 promoter (PCR 3), but not for the HA (PCR 2) and MTT1-HA-fusion (PCR 1) sequence was observed, as expected.

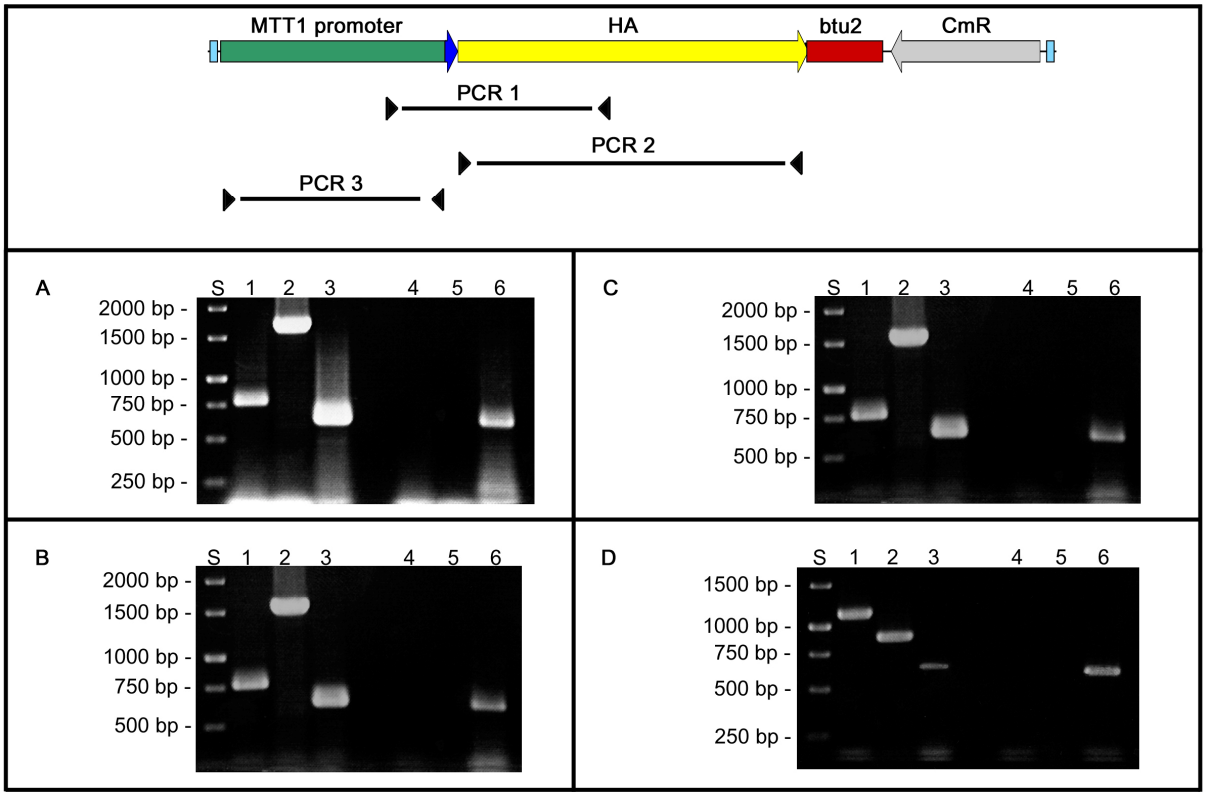

Figure 8: Analyses of representative transformed Ciliates.
Results of the PCR analysis of *T. thermophila* transformed with pAX_MTT1_NC (**A**), pAX_MTT1_NC531 (**B**), pAX_MTT1_UR (**C**), or pAX_MTT1_FL (**D**). The 1 % agarose gel electrophoresis confirms that the analyzed clones contain the correct expression modules. **1**: PCR 1 confirms the fusion of the MTT1 promoter to the HA coding region. **2**: PCR 2 depicts a PCR of the coding region of the HA. **3**: PCR 3 exhibits the MTT1 promoter. **4**, **5** and **6** are analogous, but applying wild type DNA. PCR approach was performed using MTT1 and target antigen specific primers. The negative controls show a proper signal regarding the endogenous MTT1 promoter, but no signals in respect to the target antigens. An 1 kb DNA ladder was used as standard (S).

4.1.3 Screening of transformed ciliates for HA expression by Western blot analysis

The verified transformants were screened for HA antigen expression and secretion. For this purpose, transformed *T. thermophila* cells were cultured in SPP medium under selection pressure (400 µg/mL paromomycin) for 12 hours. The expression was subsequently induced

by cadmium, setting up a final concentration of 5 µg/mL. Same expression strains were cultivated without induction of the protein expression by cadmium. To minimize digestion of the recombinant HA in the extracellular medium by endogenous proteases, the cysteine protease inhibitor E-64 was added to the cultures at point of time of cadmium induction. 24 hours after the induction of the HA expression cells were harvested by centrifugation and were subsequently solubilized in an extraction buffer (RIPA-buffer). Samples of the cell free culture supernatant were collected, as well. For the separation of the proteins 12 % SDS gel electrophoresis was performed. Extracts of cells and supernatants derived from *T. thermophila* cultures transformed with vectors of the New Caledonia derived untruncated HA (NC HA, Figure 9 A, theoretical molecular weight 62 kDa), and C-terminally truncated HA (NC531 HA, Figure 9 B, theoretical molecular weight 58 kDa), the Uruguay derived HA (UR HA, Figure 9 C, theoretical molecular weight 60 kDa), and the Florida derived HA (FL HA, Figure 9 D, theoretical molecular weight 57 kDa) were tested by SDS-PAGE and Western blot analysis. The proteins were transferred to a nitrocellulose membrane using semi-dry blot and recombinant HA antigens were detected using strain specific anti HA antibodies. A *Tetrahymena* wild type strain was applied as negative control. As a positive control a NIBSC antigen of each untruncated HA was taken. All samples were tested using non-reducing conditions, and reducing conditions where 2-mercapto ethanol was added to the sample buffer.

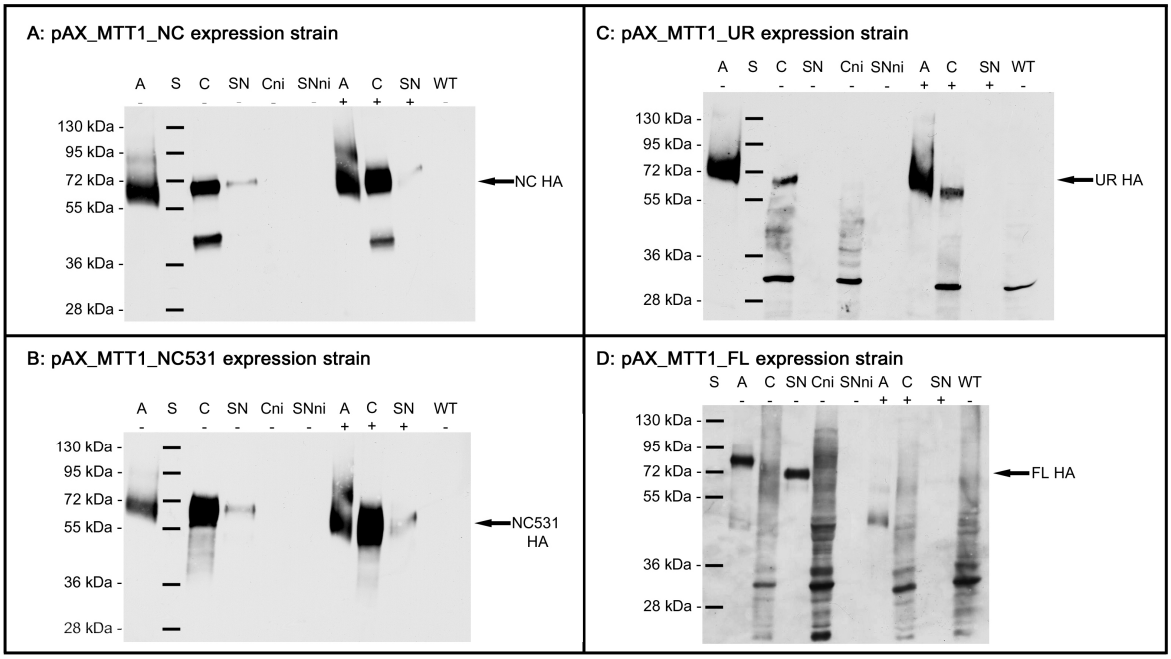

Figure 9: Expression analyses of recombinant HA in *T. thermophila*.
Aliquots of a non-transformed wild type and cell extracts of clones that were transformed with the pAX_MTT1_NC (**A**), pAX_MTT1_NC531 (**B**), pAX_MTT1_UR (**C**), and pAX_MTT1_FL (**D**) expression plasmid were separated on 12 % SDS-PAGE and transferred to a nitrocellulose membrane via semi-dry blot. Samples of the NIBSC antigen were applied using non-reducing (A-) and reducing conditions with the addition of 2-mercapto ethanol (A+). Cell extracts of every expression strain were applied, taking non-reducing conditions with induction by cadmium (C-) and without induction by cadmium (Cni-), as well as supernatants with induction by cadmium (SN-) and without (SNni-). The cell extracts were also applied utilizing reducing conditions (C+) as well as the supernatants (SN+). Cell extracts under non-reducing conditions of a non-transformed wild type strain served as negative control (WT-). As protein standard the PageRuler™ Plus Prestained Protein Ladder (S) was applied.

Analyzing the transformants of the NC expression strain HA specific bands were detected in the cell extracts indicating HA expression with an apparent molecular weight of 62 kDa using non-reducing and reducing conditions as also shown for the NIBSC NC antigen (Figure 9 A). Sample preparation with and without the addition of 2-mercapto ethanol was performed in order to test the stability of the recombinant expressed antigen. Furthermore, slight signals were detected in the supernatant of the pAX_MTT1_NC transformant (Figure 9 A). Applying the cell extract and the supernatant of wild type *Tetrahymena* cells and supernatants of a non-induced transformant (negative control) no signal was shown, as expected. Testing transformants of the NC531 expression strain lacking the TMD, strong bands were visualized in cell extracts under non-reducing and reducing conditions comprising an apparent molecular weight of 58 kDa (Figure 9 B). The applied positive control (NIBSC NC antigen) was the same as for the NC expression strain. This NIBSC antigen could be used, because the epitope where the antibody binds is located within the HA1 chain. Slight signals testing the supernatants were also visualized for the NC531 expression strain. For the non-transformed

wild type control and a non-induced transformant no HA specific band was detected. The results of the analysis of the UR expression strains are depicted in Figure 9 C. UR antigen specific signals appeared, analyzing cell extracts applying non-reducing and reducing conditions with an apparent molecular weight of 60 kDa, as well as for the NIBSC UR antigen. No signals were found testing the supernatants of pAX_MTT1_UR transformed *T. thermophila*, non-transformed wild type cells and the cell extract and the supernatant of the non-induced UR expression strain. When analyzing the cell extracts of UR transformants and non-transformed wild type cells, an additional signal appeared at approximately 30 kDa within the controls (Figure 9 C), which could be explained by a cross reaction of the used antibody. Testing the pAX_MTT1_FL transformants applying non-reducing conditions, a strong signal that points to FL derived HA was detected within the cell free culture supernatant (Figure 9 D). No band can be visualized applying the supernatant with reducing conditions. The visualized band had an apparent molecular weight of approximately 72 kDa testing the cadmium induced supernatants. No FL HA specific band could be visualized regarding the non-induced supernatant of the FL expression strain. Applying the NIBSC FL antigen resulted in a FL HA specific band of an apparent molecular weight of 90 kDa (Figure 9 D). Testing the cell extracts of the pAX_MTT1_FL transformant, nearly all bands that were detected were also seen in the non-transformed wild type cells. This can also be explained by a side reaction of the used FL HA antibody.

The fact that signals for HA expression strains were detected for cadmium induced cultures but not for non-induced cultures pointed out that the expression of the recombinant HAs was regulated properly by the cadmium inducible metallothionein 1 promoter (MTT1) system. The main focus of this initial experiment was to illustrate that HA from different influenza virus strains can be expressed by the ciliate system. Nevertheless, to give an estimation of the yield of the expressed HA, quantitative Western blot analysis was performed.

4.1.4 Lab-scale expression of HA (0.5 L) and quantification of the yield

For the development of an industrial production process, the evaluation of an optimal fermentation procedure is of critical importance, because parameters such as medium components, pH values, oxygen supply and duration of the fermentation process significantly affects product concentration, yield and productivity. For biopharmaceuticals, medium costs can substantially affect the overall process economics.

To examine the production efficiency of the developed HA expression system on lab-scale, fermentation of the recombinant strains was carried out in a 0.5 L multi fermentor at 30 °C. Applying this multi bioreactor, different *Tetrahymena* clones and different cultivation conditions can be tested parallely. Therefore, in an initial attempt the fermentation process was established setting optimal growth conditions regarding the pH values and the oxygen supply of the cells. Testing different media components, supplemented proteose peptone (SPP) medium was ascertained to be optimal for the cultivation of HA transformed *Tetrahymena* cells. The cysteine protease inhibitor E-64 was added to prevent digestion of extracellular HA. The fermentor was controlled to maintain the broth temperature at 30 °C, pH 7.0, and dissolved oxygen at 20 % of the air saturation level. For the growth of the cells, a fed-batch culture method was chosen, because it has become a platform technology for the large-scale production of therapeutic and recombinant proteins due to its simplicity, flexibility and a high final product concentration[166]. Inoculation of the culture broth was done with a cell titer of 20,000 to 40,000 cells/mL obtained from a preculture of 200 mL SPP medium under selection pressure (paromomycin 400 µg/mL). Cells were concentrated performing a centrifugational step, and the fermentor was inoculated with 10 mL of concentrated cells. In the first phase of the fermentation process, cells were grown batch-wise for approximately 43 hours exhibiting a cell density of approximately 1.5 x 10^6 cells/mL. The second phase of the fermentation procedure was initialized by a feeding step (Figure 10, feeding) with 10 mL of 50 fold concentrated SPP medium, and the final volume of the culture broth was about 0.5 L. Cells reached a cell density of approximately 1.1 x 10^6 cells/mL for the NC HA expression strain, 1.5 x 10^6 cells/mL for the NC531 HA expression strain, 2 x 10^6 cells/mL for the UR HA expression strain, and 1 x 10^6 cells/mL for the FL HA expression strain. At this point of time the cultures were induced by cadmium chloride in a final concentration of 40 µg/mL after approximately 47.5 hours of growth (Figure 10, Cd). This cadmium concentration was defined as optimal for the recombinant HA production for all expression strains. Samples were taken in minimum at about 23 and 43 hours after the induction of the HA expression (Figure 10, harvest 1 and 2), and analyzed via Western blot. Finally, cells were harvested 43 h after induction of the expression. At this point of time, the NC HA expression strains reached a cell density of 2.6 x 10^6 cells/mL, while the pAX_MTT1_UR and pAX_MTT1_FL transformed cultures reached a cell density of 2.1 x 10^6 cells/mL for the UR HA strain, and 1.46 x 10^6 cells/mL for the FL HA strain during the fermentation process. The harvest was carried out by a centrifugational step, and samples were collected for a pretesting and a following quantification and were frozen at -20 °C immediately. Figure 10 shows the

growth curve of an optimized fermentation process using a 0.5 L fermentor of pAX_MTT1_NC (blue line), pAX_MTT1_NC531 (red line), pAX_MTT1_UR (green line), and pAX_MTT1_FL (purple line).

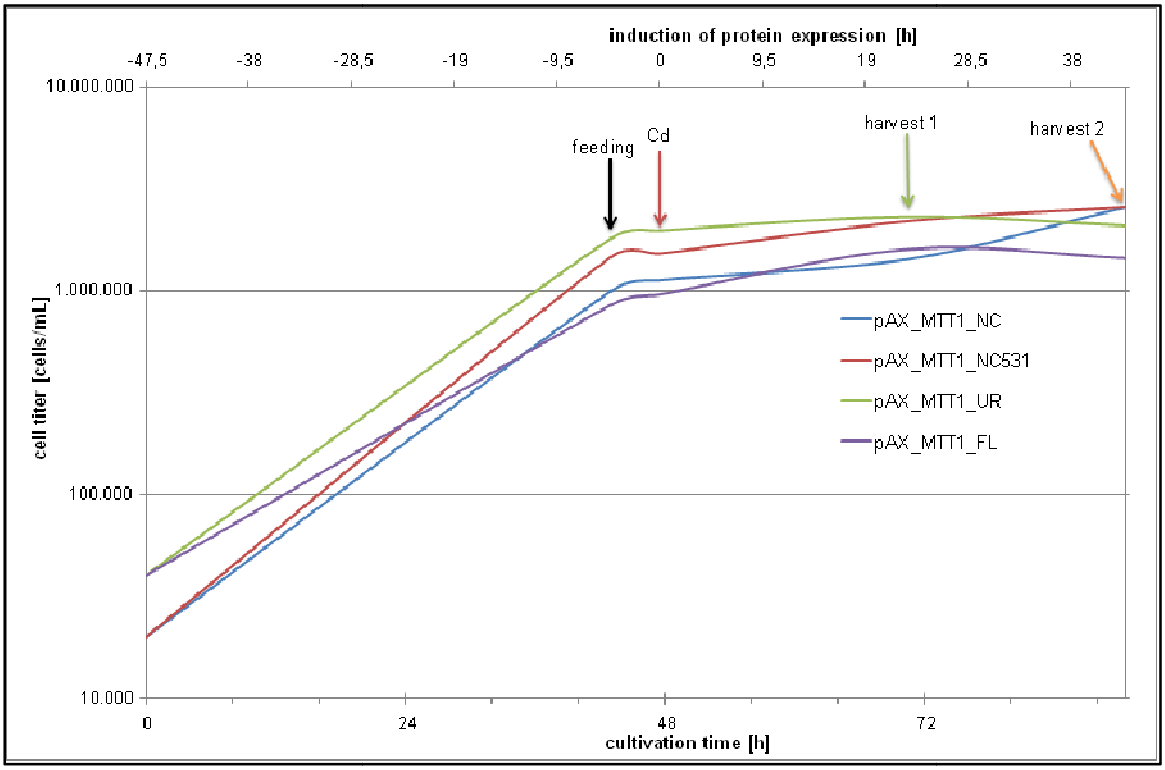

Figure 10: Fed-batch fermentation of HA expressing *T. thermophila* transformants of pAX_MTT1_NC (blue line), pAX_MTT1_NC531 (red line), pAX_MTT1_UR (green line), and pAX_MTT1_FL (purple line). The growth curves are depicted logarithmically.
The cells were grown in a 0.5 L multi fermentor using SPP medium. The fermentation broth temperature was maintained at 30 °C, regulating a pH of 7.0, and an oxygen supply of 20 % of the air saturation level. Ten milliliter of nutrient was supplemented to the culture broth (black arrow, feeding) adding 50 fold concentrated SPP medium. The cadmium induction was carried out after cells reached the stationary phase (red arrow, Cd). The harvest (green arrow harvest 1 and orange arrow harvest 2) was performed 70 hours and 90 hours after inoculation, respectively, meaning 23 and 43 hours after the induction of the expression by cadmium. Cell densities at different points of time of the fermentation process are shown as lines for the respective *Tetrahymena* expression strain.

After separating the culture from the fermentation broth by centrifugation, cell extracts as well as the supernatants were analyzed for recombinant antigen expression by SDS-PAGE and Western blot. A strong expression of NC and NC531 antigen regarding the cell extracts was demonstrated using the aforementioned fermentation conditions and points of time for harvest. For NC HA transformed *Tetrahymena* cells additional signals were detected, correlating in size with NC HA dimers and trimers at approximately 120 kDa and 270 kDa (Figure 11). Additionally, a signal appropriate to recombinant NC and NC531 HA was also detected analyzing the supernatant. Besides, the activity of an intracellular enzyme, the isocitrate dehydrogenase (ICDH), was measured. The finding that the ICDH activity in the extracellular medium exceeded 50 % of the total activity indicated that cell lysis occurred at

high rates. This pointed out that the extracellular detected signal was based on a leakage of the cells.

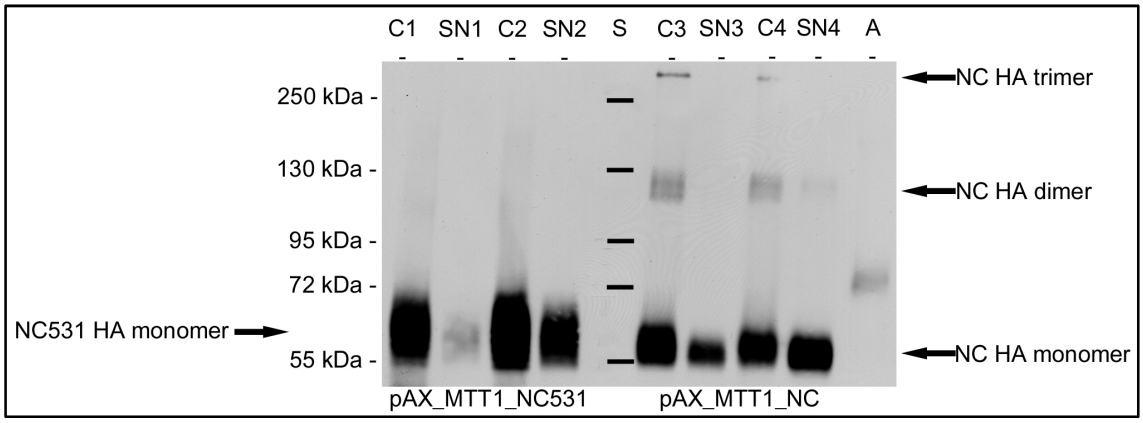

Figure 11: Analysis of the production of recombinant NC531 and NC HA during fermentation (0.5 L).
A NC531 and a NC expression strain are shown after 70 and 90 hours of fermentation. Cell extracts and supernatants were applied on an 8 % SDS-PAGE. Following, Western blot analysis was performed. Recombinant NC antigen (A-, NIBSC, 6 ng, non-reducing conditions) served as positive control. Aliquots of 250 cells and the supernatant corresponding to 1,000 cells were applied to the SDS-PAGE under non-reducing conditions.
Within the cell extract of NC531 strains after 70 hours (C1-) and after 90 hours (C2-) of the fermentation, meaning 23 hours and 43 hours after induction of the expression by cadmium, strong NC531 HA specific signals were found. Regarding the supernatants, a signal after 23 hours (SN1-) of the induction of the expression was detected that increased in intensity after 43 hours (SN2-) of induction. Analyzing the NC HA strain, a signal at point of time of harvest of 70 hours (C3-) and 90 hours (C4-) was found regarding the cell pellet and the supernatants, respectively (SN3- and SN4-). Additional signals were detected at approximately 120 kDa and 270 kDa what can be due to dimers and trimers. As protein standard the PageRuler™ Plus Prestained Protein Ladder (S) was applied.

A first estimation of the yield of the recombinant NC531 and NC antigen was made by performing a quantitative Western blot analysis. To draw up a calibration curve, different amounts of recombinant NC antigen (5 ng, 10 ng, 20 ng, 30 ng) were applied on a SDS-gel and transferred to a nitrocellulose membrane using semi-dry blot. The produced HA antigen of transformed *T. thermophila* cells obtained from the established fermentation within the multi bioreactor after 43 hours of the induction of the expression by cadmium served as samples. Densitometric analysis of performed quantitative Western blot led to the estimation that the signals correspond to 102±25 mg/L of recombinant untruncated NC antigen, and 185±25 mg/L of recombinant C-terminally truncated NC531 antigen.

Following, the UR expression strain was tested by SDS-PAGE and Western blot analysis during a fermentation process. UR HA specific signals were detected regarding the cell extracts after the induction of the recombinant UR HA expression analyzing the cell extracts (Figure 12). No signals were detected in the supernatants. The determination of the yield of

the UR transformant was performed as mentioned above and led to an expression value of 186±49 mg/L of UR derived HA. This is comparable to the expression value of the shortened New Caledonia derived HA.

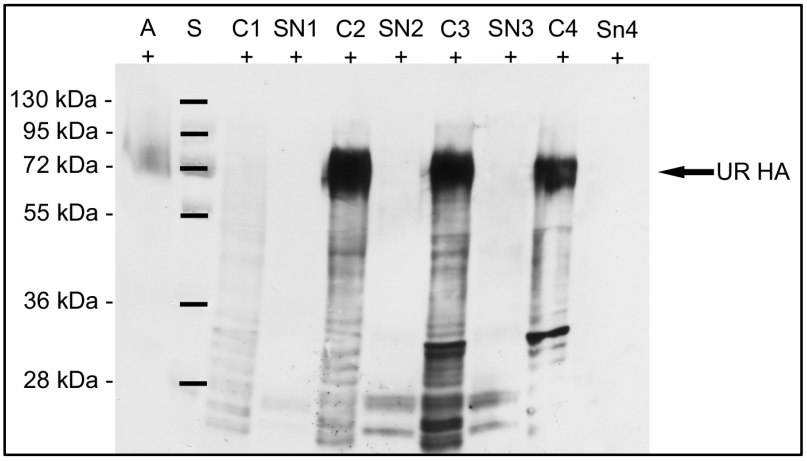

Figure 12: Analysis of recombinant UR HA expression during a fermentation process (0.5 L).
Aliquots of 10,000 cells and the supernatant corresponding to 10,000 cells were applied to a 12 % SDS-PAGE using reducing conditions. A non-induced UR HA expression strain is shown after 47.5 hours of growth (C1+, SN1+), where no specific signal was detectable. Recombinant UR antigen (A+, NIBSC, 15 ng) served as positive control. After the induction of the UR HA expression at point of time of 54 hours of growth an UR HA specific signal at approximately 70 kDa was detected in the cell extract (C2+), but not in the supernatant (SN2+). This point of time was about 6.5 hours after induction of the expression by cadmium. After 70 hours (C3+, SN3+), and after 90 hours of growth (C4+, SN4+), meaning 23 hours and 43 hours after induction of the expression by cadmium, UR HA specific signals were found in the cell extracts. As protein standard the PageRuler™ Plus Prestained Protein Ladder (S) was used.

Finally, the FL HA expression strain was tested for FL derived HA antigen expression and secretion obtained during a 0.5 L fermentation process. No FL HA specific signal could be visualized testing the cell extracts due to unspecific binding of the antibody (data not shown), however, strong FL HA corresponding signals were detected regarding the extracellular medium (Figure 13). Signals for recombinant expressed FL antigen (calculated molecular weight was about 57 kDa) were found at an apparent molecular weight of approximately 72 kDa and 90 kDa. This increased molecular weight could be due to glycosylation or folding effects of the produced protein. The same effect could be detected applying the FL specific NIBSC antigen.

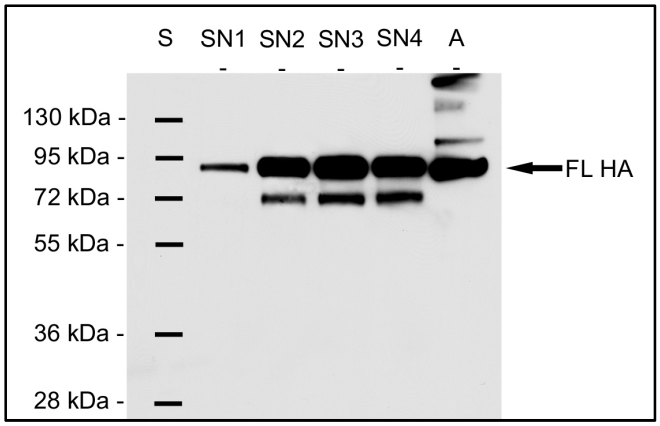

Figure 13: Analysis of the expression of recombinant HA of a pAX_MTT1_FL transformant peforming a fed-batch fermentation process (0.5 L).
Obtained supernatants were applied to a 12 % SDS-PAGE using non-reducing conditions followed by Western blot analysis. The applied NIBSC FL antigen showed a signal corresponding to approximately 90 kDa (A-, NIBSC, 45 ng). Different points of time of harvest after the induction of the FL antigen expression (SN1- 3.5 h, SN2- 20.5 h, SN3- 23 h, SN4- 43 h) did not lead to an accumulation of the detected extracellular HA derived from the Florida B strain after 20.5 h of the cadmium induction. A band at approximately 72 kDa and 90 kDa appeared. As protein standard the PageRuler™ Plus Prestained Protein Ladder was applied (S).

Additionally, a quantification of the yield applying the extracellular medium of the FL expression strain was performed and determined as 21±8 mg/L of recombinant expressed FL derived HA antigen. The measured extracellular ICDH activity was about 28 % but the estimated cell titer during the fermentation was constant from the point of time of induction of the FL antigen expression by cadmium until the harvest of the cells. Furthermore, the measured extracellular recombinant FL antigen appeared as a strong signal testing the supernatant, and no signal could be visualized analyzing the cell extracts (data not shown). Therefore, it can be concluded that the FL derived HA antigen was secreted despite the relative high measured ICDH activity. The results of the Western blot analysis are depicted in Figure 13.

Figure 14 summarizes the determined expression yields of several fermentations. While acceptable yields of truncated recombinant NC531 antigen and the recombinant expressed UR antigen could be reached, the expression yield of the untruncated version, the NC antigen, was significantly lower. The lowest expression yield was obtained for the HA derived from the FL expression strain. But it is mentionable that this FL derived HA antigen was secreted using the *T. thermophila* based expression system.

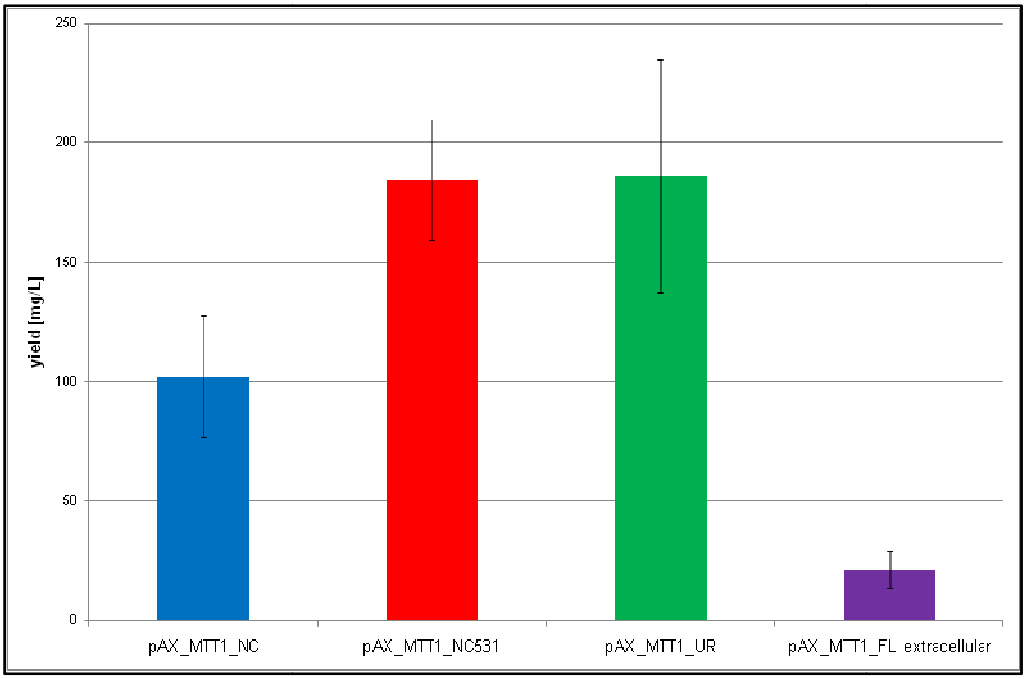

Figure 14: Expression yields of the four different recombinant HAs produced with the *T. thermophila* expression system.
The yield of intracellular recombinant expressed HA from a pAX_MTT1_NC transformant obtained in eight independent 0.5 L fermentation process was determined to be 102±25 mg/L (blue bar, n=8). The yield of intracellular NC531 HA was calculated to be 185±25 mg/L (red bar, n=6), of intracellular UR derived HA to be 186±49 mg/L (green bar, n=8), and of extracellular FL derived HA to be 21±8 mg/L (purple bar, n=9). The standard deviations are shown by error bars.

4.2 Development of an efficient production and purification process

After the quantification of the yield of recombinant expressed HA, the NC531 HA derived from the influenza virus New/Caledonia/20/99 turned out to be the ideal candidate for further investigations. First, an up-scaling process was established to obtain enough recombinant protein to enable the development of a NC531 HA specific purification process based on same cell material. The purification process itself should lead to a purity acceptable for an *in vivo* testing. Finally, it should be tested if the purified NC531 antigen exhibits a glycosylation pattern, which is necessary for immunogenicity of the antigen when using as a vaccine candidate.

4.2.1 Production of the NC antigen performing a 50 L fed-batch fermentation

The NC531 expression strain was applied for further studies because of the high yield achieved testing this transformant in lab-scale (0.5 L) and because of the absence of a transmembrane domain that might interfere with an extraction of recombinant protein from the cells. To obtain sufficient material for the evaluation of a purification process, an

up-scaling by using a 50 L fermentation vessel was necessary as a first step. The fermentation was performed as a fed-batch cultivation because this method showed best results in small scale fermentation (0.5 L). SPP medium was chosen because high cell densities were reached and a high productivity was shown as aforementioned (4.1.7). Inoculation was performed with a cell titer of 30,000 cells/mL. The feeding with 750 mL of 50 fold SPP medium was done after 43.5 hours of the fermentation procedure giving a final volume of 37 L (Figure 15). The fermentor was controlled to maintain the broth temperature at 30 °C, and pH value of 7.0 with 20 % dissolved oxygen, as found out during the 0.5 L fermentation process. After the culture reached the stationary phase, induction of the NC531 HA expression was carried out by addition of cadmium chloride in a final concentration of 40 µg/mL after 48.5 hours of growth (Figure 15). Cells reached a final cell density of 0.95 x 10^6 cells/mL and were harvested after 90.5 hours of fermentation, meaning 42 hours after induction of the expression by cadmium (Figure 15).

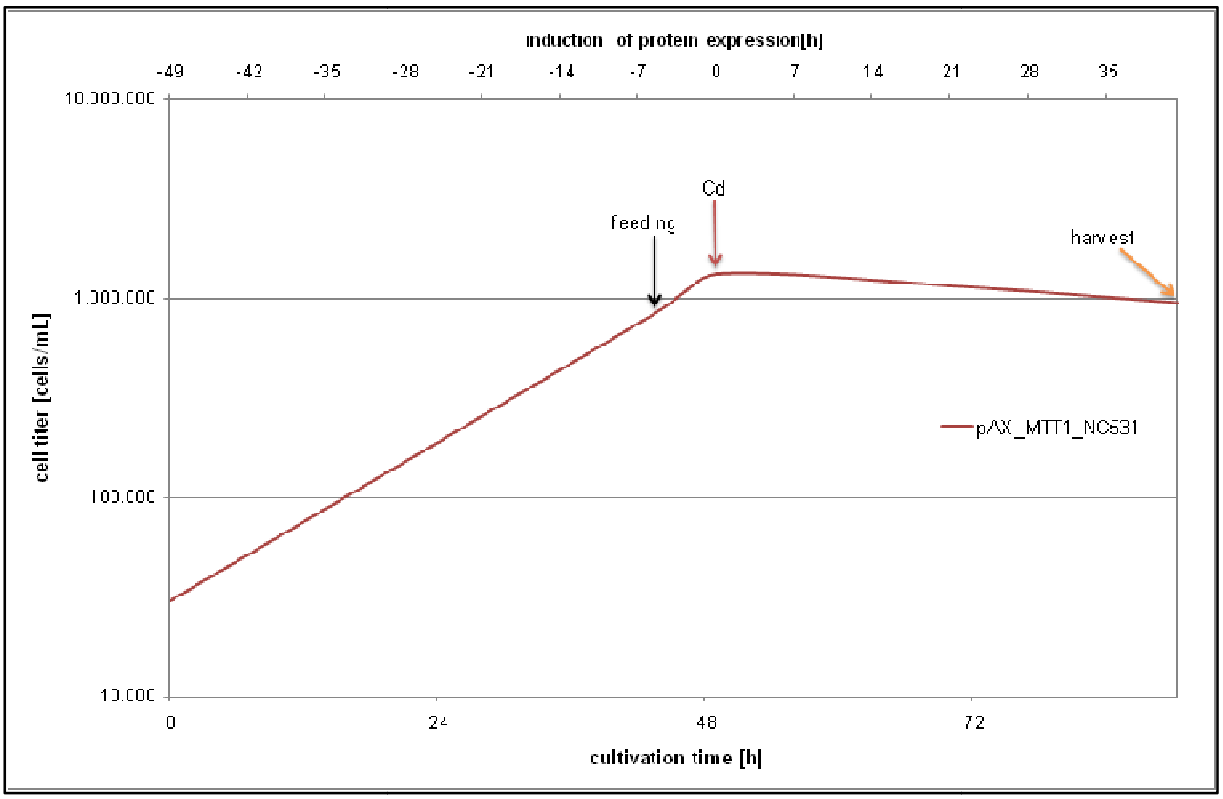

Figure 15: Fed-batch fermentation of the NC531 HA expressing *T. thermophila* production strain. Growth curve is depicted in logarithmic scale.
The cells were grown in a 50 L fermentor using SPP medium. The fermentation broth temperature was maintained at 30 °C regulating the pH at 7.0, and an oxygen supply of 20 % of the air saturation level. Feeding (black arrow) was done after 43.5 hours of growth, and the induction of the expression by cadmium was done when cells reached the stationary phase after 48.5 hours (red arrow, Cd). Harvest (orange arrow, harvest) was performed after 42 hours of the induction of the expression. The final cell densitiy was about 0.95 x 10^6 cells/mL, depicted as red line.

After 42 hours of induction of NC531 protein expression, the culture was separated from the fermentation broth applying a hollow fiber module and a following centrifugational step. To enhance the yield of the recombinant expressed NC531 antigen and to ease the purification process, the harvest was optimized by an additional step. Cells were shocked on ice and washed with ice cold phosphate buffer after the separation of the cells from the fermentation broth to remove mucocyst contents and remaining media components. Mucocysts, extrusomes of *Tetrahymena* and other ciliates, have the capacity to secrete a dozen acidic polypeptides. The function of these organelles in *Tetrahymena* is unknown, but nutrient trapping has been proposed as a possibility along with membrane accretion and defense against predators. These expanding protein structures would interfere with the purification using column chromatography. Finally, the washed and pelleted cells were frozen at -80 °C, immediately.

In order to determine the expression yield, samples corresponding to 10, 50 and 100 cells and aliquots of the supernatant were analyzed by an 8 % SDS-PAGE and a following Western blot analysis using non-reducing conditions. The cell extract (Figure 16) and the supernatant of a non-transformed wild type strain served as negative control and the recombinant NIBSC NC antigen (NIBSC, 6 ng) as positive control. NC531 HA corresponding signals were detected at approximately 57 kDa within all samples. In addition, the formation of putative dimers and trimers was shown when applying 100 cells per lane. Extracellular recombinant NC531 HA was detected as well, but the performed ICDH activity assay exceeded 50 %, indicating that the extracellular located NC531 HA was dependant on cell lysis. Quantitative Western blot analysis revealed intracellular HA expression values of ~150 mg/L obtained during a 50 L fed-batch fermentation process.

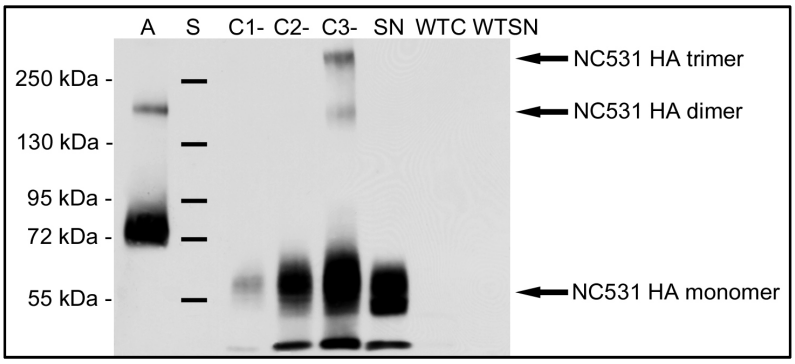

Figure 16: Expression analysis of NC531 HA obtained during a 50 L fed-batch fermentation.
An 8 % SDS-PAGE and a Western blot analysis was performed. The NC HA antigen (NIBSC, 6 ng, A) served as positive control, whereas the cell pellet (WTC) and the supernatant (WTSN) of a non-transformed wild type strain served as negative controls. Verifying the NC531 HA expression, aliquots corresponding to 10, 50 and 100 cells/lane (C1-, C2- and C3-) were applied after 42 hours of induction by cadmium. The supernatant (SN) that refers to 1,000 cells is demonstrated, as well. As protein standard the PageRuler™ Plus Prestained Protein Ladder was used (S).

4.2.2 Establishment of a purification process for HA expressed by *T. thermophila*

For industrial purposes, the purification strategies employed should be inexpensive, rapid, efficient in terms of yield, and amenable to large-scale operations. They should have the potential for a relatively high capacity and selectivity for the desired product.

To establish a purification process, the harvest was optimized as mentioned above (4.2.1). Different columns for chromatography were tested to define the best column material to use for the purification of recombinant NC531 antigen. Furthermore, different buffers, different protease inhibitors as supplementation for the buffers and different elution strategies were checked. An important point during the purification process was to maintain neutral pH values around 7.4 in order to prevent refolding of NC531 HA (personal communication by the cooperation partner).

Cell lysis and separation of cell debris.

Washed and frozen cells (80 g) obtained from the 50 L fed-batch fermentation, corresponding to 2.1 L of original culture, were thawed in loading buffer supplemented with the cysteine protease inhibitor E-64 (70 µM) and disrupted using an IKA UT T25 Ultra-Turrax with a S25N-25G rotor at 10,000 rpm for 5 minutes. This mechanical method together with the lysis by Tween®-20 was found out to be the most efficient technique to obtain completely disrupted cells without losing much recombinant NC531 antigen. After 12 hours of stirring at 4 °C, most of the NC531 HA was soluble, and the lysate was filtrated applying a hollow fiber module (0.45 µm, 850 cm², 1 L/min). The pH, which dropped to about 7.1 during the stirring step, was readjusted to 7.4 by addition of NaOH. The cell retentate of the filtration containing cell debris and with this about 30 % of the total NC531 HA was discarded.

1st column: 125 mL Capto™ Q.

The solubilized and filtrated cell lysate was used to load the first column of the purification process. Capto™ Q is a strong anion exchange chromatography material for the capture and intermediate purification of proteins from large feed volumes. This anion exchanger was chosen because the NC531 HA should be purified at neutral pH values concerning the folding properties. At neutral pH values, the NC531 antigen is theoretically negatively charged because of its theoretical isoelectric point of 6.31. All column steps were performed at 4 °C.

The column comprising a bed volume of 125 mL was equilibrated and subsequently loaded with the obtained filtrate at a flow rate of 15 mL/min. A washing step with one column volume (CV) of phosphate buffer supplemented with Tween®-20 was done. Afterwards, the washing was changed to phosphate buffer without Tween®-20 applying 2 CV to get rid of the detergence. The NC531 HA was eluted by a step to 150 mM NaCl in phosphate buffer (elution buffer 1). This salt concentration was found to be best suited for elution. In previous experiments, a NaCl gradient was used to determine this salt concentration. To check for complete elution of NC531 HA, a final step increasing the NaCl concentration to 1 M (elution buffer 2) was performed. The exemplary elution profile of this column chromatography step and corresponding amounts of NC531 HA is shown in Figure 17. This first capturing step led to a concentration of the antigen in 16 fractions by eliminating contaminating proteins. Testing the fractions of the 1 M NaCl step, nearly no HA was detected. Each collected fraction was supplemented by 11 µM E-64 to exclude the digestion by endogenous proteases of *T. thermophila*.

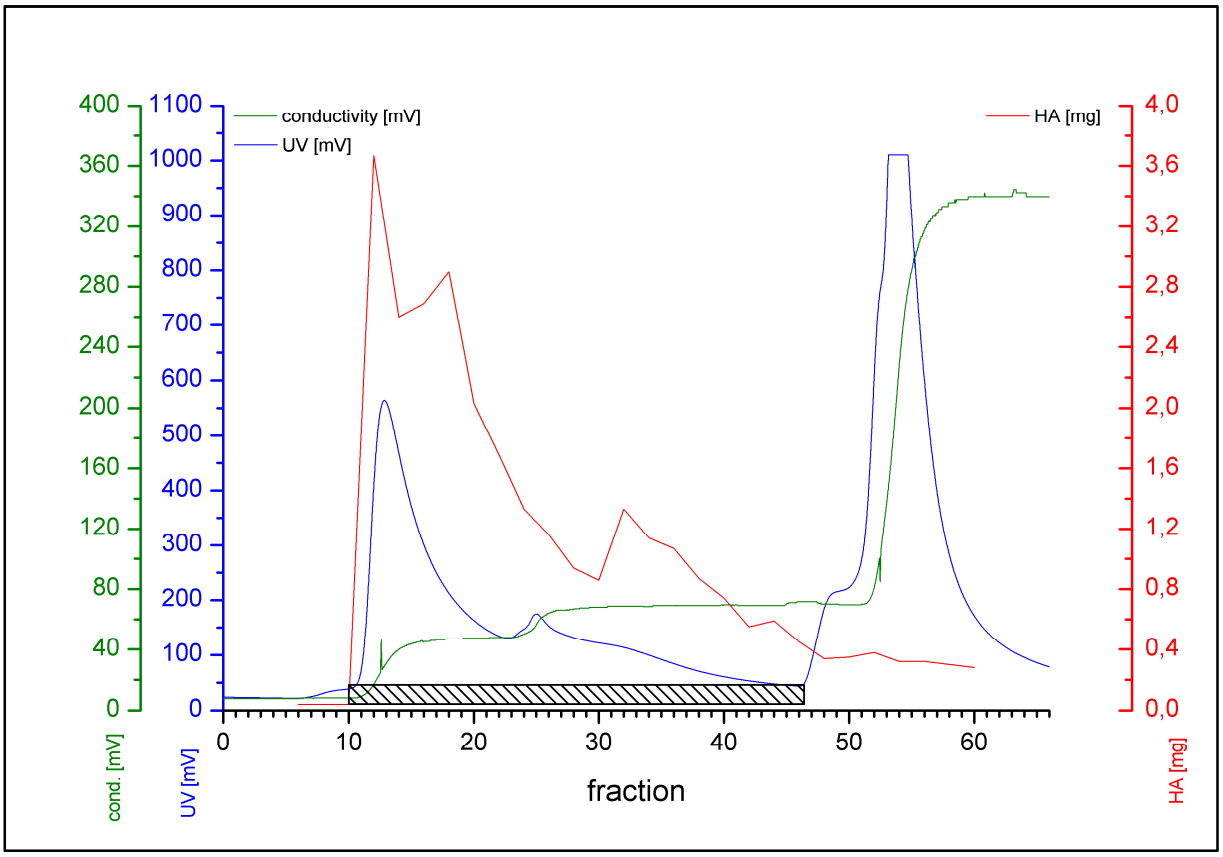

Figure 17: Elution profile of NC531 HA using a Capto™ Q column (1st column).
The red line indicates the amount of NC531 HA [mg] quantified by Western blot analysis. The blue line shows the UV values at 280 nm [relative] that reflect the whole protein content measured during the column chromatography. The green line demonstrates the conductivity [relative] related to the salt concentration of the step elution. The dashed bar indicates the fractions pooled for the 2nd column.

Second column: 25 mL Ceramic hydroxyapatite

Ceramic hydroxyapatite (CHT) overcomes the physical and chemical limitations of traditional crystalline hydroxyapatite, and provides the throughput, stability, and reproducibility required for industrial biopharmaceutical manufacturing. CHT interacts with biomolecules by multiple modes. Electrostatic interactions occur between positively charged calcium ions and negatively charged phosphate groups on the support and charged moieties on the biomolecule´s surface. Much stronger coordination complexes can form between clusters of carboxyl groups (on proteins) and calcium sites on CHT. Repulsion effects and the geometric charge distribution on CHT provide unique selectivity. Typically, proteins are bound to hydroxyapatite using a low ionic strength phosphate buffer. Elution is accomplished through the use of a phosphate gradient of increasing strength.

Fractions 10 to 47 of the Capto Q column were pooled, and a buffer exchange using running buffer 2 was carried out applying the pooled fractions to a lab-scale Tangential Flow Filtration module (TFF, 30 kDa, Pellicon, Millipore GmbH, Schwalbach, Germany) to reach a final NaCl concentration below 10 mM with a final volume of 80 mL. The pH value of this sample was adjusted to 7.5 prior to loading to the column. The column comprising a bed volume of 25 mL was loaded with a flow rate of 7 mL/min. The recombinant NC531 protein did not bind to the column but could be detected nearly quantitatively in the flow through. Performing this column step, the whole protein content could be decreased because a vast number of contaminating proteins bound to the column, and this finally led to a purer antigen (see Table 1).

Third column: 15 mL Concanavalin A.

Concanavalin A (Con A) is a tetrameric metalloprotein isolated from *Canavalia ensiformis*. Con A binds molecules containing α-D-mannopyranosyl, α-D-glucopyranosyl and sterical related residues. The binding sugar requires the presence of C-3, C-4 and C-5 hydroxyl groups for reaction with Con A. Con A coupled to Sepharose is routinely used for separation and purification of glycoproteins, polysaccharides and glycolipids. Binding of proteins occurs at a neutral pH using PBS buffer supplement with NaCl. Elution of bound substances can be achieved applying an increasing gradient (linear or step) of α-D-methylmannoside or α-D-methylglucoside.

A 15 mL Con A Sepharose 4B (GE-Healthcare) column was used. Loading and elution were performed at a flow rate of 5 mL/min. After the equilibration of the column, the flow through of the CHT column, supplemented with 150 mM of NaCl, was loaded to the Con A column. The bound NC531 HA was eluted by applying a step to 0.5 M of α-D-methylmannoside in PBS. This high concentration of sugar was discerned to be best suited for the elution of NC531 HA. An overall of 30 fractions of each 5 mL were collected and quantified via Western blot analysis. The entire yield was determined to be about 15 mg of purified NC531 HA. Figure 18 shows the elution diagram of the Con A chromatography, the corresponding NC531 HA and whole protein content. The elution was performed as a step elution. This applied step did not lead to an elution of the recombinant antigen in a circumscribed area but elutes during many fractions. Even a step to as much as 1 M α-D-methylmannoside did not result in an elution volume of 1-2 CV which indicated a very tight binding of HA to the Con A material (data not shown). The whole protein content was measured by Bradford assay, and is nearly similar to the content of NC531 HA in the fractions of the Con A column. The fractions 3 to 30, containing 0.5 M α-D-methylmannoside at a pH value of 7.4 were pooled. A buffer exchange to PBS buffer at a pH of 7.4 including a concentration of the antigen was done using a Pellicon XL biomax 50 kDa module. The starting volume was about 130 mL of the sample and finally concentrated to 25 mL. The final protein suspension showed a concentration of α-D-methylmannoside beneath in theory 0.07 mM and a yield of recombinant and purified NC531 antigen of 15 mg.

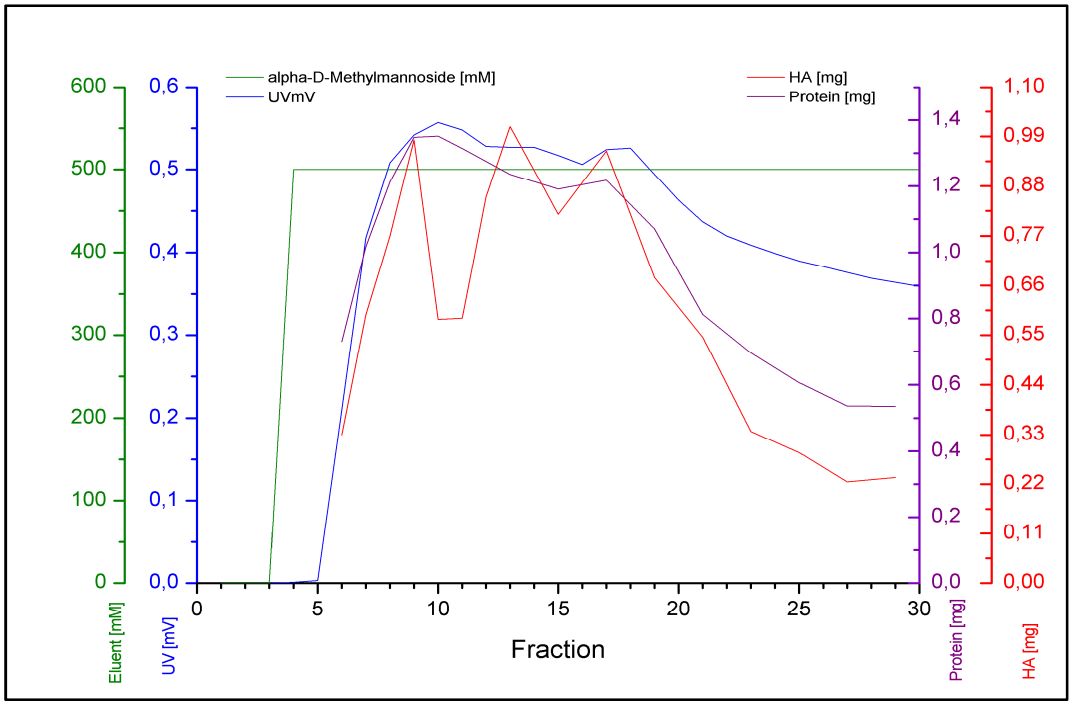

Figure 18: Elution of NC531 HA on Con A Sepharose 4B.
The red line indicates the yield of recombinant NC531 HA [mg] quantified by Western blot analysis. The blue line depicts the measured UV 280 nm [relative], and the purple line shows the whole protein content determined by Bradford assay [mg]. The green line demonstrates the estimated step of the α-D-methylmannoside.

To show the effect of the purification process, a coomassie stain was chosen to detect all protein bands of each sample. Figure 19 shows the coomassie stained SDS-gel of all samples collected during the purification process. The concentrated sample after the completed purification cascade is applied in lane 7. The purified protein fraction, eluted from the matrix and afterwards concentrated resulted in predominant bands running at a molecular weight of approximately 60 kDa. The size was in close agreement with the calculated molecular weight of NC531 HA (57 kDa). Furthermore, a band due to a NC531 HA dimer and due to a NC531 HA trimer could be visualized. To verify this result of the purification step, aliquots of this sample were separated by SDS-PAGE, blotted to nitrocellulose and finally probed with anti New Caledonia H1N1 antibody (data not shown). A clear band was detected at the same height as seen in the coomassie stained SDS-gel, which argues for an efficient purification process.

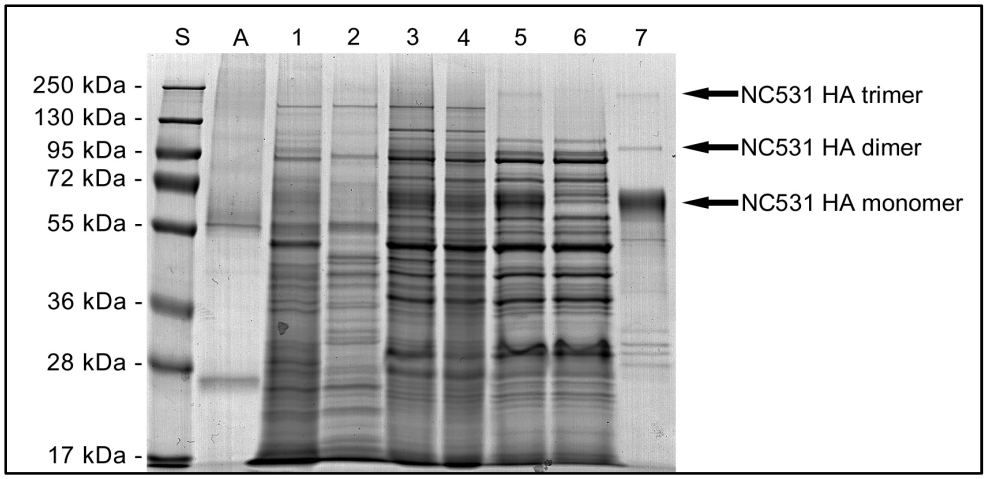

Figure 19: Purification of recombinant NC531 HA.
A coomassie stained SDS-gel (12 %) of different purification steps by Capto™ Q, Ceramic Hydroxyapatite, and Concanavalin A. As a marker the PageRuler™ Protein Ladder Plus (Fermentas) was applied (S). For the control the NC NIBSC antigen was used (**A**, 0.65 µg, NIBSC). The cell lysate after the solubilization and filtration step (lane **1**, 20 µg), the flow through of the Capto™ Q column (1st column) (lane **2**, 20 µg), the pooled fractions 10-47 of the Capto™ Q column (1st column) (lane **3**, 7 µg), the pooled fractions 10-47 of the Capto Q column after concentration and salt removal (lane **4**, 7 µg), the flow through of the Ceramic Hydroxyapatite column (2nd column) (lane **5**, 7 µg) is depicted. The flow through of the Con A column (3rd column) (lane **6**, 7 µg) showed many protein bands regarding the coomassie stained SDS-gel. After the purification cascade nearly a single band at approximately 60 kDa can be detected analyzing the pooled fractions 3-30 of the Con A column (3rd column) (lane **7**, 2 µg). A putative NC531 HA dimer and trimer can be visualized, as well.

Table 1 summarizes all quantified samples of the purification process. The purity of NC531 HA is shown in %. This purity was calculated based on a whole protein content determination using Bradford assay, and a quantitative HA antigen determination applying Western blot analysis. After all purification steps, a nearly 100 % pure *Tetrahymena* based NC531 HA was obtained. The loss of HA antigen after the concentration of the sample could be due to absorption of the antigen to the membrane of the concentration tool.

Finally, an efficient purification protocol was established. The first sample that was applied for the Capto™ Q column yielded about 4,880 mg of whole protein, with a content of recombinant HA of 56 mg. Regarding the pooled and concentrated fractions of the Con A column, a whole protein content of 16 mg was detected with 11.4 mg of NC531 HA. Concluding, 20.4 % of NC531 antigen were obtained during a 3-step column chromatography purification process, decreasing whole protein content to 0.33 %.

Table 1: Summary of the NC531 HA preparation. Protein content was determined by Bradford assay and the NC531 HA was quantified by Western blot analysis.

Sample	Protein [mg]	NC531 HA [mg]	purity of NC531 HA [%]
sample Capto™ Q	4,880	56	1.1
flow through Capto™ Q	4,537	5	0.1
fractions 10-47 Capto™ Q	376	57	15.2
sample CHT	392	42.4	10.8
sample Con A	179	33.1	18.5
flow through Con A	192	0,1	0.1
pooled Con A fractions	18	17.2	95.6
pooled concentrated Con A fractions	16	11.4	71.3

4.2.3 Characterization of the *T. thermophila* expressed NC531 antigen

Transport through ER and Golgi compartments usually is accompanied by the addition of N-glycans. Because the NC531 HA comprises a signal peptide, it is assumed that the HA was led to the ER and gets glycosylated. This point was investigated for the purified recombinant antigen NC531 and a glycosylation assay was performed by applying N-glycosidase F (PNGase F) and denaturing conditions (Figure 20). PNGase F is an amidase that cleaves between the innermost GlcNAc and asparagine residues of high mannose, hybrid and complex oligosaccharides from N-linked glycoproteins. For the test, purified NC531 HA was applied. The PNGase treated samples revealed a significant band-shift on a 12 % SDS-PAGE followed by a transfer to a nitrocellulose membrane when compared to the PNGase F untreated control (C1). Additional bands were detected regarding the PNGase F treated samples (C2 and C3), which could be due to protein degradation. The NC NIBSC antigen (A) was used as control. The significant reduction of the molecular weight of intracellular extracted NC531 HA from *T. thermophila* demonstrated that it becomes glycosylated while shuttling through the ER and Golgi. Furthermore, comparison of C1 and A showed that the NC531 antigen from *T. thermophila* and from the native virus were running at heights of comparable molecular weight. The slight reduction of the *Tetrahymena* derived NC531 HA is due to the truncation of the transmembrane domain.

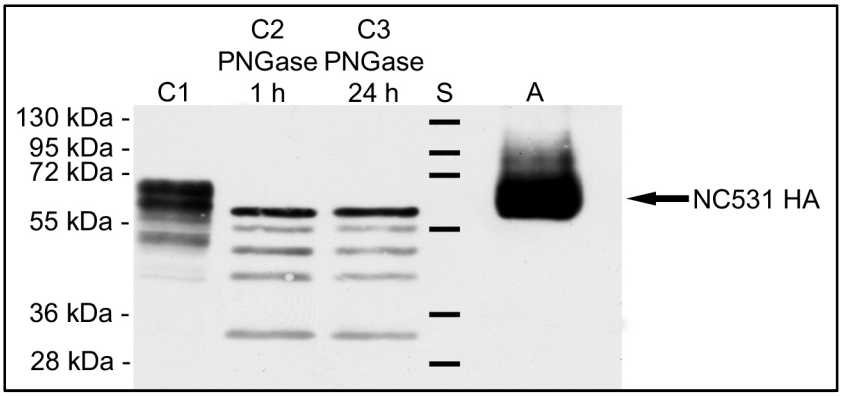

Figure 20: Deglycosylation assay of recombinant expressed and purified NC531 antigen.
The treatment with PNGase F results in a significant band-shift (C2 and C3), when compared to non-treated recombinant NC531 of *T. thermophila* (C1) and the non treated NC NIBSC antigen (A, 30 ng). The significant band-shift of the recombinant HA from *T. thermophila* demonstrates that NC531 HA becomes glycosylated while shuttling through the ER and Golgi. As protein standard the PageRuler™ Plus Prestained Protein Ladder was applied (S).

4.3 Optimization of the production process of NC531 antigen

A high expression yield of the truncated version of the New Caledonia derived HA was obtained and the purification process of this antigen was established by reaching a purity of nearly 100 %. Nevertheless, to facilitate the purification, it was addressed to the secretion improvement of the NC531 antigen as a first step. Therefore, it was tested if secretion can be reached while expressing the HA1 chain without the hydrophobic HA2 chain, the fusion domain and the transmembrane domain. By this, the purification of the secreted HA would be facilitated, and the functionality of the antigen would be preserved because the HA1 chain is responsible for immunogenicity. As a second step, a knock out strain that is not able to produce and eject mucocyst contents would ease the purification process in terms of time because additional washing steps for the harvest of the cells would not be necessary. Finally, the strong but cadmium inducible MTT1 promoter was intended to be replaced by a novel promoter candidate of equal strength due to the fact that cadmium is toxic and hence undesirable for a large scale industrial production process.

4.3.1 Improvement of the secretion of recombinant HA

Because the purification of recombinant protein from cell culture supernatant is much more preferable than the purification of intracellular protein, it was tested if the secretion of the more truncated HA1 version without the hydrophobic HA2 chain of the recombinant protein is feasible. Therefore, the HA1 coding region of the influenza virus strain New Caledonia/20/99 (aa 1-344) was amplified via PCR and afterwards cloned to the MTT1

and BTU2 coding working plasmid pDL. The generated plasmid was termed pDL_MTT1_NC344. Finally, an episomal expression plasmid was assembled by cre-dependant homologous recombination[142]. Afterwards, the generated expression plasmid pAX_MTT1_NC344, tested by restriction analysis (data not shown), was transformed into *T. thermophila* cells via biolistic bombardment. After the selection for paromomycin resistant clones and the verification of the genomic DNA by PCR analysis (data not shown), the transformants were tested for expression and secretion of the NC344 HA chain (Figure 21). To ensure that the proteases in the surrounding medium of *Tetrahymena* cells did not degrade the NC344 antigen, the cysteine protease inhibitor E-64 was added to the culture at point of time of induction of the protein expression by cadmium. To verify the results of the Western blot analysis, the degree of cell lysis during the cultivation and culture harvest was measured by determining the activity of the intracellular ICDH.

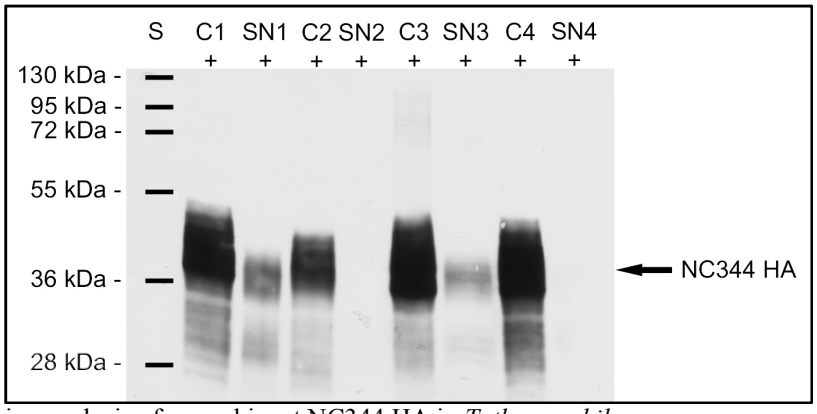

Figure 21: Expression analysis of recombinant NC344 HA in *T. thermophila*.
Aliquots of pAX_MTT1_NC344 transformed *T. thermophila* were separated on 12 % SDS-PAGE and transferred to a nitrocellulose membrane via semi-dry blot. Cell extracts of different transformants showed a strong NC344 antigen expression using reducing conditions with an apparent molecular weight of approximately 37 kDa (C1+ to C4+). Testing the supernatants regarding two clones a signal was detected (SN1+ and SN3+), and two clones showed no signal (SN2+ and SN4+). As protein standard the PageRuler™ Plus Prestained Protein Ladder was applied (S).

NC344 HA corresponding signals were detected by Western blot analysis regarding the cell pellet and the supernatant of the generated transformants at an apparent molecular weight of approximately 37 kDa (Figure 21). In the applied supernatants, NC344 HA corresponding signals were detectable in some samples. After a quantitative analysis, the yield of intracellular NC344 HA was determined to be 22±7 mg/L, whereas the quantification of extracellular NC344 HA yielded about 0.8±0.4 mg/L (Figure 22). The performed ICDH assay revealed evidence that the detected NC344 HA of the supernatant does not originate by the event of secretion because ICDH activity in the extracellular medium exceeded 15 % of the total activity. Deductively, secretion was not reached applying the HA1 chain of the

New Caledonia influenza virus strain in order to optimize the production and purification process. Therefore, it was continued using the NC531 HA, because expression yield was high and a purification process was successfully established.

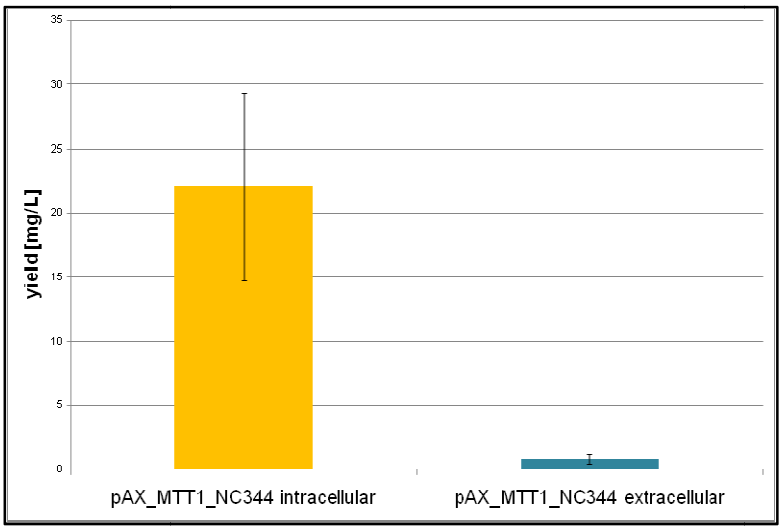

Figure 22: Intracellular and extracellular expression yields of recombinant NC344 antigen.
Expression yields were determined by quantitative Western blot analysis. Analyzing the cell extracts 22±7 mg/L of NC344 antigen were calculated (orange bar, n=10), whereas verifying the supernatant 0.8±0.4 mg/L of NC344 HA were found (cyan bar, n=9). The associated standard deviations are shown by error bars.

4.3.2 Construction of a mucocyst deficient strain to facilitate the purification of intracellular expressed recombinant NC531 HA

NC531 antigen obtained from a 50 L fed-batch fermentation process was successfully purified. Nevertheless, this process has yet to be optimized. Up to now, the purification of intracellular proteins is problematic using *Tetrahymena*, because of spring-like expansion of mucocyst contents (mucus). After the harvest of the cells, a removal step of mucus and a washing step has to be done (see Figure 23). Therefore, a knock out vector was designed resulting in a mucocyst deficient NC531 expression strain. Using this mucocyst knock out strain, it will be much easier to purify intracellularly located target proteins.

Mucocysts are vesicles specialized in the storage of highly concentrated proteins, which can be excreted in a stimulus dependent manner. It was found out that disruption of the mucocyst Granule lattice protein 3 (Grl3) leads to the disability of core formation but not to ER accumulation and failing of reaching the post-trans-Golgi network processing[92].

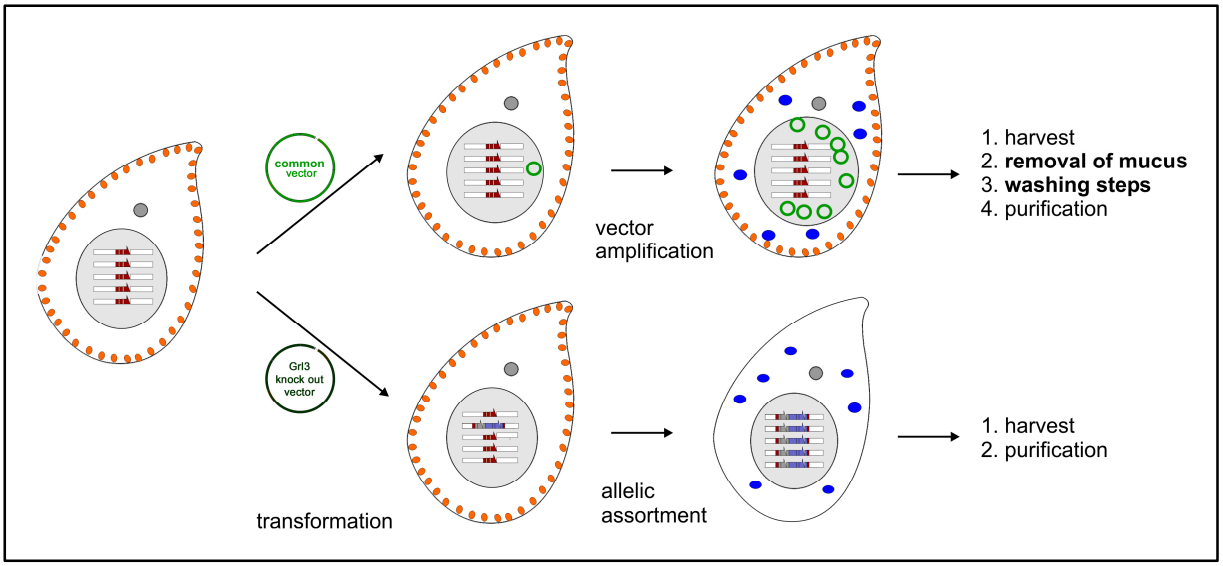

Figure 23: Comparison of the application of a mucocyst knock out strain to a common expression strain. The purification, when using the common vector system needs additional steps prior to purification, i.e. removal of mucus and additional washing steps. When using a Grl3 knock out vector, *T. thermophila* expression strains excreted no mucocyst contents. Thus, no additional steps before the purification step are necessary.

For the expression of NC531 HA in *T. thermophila*, an expression module was designed that combines the knock out of the Grl3 gene with the knock in of the expression cassette (Figure 24). The up- and downstream regions and a part of the coding region of the Grl3 gene were amplified by PCR to generate the integration flanks. Both flanks were about 1000 bp, and the cloning was performed via restriction sites that were integrated into the flanks by the primer sequence. The ligation into an equally restricted integrative vector pKOIX that was kindly provided by the Cilian AG[142] replaced its common integration sites, so that the *T. thermophila* selection marker (neo2) and the lox sites were framed by the newly designed Grl3 knock out cassette (Figure 24). Making use of cre-dependant homologous recombination, the expression module for NC531 was integrated between the two Grl3 flanks. The selection of positive transformants was done using the paromomycin resistance. After the construction, the plasmid was transformed into at least one allele of the 45 MAC alleles of the Grl3 locus. For the complete replacement of all 45 wild type alleles against the disrupted version of the Grl3 gene, the process of allelic assortment was necessary[99]. This allelic assortment took place for a duration of approximately six weeks applying increasing paromomycin concentrations up to 1,000 µg/mL cultivating the transformants in SPP medium. By the application of this integrative expression plasmid, stable transformants were obtained after the process of allelic assortment that expressed NC531 HA without producing mucocyst contents. Furthermore, after the stable integration of the expression cassette into the

T. thermophila genome the addition of the antibiotic is not necessary anymore. This is an important argument for an industrial production purpose.

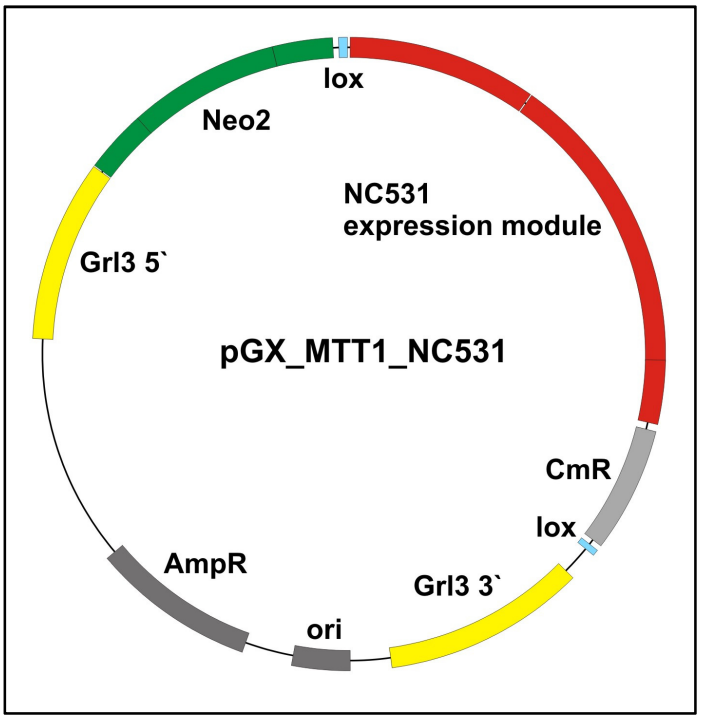

Figure 24: Grl3 knock out vector.
The NC531 HA expression cassette and the selection marker (neo2) is flanked by Grl 3' and 5' sites that promote the integration into the *Tetrahymena* macronuclear Grl3 locus.

The verification of a proper knock out of the Grl3 gene was done using RNA isolation, including a following analysis of the cDNA using specific primers for a PCR. Figure 25 shows exemplarily the PCR analysis of a Grl3 knock out transformant and of a non-transformed *Tetrahymena* wild type strain. The amplification of a small fragment of an endogenous *T. thermophila* enzyme, the β-hexasominidase (β-hex), served as positive control. The agarose gel electrophoresis of the PCR applying the non-transformed wild type led to a signal for the Grl3 (WT Grl3) gene and for the β-hex (WT hex), as expected. Analyzing the knock out transformants, a cDNA fragment corresponding to the Grl3 gene was not detectable (ko Grl3), indicating that the gene was properly knocked out and the mRNA could not be synthesized. It could be excluded that a cDNA preparation was not successful because a signal for the β-hex was demonstrated properly regarding the knock out transformant (ko hex).

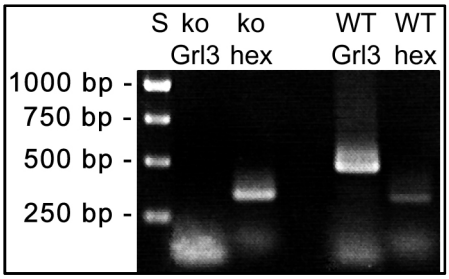

Figure 25: Agarose gel electrophoresis (1 %) of the isolated RNA of a Grl3 knock out transformant and a non-transformed wild type strain via cDNA synthesis and a following PCR. Specific primer pairs were used to verify the proper knock out. The use of the wild type cDNA led to a specific band for the Grl3 (WT Grl3) and the β-hex (WT hex). The anlogous approach was done applying the cDNA of a Grl3 knock out transformant. Here, a specific signal was detected reagarding the β-hex PCR (ko hex) but no signal was found for the Grl3 PCR (ko Grl3). A 1 kb DNA ladder was used as standard (S).

Additionally, it was tested if the excretion of mucocyst contents can be initiated by dibucaine stimulation. Synchronous secretion of mature mucocysts in *Tetrahymena* can be induced by the local anesthetic dibucaine[167,168]. The dibucaine treatment should stimulate the biosynthetic processes related to mucocyst biogenesis. Therefore, a non-transformed wild type strain and a Grl3 knock out transformant were grown in SPP medium for three days. The cells were harvested, washed, and subsequently stimulated by the addition of dibucaine. After an additional centrifugational step, no mucocysts were visible regarding the Grl3 knock out transformant but for the wild type control (Figure 26).

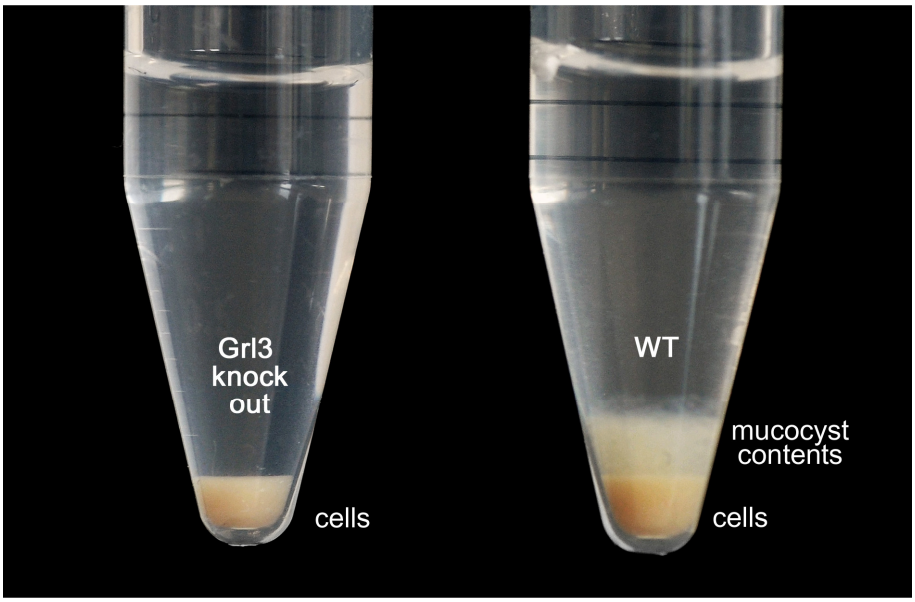

Figure 26: Dibucaine stimulation of a Grl3 knock out transformant and a wild type strain.
The addition of dibucaine to a Grl3 knock out transformant did not lead to the excretion of mucocyst contents (Grl3 knock out). For the non transformed wild type strain mucocyst contents could be properly visualized (WT).

Finally, a Western blot analysis of a Grl3 knock out NC531 expression strain was performed. A proper signal of NC531 HA was detected intracellularly (data not shown). The yield of

intracellular recombinant NC531 antigen, applying an integrative knock out expression cassette, was determined to be 24±10 mg/L.

4.3.3 Identification of promoter candidates for NC531 HA expression

The MTT1 promoter of *T. thermophila* is a strong and cadmium inducible promoter. Nevertheless, the use of this cytotoxic heavy metal is not suitable for an industrial production process. For this reason, different promoter candidates were tested for the expression of NC531 HA. The MTT5 promoter that is strongly inducible by the addition of several heavy metals, was described for the homologous expression as most inducible by the addition of divalent zinc ions. To test if the MTT5 promoter is also applicable for the heterologous expression of the NC531 HA, 1200 bp of the upstream region of the macronuclear MTT5 locus were cloned directly upstream to the coding region of NC531 HA. The native zinc inducible MTT5 promoter turned out to be the strongest promoter when pretested in small-scale. For this approach, cultures were grown using SPP medium under selection pressure (400 µg/mL paromomycin) and the expression of the NC531 antigen was induced after 12 hours of growth by the addition of zinc in a final concentration of 2 mM. Based on these findings, a 0.5 L fermentation was performed, applying same parameters as described in 4.1.7 for the cultivation of the pAX_MTT1_NC531 strain. A feeding step was done after 43 hours of growth, and after additional 4.5 hours of growth the induction of the NC531 antigen expression was performed by the addition of zinc in a final concentration of 4 mM. This higher concentration of zinc turned out to be important in order to reach higher yields of the recombinant antigen during a fermentation procedure. Cell extracts and supernatants were tested by Western blot analysis, where a strong NC531 HA specific signal was detected after 43 hours of the induction of the expression (data not shown). Following, a quantification of the yield was done. The yield of NC531 HA using the MTT5 promoter was determined as 162±9 mg/L. By this, an expression rate of this antigen was reached which is comparable to the MTT1 based NC531 HA expression rate(185±25 mg/L, see Figure 27). Hence, a promoter was found to replace the MTT1 promoter.

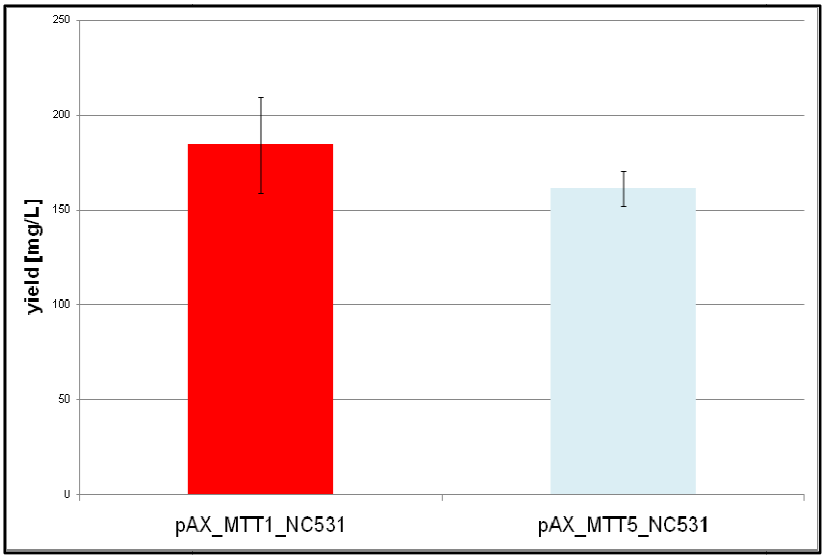

Figure 27: Analysis of expression yields of NC531 HA using the cadmium inducible MTT1 promoter or the zinc inducible MTT5 promoter. The quantification of the pAX_MTT1_NC531 transformant led to the estimation of the yield of intracellularly expressed NC531 HA of 185±25 mg/L (red bar, n=6). The yield of recombinant expressed NC531 HA applying the MTT5 promoter led to 162±9 mg/L of NC531 HA (light blue, n=3). The associated standard deviations are depicted as error bars.

5. Discussion

Heterologous protein expression becomes more and more important in the field of the production of pharmaceutical relevant agents. Recently, bacterial and yeast expression systems are used for heterologous protein expression, due to the abundance of simple genetic tools and because they can be cultivated to high cell densities including short generation times. On the other hand these systems often fail in the field of required posttranslational modifications like glycosylation and disulfide bonds which are easily integrated by mammalian cells[76]. But, this is accompanied by high costs and the risk of contaminations by human and animal pathogens and prions[169]. The possibility to use the innovative *Tetrahymena thermophila* expression system for viral antigen production is based on many favorable features, including the large-scale fermentation using inexpensive axenic media, gaining high cell densities, and the ability of *T. thermophila* to produce and secrete complex homogenously glycosylated proteins.

5.1 Establishment of a heterologous expression system for viral antigens

A proven, effective way to combat influenza is through vaccination of the public using the trivalent vaccine produced in embryonated chicken eggs. In the current process three influenza strains selected by WHO/CDC are propagated in chicken eggs, chemically inactivated, and semi-purified. The egg-based technology, however, is unable to respond to a pandemic crisis. Vaccine development and production takes several months following identification of potential target strains and typically requires the re-assortment of the target strain with a high yield strain to obtain suitable growth properties[170]. Therefore, there is a pressing need to establish innovative and effective expression systems for the production of recombinant antigens. In the recent years, two new approaches, a cell-culture based and a recombinant protein (antigen)-based system have shown promise to replace the egg-based technique[170]. But, due to several limitations the development of other expression platforms is inevitable. The cell culture-based approach involves the production of influenza viruses in cell culture followed by the current (egg-based) virus inactivation and purification for the down-stream processing. The advantages are: cell cultures are easier to handle and can be scaled up in a short period of time. Additionally, the influenza vaccines produced within this approach have been tested in Phase I and Phase II clinical trials were found to be safe and at least as effective as the vaccines produced in embryonated chicken eggs[22,171,172]. A limitation

of the cell culture-based approach is the requirement of high-yielding re-assorted viruses. This process also may introduce cell line specific mutations in the genes that can lead to the selection of variants characterized by antigenic and structural changes in the HA protein, potentially resulting in less-efficacious vaccines[173,174]. Additional hurdles include that the production and handling of dangerous virus requires the availability of a high containment facility; mammalian cells can harbour animal viruses that may lead to safety concerns. Besides, the residues from the expressing cells may cause some unknown side-effects since no thorough purification process has been introduced into the manufacturing process. The second recombinant protein-based approach involves production of viral antigens such as hemagglutinin (HA) and neuraminidase (NA) in cell culture with recombinant DNA technology and utilization of the purified antigens as the active ingredients in the vaccine. The recombinant HA influenza vaccine developed using the baculovirus/insect cell expression system, named FluBlok®, has been tested in several Phase I and Phase II human clinical trials that demonstrated safety, immunogenicity and efficacy[175-179]. In elderly adults, this recombinant HA vaccine is equally or more immunogenic than the egg-based vaccine[179]. Poor response rates to the FluBlok® were observed in the 6-35-month-old age group. One possibility may be the differences between FluBlok® and conventional egg-derived vaccines in the glycosylation and cleavage of the HA molecule, which could result in less antigenicity and less efficient presentation to the immune system in young children[180]. HA produced in insect cells is generally less glycosylated than HAs produced in chicken eggs, and, more specifically, it has been reported the complex oligosaccharides found on HAs derived from vertebrate hosts are replaced by small truncated side chains on HAs derived from insect cells[181]. In addition, while HA produced in insect cells remains mostly in an "uncleaved" form, the HA present in the egg based vaccine is primarily in the form of the cleaved HA1 and HA2 fractions[182].

To meet the challenge of a potential influenza pandemic, however, a reliable expression system and a quick, efficient downstream purification process is needed. Additional challenging issues are to create safe vaccines. Ciliates combine molecular genetic manipulations and growth characteristics of prokaryotic organisms together with the sub-cellular machinery for performing post-translational protein modifications. Moreover, ciliates have been a model organism in many research fields for decades and there is no evidence that they contain any viruses, endoparasites or DNA sequences with oncogenic potential. The differences between humans and ciliates during evolution resulted in significant differences concerning the codon bias and the whole genome organisation, indicating that it is

implausible that ciliates could be a host for human pathogens. The ciliate expression system is safe and ciliates are GRAS organisms (generally regarded as safe)[132]. Neither from the scientific literature, nor own studies argue for any bio safety problems. In addition to that, nothing has been reported about toxins. In expression hosts like fertilized chicken eggs and cell lines most of the antigens are difficult and cost intensive to produce, due to bio safety aspects (contaminations in mammalian cell lines), but additionally due to the missing up-scalability and processing time. In general, ciliates exhibit many features essential for an excellent expression system, but the potential to produce candidate vaccines provides additional and unique features that should be considered[124]. Ciliates are economical, usually giving higher yield and are less demanding in terms of time and effort relative to complex eukaryotic systems such as Chinese hamster ovary cells and baculovirus infected cell lines. Large cultures of ciliates can be grown in cheap and simple medium compositions to high cell densities up to 2 x 10^7 cells/mL, leading to dry mass weights of 60 to 70 g per litre[131]. The recombinant proteins can be targeted into the cytosol (intracellular expression), into the medium (expression and secretion), or to the surface (surface display)[58,59,113,135]. This allows an easy and fast testing of different protein variants and putative vaccine candidates, respectively. These features along with easy maintenance, easy scale-up, and inexpensive growth requirements makes *T. thermophila* as a very useful protein expression system.

Because of safety issues, and cost effectiveness, the *Tetrahymena* based expression system was addressed to express three different HAs, two derived from A strain influenza virus, and one derived from B strain influenza virus during this work. Therefore, the HA of the influenza A virus New Caledonia/20/99 (H1N1), the HA of the influenza A virus Uruguay/716/2007 (H3N2), and the HA of the influenza B virus Florida/4/2006 were chosen because they composite a potential trivalent vaccine against influenza. *Tetrahymena* is characterized by both very AT-rich genome and an unusual codon usage[107,108,119,183]. Consequently, to avoid problems in heterologous HA expression, codon adapted genes in which critical triplets were changed to the commonly used triplets of highly expressed genes in *T. thermophila* were used for the target antigens (see appendix). AT-rich genomes can cause difficulties in handling the DNA sequences. Also the AT-richness of the *T. thermophila* genome causes the main challenge in altering the ciliates phenotype. Previously described episomal expression plasmid consists of a pUC backbone that enables propagation in *E. coli*, two 1.9 kb parts forming the *T. thermophila* rDNA origin and the paromomycin resistance cassette (neo 2) that allows the selection for transformed ciliates. The empty vector is sized about 8.4 kb. Especially the rDNA origin duplicates are (3.8 kb) AT-rich sequences[147]. These

sequences are probably one reason why this plasmid tends to recombinate into the highly amplified endogenous rDNA chromosomes of the host cell[144]. Additionally, the uptake of ligation reactions and the subsequent amplification in *E. coli* often resulted in reduced and fragmented backbones and the loss of the expression cassette in the past[113,142]. The adaption of the cre-recombinase system to the ciliate expression system that comprises the assembly of expression modules within small cloning plasmids and the subsequent construction of final expression vectors by an easy and robust shuttling of the expression module was established by Weide *et al.*[142]. The cloning plasmid pDL lacks AT-rich sequences like the rDNA origin and the expression module of the HAs could be easily generated and afterwards transferred by a time-saving cre-dependant homologous recombination into the expression plasmid pAX[142]. The created constructs were named pAX_MTT1_NC for the New Caledonia (NC) derived HA, pAX_MTT1_NC531 for a C-terminally truncated version of the NC derived HA, pAX_MTT1_UR for the Uruguay (UR) derived HA, and pAX_MTT1_FL for the Florida (FL) derived HA. For the regulation of the recombinant antigen expression the strong and cadmium inducible metallothionein 1 promoter (MTT1), and the beta tubulin 2 terminator (btu2) were applied. The MTT1 promoter was chosen because it is the strongest promoter described for *T. thermophila*[134,135]. Furthermore, its inducibility is beneficial for certain approaches. Regulated systems are particularly useful, because they allow a control on timing and levels of gene expression, and are preferred to be applied in many expression systems. For the methylotropic yeast *Pichia pastoris* several promoters are described. For instance, the promoter of the alcohol oxidase I gene (*AOX1*) has been employed in most studies and applications[184]. Being tightly repressed by glucose, its transcription is highly induced upon shift to methanol as sole carbon source[185,186]. Other carbon sources which represses transcription from *AOX1* gene are glycerol and ethanol, while very low derepression of transcription occurs upon carbon starvation[186]. Furthermore, some examples of yeast promoters are those of the *MET3* gene, negatively regulated by methionine; the *PHO5* gene, negatively regulated by inorganic phosphate; the metallothionein-encoding *CUP1* gene, activated by Cu^{2+} ions, and the *GAL1*, *Gal7* and *Gal10* gene, activated by galactose and repressed by glucose[187]. The yeast metallothionein (*CUP1*) gene is a metallothionein gene like that of the *Tetrahymena* MTT1 promoter, that belongs to a group of promoters of metal-binding proteins that are widely distributed among eukaryotes[188].

For the expression in *T. thermophila* the promoter region of the metallothionein gene 1 was described[156]. This MTT1 promoter was shown to be able to highly overexpress both homologous and heterologous genes[134]. It was demonstrated that a region of approximately

0.9 kb of this MTT1 promoter is responsible for high transcription rates and can be used to increase the efficiency of most of the commonly used types of DNA-mediated transformation in this organism. During this work, the MTT1 promoter turned out to regulate a strong expression of all recombinant expressed HA antigens, without loss of viability of the cell cultures. Obtained expression yields during a fermentation process were 102±25 mg/L for the NC antigen, 185±25 mg/L for the NC531 antigen, 186±49 mg/L for the UR antigen, and 21±8 mg/L for the FL antigen. Although the MTT1 is a strong promoter, the toxicity of cadmium leads to problems when applying in an industrial production process. Consequently, several other inducible as well as constitutive promoters were tested during this work for the expression of the NC531 HA to replace the MTT1 promoter. Recently, a 1.5 kb fragment from the copper-inducible MTT2 promoter region of *T. thermophila* has been shown to drive a reporter gene expression to levels roughly equivalent to those obtained with MTT1 in this system[157]. This effect could not be demonstrated for the expression of recombinant HA (data not shown). Besides, a 1.3 kb fragment of the copper inducible MTT4 promoter region was reported to act as a strong promoter in homologous protein expression[155]. Applying this promoter for heterologous expression of HA, no functionality could be demonstrated, neither by cadmium nor by copper induction (data not shown). A 1.2 kb part of the upstream region of the MTT5 gene of *T. thermophila* was also described to be useful in homologous protein expression[155]. During this work this zinc inducible MTT5 promoter was shown to be efficient in NC531 HA expression. For the first time the MTT5 promoter was shown to be highly productive in heterologous protein expression comprising an expression yield of NC531 HA of 162±9 mg/L. Likewise, this MTT5 promoter might be an excellent tool for the overexpression of viral antigens for biotechnological application.

Also constitutive promoters are widely distributed for the expression of proteins. The *GAP* promoter, for instance, is derived from the *P. pastoris* glyceraldehydes-3-phosphate dehydrogenase (GAP) gene[189]. The advantage of using this promoter is that neither methanol is required for induction, nor it is necessary to shift cultures from one carbon source to another, making strain growth and protein expression more convenient and straightforward. The disadvantage of a constitutive promoter is that it is undesirable for the production of proteins that may be toxic to the host cell[190]. This could also be true for the expression of recombinant HA by *T. thermophila*. Because of the high expression yield, the recombinant protein is maybe toxic for the host cells, and therefore an inducible promoter like the MTT5 or the MTT1 turned out to be more efficient for the expression of HA. The use of a constitutive promoter would be necessary when applying a high cell density continuous

fermentation procedure, as it was described before[131]. Different constitutive promoters like the phospholipase A1 promoter were tested during this work, but none turned out to initiate a comparable strong expression of HA antigen[191].

Technical tools that are important to establish a sophisticated system have been developed and improved for the *T. thermophila* expression system. This includes techniques for the ciliate transformation into vegetative and conjugating cells and set up of a promoter and secretion test system. Additionally, a variety of molecular biology tools and methods like new promoters, use of a cre dependant homologous recombination system, different episomal and integrative expression vectors, different markers for the selection of clones, including an auxotrophic marker enzyme, and the use of proprietary signal and pro peptides were established[136,142,144,147]. This allows easy and fast testing of different protein variants and putative vaccine candidates. All constructed expression plasmids enabled the transformation of *T. thermophila* cells. After a selection of paromomycin resistant clones the proper uptake of the episomal expression plasmids could be confirmed by testing of whole cell DNA of transformed *Tetrahymena* cells. All target genes comprised a signal peptide to target the antigens to the cell membrane or even to reach secretion. *T. thermophila* is able to secrete a wide range of acid hydrolases into the extracellular medium. The largest clusters of expanded proteases in *T. thermophila* are all cysteine proteases, which comprise 44 % of the total protease complement[108]. Evidence has accumulated for the presence of multiple forms of cysteine proteases in cellular extracts and secretions from *Tetrahymena*[192]. The most conclusive evidence for a predominance of cysteine proteinases came from the observation that E-64, an active site specific and irreversible inhibitor of cysteine proteinases, also strongly inhibited azocasein digestion by extracts[192]. For the protection of the recombinant expressed antigens protease inhibitors were added to prevent the digestion of the membrane bound or secreted antigens. For all antigens the addition of the cysteine protease inhibitor E-64 (2.5 µg/mL) turned out to be most effective for the inhibition of extracellular proteases of *T. thermophila*. Furthermore, E-64 with low cell permeability may also stabilize the intracellular antigen expression. Performing a protease assay, no extracellular protease activity was measured when adding E-64, but growth of the cells and protein production was not influenced (data not shown). The fact that protease inhibitors like leupeptin, antipain, a mix of protease inhibitors (complete™, Roche) are less effective as E-64 was confirmed within this work, as shown for *T. thermophila* before[192]. Similar results were found for the insect cell/baculovirus expression system. It has been demonstrated that leupeptin and E-64 can be used to effectively inhibit the *in vitro* proteolysis of a human recombinant enzyme[193].

Analysis of the cell extracts revealed signals by Western blot analysis due to NC and UR HA expression. For the HA of the FL expression strain a signal was detected analyzing the supernatants. The respective detected bands were appropriate to the calculated molecular weight, except of the FL HA. A signal with an apparent molecular weight of approximately 90 kDa was detected; this was not in agreement with its theoretical molecular weight value of 57 kDa. Differences between apparent and real molecular weights are not uncommon due to differences in folding properties and differences in protein surface charges or the glycosylation pattern (personal communication by Dr. Ingo Aldag, Cilian AG). The fact that the secreted FL antigen could be detected within the cell free supernatant when adding the protease inhibitor E-64, suggests evidence that the other HAs that are anchored by the transmembrane domain (TMD) did not go the pathway of secretion after the passage through the ER and Golgi apparatus. For instance, yeast organisms are favoured for producing heterologous proteins by recombinant DNA technology due to their ability to secrete expressed proteins from the cell. Secretion of the produced protein has several advantages. First, secretion avoids toxicity, which is often a problem when recombinant proteins accumulate inside the cell. Second, many proteins require passage through the secretory pathway in order to obtain proper conformational folding, post translational modifications, and full biological activity. Third, biochemical purification of the recombinant protein is facilitated by secretion since higher initial degrees of purity are obtained in the culture broth than in a cell lysate, and fewer purification steps are needed. The secreted and soluble proteins can be directly recovered by clarification of culture media by centrifugation as it was shown for the production of proteins by *Pichia pastoris*. Samples can be concentrated and purified by subjecting the supernatant to ultrafiltration, precipitation, and/or adsorption/elution chromatography[194]. Because secretion is an important point in the choice of a production platform, it was tested if the HA of the New Caledonia strain is anchored by its TMD into the membrane or if it is intracellularly accumulated applying the *T. thermophila* expression system. Therefore, a C-terminally truncated version of the NC derived HA was cloned, where the TMD and the fusion peptide were excluded. It was possible to exclude the TMD, because the HA1 chain contains receptor binding and major antigenetic sites which should introduce immunogenicity when testing *in vivo*[195]. For the expression of a glycoprotein in *Saccharomyces cerevisiae* the deletion of the transmembrane domain together with the deletion of the autologous signal sequence led to an efficient secretion of the protein[196]. Nevertheless, truncation of the TMD did not lead to the efficient secretion of the C-terminally truncated NC derived HA (NC531) applying the *T. thermophila* expression system but led to a

1.6 x enhanced intracellular protein yield of approximately 180 mg/L. Furthermore, it was tested if the secretion of the NC derived HA can be achieved by the additional exclusion of the hydrophobic HA2 chain. The HA2 chain is primarily responsible for fusion and due to its hydrophobicity can hinder an efficient secretion[195]. Therefore, the open reading frame of the NC HA was reduced to the first 344 amino acids. Cells expressing the NC344 chain were tested for expression and secretion. Intracellular HA of 22±7 mg/L was yielded but no extracellular located protein could be measured reproducibly. As shown for the FL derived HA, it seems to be a clone or HA strain specific phenomenon that enables secretion. Therefore, a clone screening has to be evaluated for the production platform of *T. thermophila*. Cell line screening is an important parameter in the generation of the final Chinese hamster ovary (CHO) hosts for the production of monoclonal antibodies, for instance. However, it is a time-consuming and labor-intensive process, with hundreds of clones screened over the course of many weeks[197]. Furthermore, to screen large number of clones, a high-throughput, flexible system is required. For the expression of recombinant HA, the evaluation of a clone screening remains difficult, because at the moment no suitable test is available and has to be established if necessary. Moreover, there is a high time pressure in producing influenza vaccines which argues against a time-consuming clone screening approach.

A further approach to reach secretion reflects the application of an optimal ER-leader sequence. The mechanism of protein secretion is conserved among all eukaryotes and depends on the recognition of this N-terminal signal sequence and the subsequent co-translational insertion of the nascent polypeptide into the lumen of the ER[198]. Afterwards, the protein is transported through the Golgi apparatus into secretory vesicles. In *T. thermophila*, vesicles may fuse with incoming phagosomes to form phagolysosomes, which are intracellularly processed, and whose residual contents are finally egested. Alternatively, vesicles may fuse with the plasma membrane at so far unknown sites, releasing soluble enzymes into the culture medium. By this pathway, secretory but also membrane-bound forms of lysosomal enzymes may be incorporated into the plasma membrane[76]. For all HA expression constructs the respective native ER leader sequence was applied during this work. The usage of the original signal peptide for recombinant protein production in *T. thermophila* was shown before and resulted in an efficient protein secretion[113,135]. For all expressed HA antigens no secretion was shown using this native signal peptide, except surprisingly for the Florida derived HA antigen containing its TMD. For many secreted proteins, cleavage of the signal peptide results in production of the mature form of the protein. However, there are other secreted proteins,

including, e.g., human insulin-like growth factor-I, that require, when they are intended to be secreted by yeast cells, a sequence between the carboxy-terminus of the signal peptide and the amino-terminus of the mature protein that is commonly known as "pro" sequence. In the absence of a pro sequence some proteins are retained as intracellular, inactive precursors[199]. In order to test, if the secretion of recombinant HA by *T. thermophila* can be achieved using a *T. thermophila* based pro sequence, the pro sequence of the phospholipase A1 gene that led to the secretion of a functional human enzyme, the DNAse I, was tested[113,135]. The phospholipase A1 gene comprises a pre and a pro sequence. For the expression of the recombinant HA, it did not lead to secretion. The choice of the right ER leader sequence remains a complex field, where many sequences have to be tested for every heterologous expressed protein. Recently, it was shown that special combinations of a N-terminal region of the mature *P. pastoris* acid phosphatase gene, in combination with signal and cleavable spacer sequences, was most effective for directing high expression and secretion levels[200]. Therefore, it could be obtained an efficient secretion for antigens by *T. thermophila*, when determining the appropriate signal sequence.

Because of the high intracellular expression values of HA, purification of intracellular located recombinant protein was aimed to be established. To facilitate the purification process, a strain was developed, that is not able to excrete mucocyst contents into the surrounding medium because of a stable integration of a Grl3 gene knock out cassette into the macronuclear genome of *T. thermophila*. Mucocysts or dense core granules (DCGs) are vesicles specialized for storage of highly concentrated proteins, which can be secreted via exocytotic fusion with the plasma membrane in response to extracellular stimuli[201]. DCGs cargo in *T. thermophila* consists principally of a family of granule lattice (Grl) proteins. Upon exocytosis, the lattice remains intact, but undergoes expansion to propel the DCG cargo from the cell[82]. The disruption or silencing of any of the core protein showed, that these proteins act as an essential structural element of the formation of the granule cores[202]. The Grl3 gene was chosen to be best suited for the disruption during this work because of its non redundant role in core formation. Additionally, the disruption of Grl3 just partially inhibited Grl1 protein processing and more important, the transport from ER to the Golgi apparatus of other proteins was not disturbed[92]. By contrast, the disruption of other mucocyst proteins encoding genes like the Grl4 gene leads to the failure of the Grl1 protein to reach the post-trans Golgi network processing compartment due to its retention in the ER. For the Grl6 knock out the mucus after dibucaine treatment was reduced to 40-70 %, whereas for the Grl3 knock out, no mucus could be detected[92]. Therefore, a vector was constructed that combined the knock out of the Grl3

gene with a knock in of the HA expression cassette. This integrative vector had the advantage that after the process of allelic assortment it is stably integrated into the genome of the host cell and no addition of an antibiotic is necessary in terms of large-scale production. This reduces the costs of a production process of the recombinant antigen. Furthermore, the addition of an antibiotic is not desirable for the production of a vaccine. Nevertheless, the process of the allelic assortment needs additional time because all 45 Grl3 alleles have to be replaced by the disrupted copy. After the cells were transformed with the Grl3 knock out vector, *T. thermophila* cells underwent the process of allelic assortment for at least 200-300 fissions. Following analysis by RNA isolation and a PCR analysis of the cDNA, confirmed a proper knock out of the Grl3 gene. Resultant, these strains were not able to excrete mucocyst contents upon stimulation by dibucaine. Secretion of mature mucocysts in *Tetrahymena* can be induced by the local anesthetic dibucaine[167,168]. The dibucaine treatment should stimulate the biosynthetic processes related to mucocyst biogenesis. Although the expression of the NC531 was reduced to 24±10 mg/L this strain can ease the production and purification process, because additional washing steps are not necessary to get rid of the mucus produced by wild type *T. thermophila* cells that could hinder the passage of the proteins through a liquid column chromatography. Furthermore, mucus is able to adsorb target protein which gets lost in the subsequent washing steps (personal communication of Prof. Dr. A. Tiedtke). For large scale production, integration of the recombinant DNA into the chromosome is preferred in order to avoid DNA loss during growth of the cultures. By contrast, problems of genetic instability with episomal templates occur because selective pressure cannot be maintained in large volume cultures due to high costs of the added antibiotic. This results in a reduction in episome copy number due to the selective growth advantage of plasmid-free cells. Generally, a disadvantage associated with integrative systems is low gene dosage, which in turn leads to poor expression[200]. In the knock out strain the gene for the expression of the recombinant HA was present at 45 copies, whereas the episomal expression plasmid in former expression strains was present with about 9,000 copies[109,110]. Gene copy number is the first critical parameter of an expression system, and is therefore also a valuable indicator for potential bottlenecks downstream. This effect may explain the reduction of the expression yield to about 13 %. Numerous examples have been published indicating that an increase of gene copy number can significantly increase productivity. Examples for *P. pastoris* have been described, where a linear correlation between copy number and expression up to a definite upper limit was found[203]. Nevertheless, performing a clone screening approach better transformants comprising higher expression yields could be detected.

5.2 Fermentation process

Besides, another argument to use a ciliate based system is the improved fermentation technology. Batch, fed-batch as well as continuous fermentation techniques are available. Cell densities in continuous cultures of *T. thermophila* can reach up to 2 x 10^7 cells/mL, leading to dry mass weights of 60 to 70 g per Liter[131]. The up-scalability has already been shown by demonstrating the cultivation of *T. thermophila* in an 1.5 m^3 airlift bioreactor[133]. Like yeast cells, *T. thermophila* has a short generation time (1.5 to 3 hours) and is capable of growing in simple and cheap media. A further advantage is the possibility to use standard bioreactor infrastructure designed for micro organisms (yeast or *E. coli*) to cultivate *T. thermophila*[76,131,132,204]. Up-scalability is one of the most important problems in producing vaccines, because millions of vials are necessary to start and continue vaccination programs, as for example it has already been done for measles or the hepatitis B virus.

Producing an influenza vaccine, all HA expressing *T. thermophila* strains were cultivated using a fermentation process. To obtain high yields in a scalable and cost-effective system, the choice of the production process is of great importance. Fed-batch technology has become a platform technology for the large-scale production of therapeutic and recombinant proteins due to its simplicity of operation, flexibility to be implemented and a high final product concentration. Fed-batch culture is a frequently used technology for the industrial manufacturing of cell-culture based subunit vaccines. It was demonstrated before that the productivity of the cell culture process was increased dramatically by optimizing the fed-batch technology and particularly improving the formulation and the feeding strategy. For instance, an antibody production was 40-fold improved using a fed-batch process[166]. An important consequence of this increase in productivity is the decrease in product manufacturing cost. The four different HA expressing *T. thermophila* strains of NC HA, NC531 HA, UR HA and FL HA were tested applying a 0.5 L fermentor. Different conditions for the cultivation were tested using a Sixfors® multifermentor. This fermentor comprises the possibility to test up to six different culture conditions in parallel. The best fermentation medium was found out to be SPP medium. Using skimmed milk medium or wheat peptone medium turned out to be less effective for the HA production process (data not shown). Firstly, the fermentation was performed as a batch process, and induction of the HA expression by cadmium in a final concentration of 20 µg/mL was done. This procedure neither leads to high expression yields, nor to high cell densities. Therefore, the fermentation process was changed to a fed-batch fermentation process. The best suited feeding solution was concentrated SPP medium, which was added to the fermentation broth when cells

reached the end logarithmic phase. The final concentration of this feeding solution was tenfold. This feeding step together with an induction of the HA expression by cadmium in a final concentration of 40 µg/mL turned out to be ideal for obtaining high expression yields. Thus, an intracellular expression yield for recombinant NC HA was estimated to be 102±25 mg/L, for NC531 HA 185±25 mg/L, and for UR HA 186±48 mg/L. These high expression yields depict an advantage of the *T. thermophila* expression system. Indeed, some mammalian cell-based influenza vaccines have obtained license for commercialization, but low production yields are often the major limitations of using this technology[205]. Insect cell culture-based technology is an alternative for manufacturing candidate influenza vaccines. It takes advantages of the progress in recombinant DNA technology and the safety profile of insect cell cultures to successfully produce sub-unit vaccines such as Virus Like Particle (VLP) or recombinant viral proteins[56]. When applying the insect cell/baculovirus expression system expressing the NC derived HA an expression yield of 47±2 mg/L in an optimized fed-batch culture was determined[205]. This means, that using the *T. thermophila* expression platform a 4-fold increased expression with the New Caledonia truncated version of the HA and a 2-fold increased expression by applying the untruncated NC antigen was reached in comparison to the insect cell/baculovirus expression system. Consequently, production costs are lowered by the *T. thermophila* expression system. The major advantage of the insect cell/baculovirus expression system derived sub-unit vaccine is that the manufacturing of the HA protein does not require the cultivation and subsequent inactivation of functional influenza viruses as required for embryonated eggs or mammalian cell production systems[12]. This is also an advantage of the *T. thermophila* expression system. During this work high expression levels were demonstrated and safety issues are low when using *Tetrahymena* as a production platform. Furthermore, it was possible to demonstrate secretion for the recombinant expressed HA of the Florida strain, what is not common for proteins comprising a TMD. The estimation of the yield led to 21±8 mg/L of recombinant FL HA. When applying the recombinant insect cell/baculovirus expression system, the recombinant HA is expressed and transported to the cell membrane. To extract the HA from the cell membranes, a non-ionic detergent is necessary, what could result in problems in terms of folding properties of the recombinant protein[11]. This step is not needed when purification of a secreted HA antigen would be performed.

In conclusion, a simple, reliable and robust fed-batch process was successfully developed for HA production by *T. thermophila* cells. An important feature of the established fed-batch process is a simple formulation of the single feeding solution. It is added to the culture in only

one pulse to sustain high cell viability and recombinant HA productivity. The fed-batch process was extended to produce four different HA proteins. One protein, the Florida derived HA was successfully secreted into the surrounding medium what could facilitate a purification process.

5.3 Down-stream processing

High-cell-density processes are one of the most effective ways of increasing cell efficiency as well as product yield. To obtain sufficient material to establish a purification process the fermentation was scaled up to 50 Liter. As production strain the pAX_MTT1_NC531 turned out to be best suited because of the high yield, and because the excluded TMD may facilitate the extraction of the recombinant antigen from the cells. During this work, the 50 L fermentation process was done by setting optimal parameters defined during the 0.5 L fermentation. The cells reached 0.95×10^6 cells/mL what is less as achieved during the small-scale fermentation procedure. This cell density may be increased by changing the feeding solution, what was demonstrated for instance for the insect cell/baculovirus expression system. An optimized feeding solution, a different feeding strategy as well as an optimized cell density at infection maximized the final recombinant HA production yield between 2- and 3-fold, meaning an expression yield of 47±2 mg/L applying the insect cell/baculovirus expression system[205]. This optimization process has still to be performed for the HA expression strain of NC531 in large-scale. Nevertheless, a high expression yield (150 mg/L) was determined using the *T. thermophila* expression platform and sufficient material was obtained from the fermentation process to establish a purification strategy with high consistency of the starting material for every purification batch.

A purification process has been developed to quickly purify the recombinant HA from the bulk harvested of the bioreactor while retaining its biological activity. In a few instances, a single liquid chromatography step accompanied by other non chromatographic separation techniques may be enough to accomplish protein purifications with a purity ≥95 %, but in most circumstances, a combination of several liquid chromatography steps has to be used[206]. A crucial point for the purification of proteins is the careful retention of their structure during the purification process. For example, the pH had to be fixed to neutral values during the whole purification process to prevent its characteristical conformational change at low pH values which may interfere with subsequent *in vivo* testing of the recombinant HA (personal communication by the cooperation partner).

The main advantage of ion exchange column chromatography (IEC) in purifying proteins, which was used as first purification step for the *T. thermophila* expressed NC531 HA, is that most of the proteins remain bioactive, indicating its usefulness for recombinant protein purification[206]. This first purification step led to an entire concentration of the HA by eliminating contaminating proteins. As material Capto™ Q was defined to be best suited for the first capturing step of the NC531 antigen. Capto™ Q is a strong ion exchange media especially for capture and intermediate purification of proteins. It provides fast, reproducible and easy separations in a convenient format. Furthermore, Capto™ Q leads to a moderate enrichment of target protein. Applying this column chromatography step whole protein contents decreased about 92 %. Nearly no NC531 antigen was lost during this chromatography step. When testing a hydrophobic interaction chromatography material, phenyl sepharose, no purification was possible. Using this column, the pH had to be dropped to acidic pH values, what was not possible for the purification of HA. Applying the cell extracts at neutral pH values, the NC531 antigen did not bind. Testing a size exclusion chromatography, no enrichment of the NC531 antigen was feasible (data not shown). Additionally, size exclusion would not be suited for large industrial-scale processes, because of volume limitations. The second purification step using Ceramic Hydroxyapatite (CHT) as column material was done to capture contaminating protein. The use of a CHT column allows the adaption to industrial-scale columns. The NC531 antigen did not bind to this column material, but was quantitatively detected in the flow through. This depicts another big advantage of this column, because no addition of salt for an elution is necessary and the sample can be loaded directly without performing a buffer exchange to the next column. The whole protein content could be decreased about 45 % applying the CHT chromatography step, and the NC531 antigen decreased just to 80 %. As a last column a Concanavalin A column was chosen. Con A represents the separation principle of proteins by affinity chromatography what is based on the interactions between a protein and the affinity chromatography stationary phase, which has a high selectivity or specifity. It is only one of the non-universal types of liquid chromatography. The selective interaction may occur between protein and low molecular mass substance or between several biopolymers[206]. The disadvantages are high cost and the risk of fouling and irreversible chemical denaturation and notable proteolytic degradation. Con A requires for binding to a glycoprotein at least two non-substituted α-mannosyl residues or extended residues with free hydroxyl groups at C3, 4 and 6 position like in biantennary glycosylated structures. The influenza HA possesses some of such structure[207]. *T. thermophila* is able to add oligo mannose structures to the HA that possesses

3-9 N-linked glycosylation sites[208]. Therefore, Con A binds the recombinant HA and the most significant enrichment of the NC531 antigen was reached. Other *T. thermophila* based proteins did not bind to this column, because maybe the bulk of glycosylated proteins was present in the extracellular medium. Furthermore, the NC531 antigen is highly expressed intracellularly and other contaminating protein is removed during the two purificational steps before. By performing the purification procedure of the NC531 HA a purity of 95 % was reached. Only some minor bands in a coomassie stain were not related to the recombinant NC531 antigen. The HA purity was determined by the combination of Bradford assay (whole protein content) and Western blot analysis. After the concentration of the purified NC531 antigen and a buffer exchange a 71 % pure protein was obtained, and an overall yield of 20 % was reached, decreasing whole protein content to 0.03 %. The reduced purity after the concentration step was maybe due to a binding of HA to the membrane of the spin filter and a consequent loss of the recombinant antigen. Even putative dimers and trimers of the purified NC531 antigen were detected in the coomassie stain, what may enhance the immunogenicity of the antigen when testing *in vivo*. Additionally, a simple test regarding the glycosylation of the NC531 was also shown during this work. The HA monomer consists of a globular head and a fibrous stalk domain, both being N-glycosylated. Typically, 3-9 N-linked glycans are attached to the intact HA protein backbone[208]. Further, it is known that N-linked oligosaccharides attached to the stalk region are highly conserved, whereas those at the tip of the molecule vary considerably in structure and number among different influenza viruses[209]. Applying a PNGase F treatment to the *Tetrahymena* expressed NC531 HA a reduction in molecular weight was detected. Therefore, it seems that the recombinant antigen becomes glycosylated while shuttling through the ER and Golgi apparatus. Glycosylation is of interest since it is involved in protein folding, the stability of the protein in the bloodstream, and protein-protein interactions, as well as contributing to the activity of some proteins[210]. Subtle differences in glycan structure can cause an immune response, as evidenced by the generation of antibodies towards that protein in human sera[211].

In comparison it was also reported a rapid process for production of the NC derived HA with the insect cell/baculovirus system comprising a 95 % pure NC antigen. Three columns were used for the purification of HA derived by the insect cell/baculovirus system, reaching an overall yield of 53 %. The first step was an ion exchange chromatography combined with a Sepharose column material in tandem, combining two chromatographic steps into one. The second was a CHT column followed by a buffer exchange and concentration step via ultrafiltration[11]. The starting material was lower than in the here presented work. An amount

of 12 mg of recombinant HA was loaded to the first column. For the extraction and loading to the first column a pH of 5.89 was needed[11]. This argues for problems regarding the folding properties of HA.

A recombinant protein-based vaccine, such as the *T. thermophila* based HA antigen or the insect cell/baculovirus HA, has some advantages over the traditional egg-based vaccines. It consists solely of three antigens (proteins) stored in sterile phosphate buffered-saline and without preservatives such as thimerosal (a mercury derivative currently used in the egg-based vaccine), antibiotics or adjuvants. Unlike the egg-based vaccines, no functional influenza viruses, biocontainment facilities or harsh chemicals such as formaldehyde are used in manufacturing. Therefore, a reliable, effective, and affordable recombinant protein-based influenza vaccine can be and should be developed to meet the challenge of a potential influenza pandemic.

5.4 Development of a production process for vaccines by *T. thermophila*

Current influenza vaccines are administered to targeted "high risk" groups in order to protect against subtypes of influenza that circulate in humans, namely influenza A subtypes H1N1 and H3N2 and influenza B. Because of the high mutation rate of influenza virus resulting in antigenic drift, each of the three components in the trivalent influenza vaccine is updated regularly to match the vaccine antigens closely to the predominant or emerging strains. For each influenza A subtype, a seed reassortant virus is generated in the laboratory and a "backbone" virus with suitable growth properties for high growth in eggs has to be found. The subsequent selection of the desired reassortant virus can be time consuming and unpredictable and for influenza B virus no suitable high growth strain has been identified[212]. It has long been known that propagation of human influenza isolates in eggs often leads to the selection of mutations, some of which give rise to mismatch in antigenicity between the egg-grown virus and the naturally circulating virus[174,213]. For influenza B viruses, this often manifests as changes in the glycosylation pattern of the HA which may affect the antigenicity[174,214]. Moreover, for some viruses, propagation in eggs is altogether impossible and in this case, reliance on eggs for vaccine manufacturer may result in the exclusion of the most suitable strain from incorporation into the vaccine or a delay in vaccine manufacture. Several cell lines are currently being explored as alternative substrates for influenza vaccine production because of these reasons. These include the MDCK, Vero and PER.C6 cell lines[212]. The application of cell culture systems would add substantial flexibility to the influenza vaccine production

process, particularly important in times requiring rapid scale-up, such as onset of a pandemic. Furthermore, the use of eukaryotic cell lines for propagation of vaccine strains reduces likelihood of selecting variants with altered antigenicity since HA mutations are not necessarily selected during replication in cell culture. Moreover, the recombinant protein expressed by cell lines like *T. thermophila* or insect cells are highly purified and do not contain egg protein, which should eliminate hypersensitivity reactions in egg-sensitive individuals. While there is an already existing recombinant cell culture based expression system, the insect cell/baculovirus system, there are some disadvantages by this platform. In clinical trials there were poor response rates to the insect cell/baculovirus derived vaccine observed in the 6 to 35-month-old age group and expression yields were low. In addition, high license fees have to be paid when using this system.

The main advantages when using the here developed *T. thermophila* expression system are that it is cost effective. Further advantages of the system include the strong inducible MTT promoters, which have transcription characteristics useful for regulating heterologous protein expression, the well-developed methods for classical and molecular-genetic manipulation of the organism and the technology for the growth of an expression strain in large high-densitiy fermentor cultures. Furthermore, secretion of hemagglutinin was shown for the first time using a heterologous expression system. This feature can ease the purification process, and therefore lowers the costs of the production process. In addition, the glycosylation added by *T. thermophila* generates more homogenous protein what could be an advantage for the immunogenicity in younger human. The high expression yields shown in this work for three different HAs, one of the B strain and two of A strain influenza viruses, comprise a big advantage over all other expression systems currently used for the production of influenza vaccines.

5.5 Future prospects

First *in vivo* testings in mice were performed by the cooperation partner to examine the immunogenicity of the NC531 antigen produced by *Tetrahymena thermophila*. Preliminary results indicated slight immunogenicity when testing the NC531 antigen. Hence, *T. thermophila* offers a big potential for producing foreign eukaryotic proteins. However, further development has to be done to deplete some problems and enable a highly efficient application of *T. thermophila*. Following points may be considered for this purpose:

1. Enhancement of the transformation rate of the host organism.
2. Identification of new strong inducible and constitutive promoters.
3. Optimization of the secretion.
4. Analysis of the glycosylation pattern of the expressed antigen.
5. Establishment of an enhanced fermentation and purification process.

Transformation of *T. thermophila* can be carried out by e.g. electroporation or the biolistic bombardment. Nevertheless, the transformation rate is relatively low[136]. Therefore, the design of smaller sized expression vectors would be necessary as a first point. Moreover, the utilization of stronger constitutive promoters for the selection marker would enhance the transformation rate. Recently, the replacement of the histone (H4-1) promoter within the Neo2 cassette was tested in order to obtain a higher transformation rate. First promising results indicated an enhancement of the transformation rate.

In addition to the availability of optimal techniques, also the strain improvement and a fermentation process optimization including the improvement of the downstream process are very important. Experience shows that choices made early in the process of development often are difficult to change due to regulatory constraints and process modification. This may hamper the profitability and increase costs. Consequently, during this work a strong and zinc inducible promoter was applied to control a high expression level of hemagglutinin and to replace the cadmium inducible MTT1 promoter. In order to obtain a production process where the addition of any heavy metal is necessary, further promoter candidates have to be identified. For instance, heat inducible promoters of the hsp family could be an alternative. Besides, high cell density fermentation procedure applying continuous fermentation was successfully established for the cultivation of *T. thermophila* cells[131]. In order to produce a vaccine applying continuous fermentation process, strong constitutive promoters have to be

identified. Some candidates were proven to be functional during this work, but up to now none of the tested promoters were strong enough to drive high antigen expression levels.

An optimization of the secretion of specific enzymes by *T. thermophila* was obtained due to the development of hypersecretory mutants through random mutagenesis[158,215]. During this work one target antigen was also transported via the secretory pathway and was detected in the extracellular medium comprising high expression levels. To establish an efficient production process for a trivalent vaccine maybe the identification of novel ER-leader sequences that are capable of directing the secretion of the two resting HA antigens is feasible. Additionally, for better understanding of the localization of currently intracellularly expressed HAs immunofluorescent studies could be done. Moreover, the amino acid sequence of the secreted antigen could be analyzed and optimized in terms of the deletion of the transmembrane domain and hydrophobic regions. Besides, a codon harmonization can be done in order to obtain a more optimized folding of the recombinant protein and maybe a facilitated secretion of the antigen.

Furthermore, the glycosylation pattern of the recombinant *Tetrahymena* based antigens could be analyzed. Most human proteins are glycosylated: they have one or more oligosaccharide chains attached to the polypeptide backbone. To analyze the glycosylation pattern, a method was published by Schwarzer *et al*.[216]. This method could be used for the *T. thermophila* expressed hemagglutinin, giving more insights regarding the glycosylation machinery of this ciliate. Besides, the folding properties of the recombinant expressed protein could be investigated.

Additionally, the production process could be improved in order to maximize the capability of the expression strains. Strategies commonly employed to optimize the fermentation process are testing of different compositions of cultivation media and point of time of induction of the expression. First results indicated, that the choice of the medium is of critical importance, especially for the secreted FL antigen. Moreover, for the applied fed-batch fermentation procedure, the composition of the feeding solution has further to be developed. Moreover, it could be considered a continious feeding strategy. By testing different feeding strategies, the insect cell/baculovirus expression system was able to increase the yield 2- to 3-fold[205]. The development of an optimized downstream process plays a critical role in recombinant protein production. Therefore, the purification process should further be advanced by identifying more convenient column materials and elution strategies. First of all, an elution of the antigen

in one column volume is intended to be reached. Besides, the concentration of the sample combined by a buffer exchange should be optimized to minimize the losses during this step.

6. Summary

Within this work, the expression of three different hemagglutinins (HAs) representing the composition of a trivalent influenza vaccine was successfully shown applying the *Tetrahymena thermophila* based expression platform. The three different hemagglutinins, originated from the influenza A virus strain New Caledonia 20/99 (H1N1), from the influenza A virus strain Uruguay/716/2007 (H3N2), and from the influenza B virus strain Florida/4/2006 and were produced using a codon optimized synthetic gene.

Secretion of about 21±8 mg/L HA of the influenza B virus strain Florida/4/2006 was reached, while the HA of the influenza A strain was expressed intracellularly, as expected. High protein yields of about 102±25 mg/L and 186±49 mg/L were attained, expressing the HA of the influenza A virus strain New Caledonia 20/99 and of the HA of the influenza A virus strain Uruguay/716/2007, respectively.

In further analyses, it was demonstrated that the expression of C-terminally truncated HA of the New Caledonia strain, showing no transmembrane domain does not support the export of the recombinant protein but does lead to higher intracellular expression yields of about 185±25 mg/L.

The optimized truncated HA version was finally used to establish an efficient production and purification process. Within the framework of the process, optimization the strong but cadmium-inducible MTT1 promoter could be replaced by the zinc-inducible MTT5 promoter, showing nearly the same expression yield of about 162±9 mg/L.

Furthermore, a *Tetrahymena thermophila* expression strain was designed and developed to ease the purification of intracellular located target proteins due to its inability to produce the undesirable mucocyst proteins.

Concurrently, the fermentation process was optimized and scaled up to 50 L. The truncated New Caledonia derived HA was successfully purified, comprising a purity of 95 % and an overall yield of 20 %. The purified antigen was glycosylated, which is an important feature to introduce neutralizing antibodies during an immunogenicity study.

7. Reference List

[1] World Health Organization. 209.
Ref Type: Internet Communication

[2] Hansson M, Nygren PA, Stahl S. Design and production of recombinant subunit vaccines. Biotechnol Appl Biochem 2000 Oct;32 (Pt 2):95-107.

[3] Clark TG, Cassidy-Hanley D. Recombinant subunit vaccines: potentials and constraints. Dev Biol (Basel) 2005;121:153-63.

[4] Pastoret PP, Vanderplasschen A. Poxviruses as vaccine vectors. Comp Immunol Microbiol Infect Dis 2003 Oct;26(5-6):343-55.

[5] Heppell J, Davis HL. Application of DNA vaccine technology to aquaculture. Adv Drug Deliv Rev 2000 Sep 15;43(1):29-43.

[6] Swartz JR. Advances in Escherichia coli production of therapeutic proteins. Curr Opin Biotechnol 2001 Apr;12(2):195-201.

[7] Griffith DA, Delipala C, Leadsham J, Jarvis SM, Oesterhelt D. A novel yeast expression system for the overproduction of quality-controlled membrane proteins. FEBS Lett 2003 Oct 9;553(1-2):45-50.

[8] Sakamoto S, Ide T, Tokiyoshi S, et al. Studies on the structures and antigenic properties of rabies virus glycoprotein analogues produced in yeast cells. Vaccine 1999 Jan 21;17(3):205-18.

[9] Leong JC, Anderson E, Bootland LM, et al. Fish vaccine antigens produced or delivered by recombinant DNA technologies. Dev Biol Stand 1997;90:267-77.

[10] Lorenzen N, Olesen NJ. Immunization with viral antigens: viral haemorrhagic septicaemia. Dev Biol Stand 1997;90:201-9.

[11] Wang K, Holtz KM, Anderson K, Chubet R, Mahmoud W, Cox MM. Expression and purification of an influenza hemagglutinin--one step closer to a recombinant protein-based influenza vaccine. Vaccine 2006 Mar 15;24(12):2176-85.

[12] Cox MM, Hollister JR. FluBlok, a next generation influenza vaccine manufactured in insect cells. Biologicals 2009 Jun;37(3):182-9.

[13] Cox MM, Patriarca PA, Treanor J. FluBlok, a recombinant hemagglutinin influenza vaccine. Influenza Other Respi Viruses 2008 Nov;2(6):211-9.

[14] Huber VC, McCullers JA. FluBlok, a recombinant influenza vaccine. Curr Opin Mol Ther 2008 Feb;10(1):75-85.

[15] Schillberg S, Twyman RM, Fischer R. Opportunities for recombinant antigen and antibody expression in transgenic plants--technology assessment. Vaccine 2005 Mar 7;23(15):1764-9.

[16] Shoji Y, Farrance CE, Bi H, et al. Immunogenicity of hemagglutinin from A/Bar-headed Goose/Qinghai/1A/05 and A/Anhui/1/05 strains of H5N1 influenza viruses produced in Nicotiana benthamiana plants. Vaccine 2009 May 26;27(25-26):3467-70.

[17] Shoji Y, Bi H, Musiychuk K, et al. Plant-derived hemagglutinin protects ferrets against challenge infection with the A/Indonesia/05/05 strain of avian influenza. Vaccine 2009 Feb 11;27(7):1087-92.

[18] Nerome K, Kumihashi H, Nerome R, et al. Evaluation of immune responses to inactivated influenza vaccines prepared in embryonated chicken eggs and MDCK cells in a mouse model. Dev Biol Stand 1999;98:53-63.

[19] Wood JM, Oxford JS, Dunleavy U, Newman RW, Major D, Robertson JS. Influenza A (H1N1) vaccine efficacy in animal models is influenced by two amino acid substitutions in the hemagglutinin molecule. Virology 1989 Jul;171(1):214-21.

[20] Kistner O, Barrett PN, Mundt W, et al. A novel mammalian cell (Vero) derived influenza virus vaccine: development, characterization and industrial scale production. Wien Klin Wochenschr 1999 Mar 12;111(5):207-14.

[21] Palache AM, Scheepers HS, de R, V, et al. Safety, reactogenicity and immunogenicity of Madin Darby Canine Kidney cell-derived inactivated influenza subunit vaccine. A meta-analysis of clinical studies. Dev Biol Stand 1999;98:115-25.

[22] Percheson PB, Trepanier P, Dugre R, Mabrouk T. A Phase I, randomized controlled clinical trial to study the reactogenicity and immunogenicity of a new split influenza vaccine derived from a non-tumorigenic cell line. Dev Biol Stand 1999;98:127-32.

[23] Palache AM, Brands R, van Scharrenburg GJ. Immunogenicity and reactogenicity of influenza subunit vaccines produced in MDCK cells or fertilized chicken eggs. J Infect Dis 1997 Aug;176 Suppl 1:S20-S23.

[24] Kessler N, Thomas-Roche G, Gerentes L, Aymard M. Suitability of MDCK cells grown in a serum-free medium for influenza virus production. Dev Biol Stand 1999;98:13-21.

[25] Merten OW, Hannoun C, Manuguerra JC, Ventre F, Petres S. Production of influenza virus in cell cultures for vaccine preparation. Adv Exp Med Biol 1996;397:141-51.

[26] Kistner O, Barrett PN, Mundt W, et al. Development of a Vero cell-derived influenza whole virus vaccine. Dev Biol Stand 1999;98:101-10.

[27] Kistner O, Barrett PN, Mundt W, Reiter M, Schober-Bendixen S, Dorner F. Development of a mammalian cell (Vero) derived candidate influenza virus vaccine. Vaccine 1998 May;16(9-10):960-8.

[28] Pau MG, Ophorst C, Koldijk MH, Schouten G, Mehtali M, Uytdehaag F. The human cell line PER.C6 provides a new manufacturing system for the production of influenza vaccines. Vaccine 2001 Mar 21;19(17-19):2716-21.

[29] Govorkova EA, Murti G, Meignier B, de TC, Webster RG. African green monkey kidney (Vero) cells provide an alternative host cell system for influenza A and B viruses. J Virol 1996 Aug;70(8):5519-24.

[30] Schnitzler SU, Schnitzler P. An update on swine-origin influenza virus A/H1N1: a review. Virus Genes 2009 Nov;39(3):279-92.

[31] Michaelis M, Doerr HW, Cinatl J, Jr. Novel swine-origin influenza A virus in humans: another pandemic knocking at the door. Med Microbiol Immunol 2009 Aug;198(3):175-83.

[32] Peiris JS, Tu WW, Yen HL. A novel H1N1 virus causes the first pandemic of the 21st century. Eur J Immunol 2009 Nov;39(11):2946-54.

[33] Krammer F, Nakowitsch S, Messner P, Palmberger D, Ferko B, Grabherr R. Swine-origin pandemic H1N1 influenza virus-like particles produced in insect cells induce hemagglutination inhibiting antibodies in BALB/c mice. Biotechnol J 2010 Jan;5(1):17-23.

[34] Hilleman MR. Realities and enigmas of human viral influenza: pathogenesis, epidemiology and control. Vaccine 2002 Aug 19;20(25-26):3068-87.

[35] Glezen WP. Serious morbidity and mortality associated with influenza epidemics. Epidemiol Rev 1982;4:25-44.

[36] Cox NJ, Subbarao K. Influenza. Lancet 1999 Oct 9;354(9186):1277-82.

[37] Gerdil C. The annual production cycle for influenza vaccine. Vaccine 2003 May 1;21(16):1776-9.

[38] Oxford JS, Corcoran T, Knott R, et al. Serological studies with influenza A(H1N1) viruses cultivated in eggs or in a canine kidney cell line (MDCK). Bull World Health Organ 1987;65(2):181-7.

[39] Oxford JS, Newman R, Corcoran T, et al. Direct isolation in eggs of influenza A (H1N1) and B viruses with haemagglutinins of different antigenic and amino acid composition. J Gen Virol 1991 Jan;72 (Pt 1):185-9.

[40] Robertson JS, Bootman JS, Newman R, et al. Structural changes in the haemagglutinin which accompany egg adaptation of an influenza A(H1N1) virus. Virology 1987 Sep;160(1):31-7.

[41] Nayak DP, Balogun RA, Yamada H, Zhou ZH, Barman S. Influenza virus morphogenesis and budding. Virus Res 2009 Aug;143(2):147-61.

[42] Tatulian SA, Hinterdorfer P, Baber G, Tamm LK. Influenza hemagglutinin assumes a tilted conformation during membrane fusion as determined by attenuated total reflection FTIR spectroscopy. EMBO J 1995 Nov 15;14(22):5514-23.

[43] Stegmann T, Delfino JM, Richards FM, Helenius A. The HA2 subunit of influenza hemagglutinin inserts into the target membrane prior to fusion. J Biol Chem 1991 Sep 25;266(27):18404-10.

[44] Durrer P, Galli C, Hoenke S, et al. H+-induced membrane insertion of influenza virus hemagglutinin involves the HA2 amino-terminal fusion peptide but not the coiled coil region. J Biol Chem 1996 Jun 7;271(23):13417-21.

[45] Wang S, Taaffe J, Parker C, et al. Hemagglutinin (HA) proteins from H1 and H3 serotypes of influenza A viruses require different antigen designs for the induction of optimal protective antibody responses as studied by codon-optimized HA DNA vaccines. J Virol 2006 Dec;80(23):11628-37.

[46] Horimoto T, Kawaoka Y. Influenza: Lessons from past pandemics, warnings from current incidents. Nature Reviews Microbiology 2005;3(8):591-600.

[47] Wood JM. Standardization of inactivated influenza vaccine. London: Blackwell Science 1998;333-45.

[48] D'Aoust MA, Lavoie PO, Couture MM, et al. Influenza virus-like particles produced by transient expression in Nicotiana benthamiana induce a protective immune response against a lethal viral challenge in mice. Plant Biotechnol J 2008 Dec;6(9):930-40.

[49] Kang SM, Song JM, Quan FS, Compans RW. Influenza vaccines based on virus-like particles. Virus Res 2009 Aug;143(2):140-6.

[50] Bright RA, Carter DM, Daniluk S, et al. Influenza virus-like particles elicit broader immune responses than whole virion inactivated influenza virus or recombinant hemagglutinin. Vaccine 2007 May 10;25(19):3871-8.

[51] Mahmood K, Bright RA, Mytle N, et al. H5N1 VLP vaccine induced protection in ferrets against lethal challenge with highly pathogenic H5N1 influenza viruses. Vaccine 2008 Oct 3;26(42):5393-9.

[52] Ross TM, Mahmood K, Crevar CJ, Schneider-Ohrum K, Heaton PM, Bright RA. A trivalent virus-like particle vaccine elicits protective immune responses against seasonal influenza strains in mice and ferrets. PLoS One 2009;4(6):e6032.

[53] Pushko P, Tumpey TM, Bu F, Knell J, Robinson R, Smith G. Influenza virus-like particles comprised of the HA, NA, and M1 proteins of H9N2 influenza virus induce protective immune responses in BALB/c mice. Vaccine 2005 Dec 30;23(50):5751-9.

[54] Quan FS, Huang C, Compans RW, Kang SM. Virus-like particle vaccine induces protective immunity against homologous and heterologous strains of influenza virus. J Virol 2007 Apr;81(7):3514-24.

[55] Galarza JM, Latham T, Cupo A. Virus-like particle vaccine conferred complete protection against a lethal influenza virus challenge. Viral Immunol 2005;18(2):365-72.

[56] Cox MM. Progress on baculovirus-derived influenza vaccines. Curr Opin Mol Ther 2008 Feb;10(1):56-61.

[57] Perrone LA, Ahmad A, Veguilla V, et al. Intranasal vaccination with 1918 influenza virus-like particles protects mice and ferrets from lethal 1918 and H5N1 influenza virus challenge. J Virol 2009 Jun;83(11):5726-34.

[58] Peterson DS, Gao Y, Asokan K, Gaertig J. The circumsporozoite protein of Plasmodium falciparum is expressed and localized to the cell surface in the free-living ciliate Tetrahymena thermophila. Mol Biochem Parasitol 2002 Jul;122(2):119-26.

[59] Gaertig J, Gao Y, Tishgarten T, Clark TG, Dickerson HW. Surface display of a parasite antigen in the ciliate Tetrahymena thermophila. Nat Biotechnol 1999 May;17(5):462-5.

[60] Frankel J. Cell biology of Tetrahymena thermophila. Methods Cell Biol 2000;62:27-125.

[61] Allen RD. Fine structure, reconstruction and possible functions of components of the cortex of Tetrahymena pyriformis. J Protozool 1967 Nov;14(4):553-65.

[62] Satir BH, Wissig SL. Alveolar sacs of Tetrahymena: ultrastructural characteristics and similarities to subsurface cisterns of muscle and nerve. J Cell Sci 1982 Jun;55:13-33.

[63] Aufderheide K. Mitochondrial associations with specific microtubular components of the cortex of Tetrahymena thermophila. I. Cortical patterning of mitochondria. J Cell Sci 1979 Oct;39:299-312.

[64] Kurz S, Tiedtke A. The Golgi apparatus of Tetrahymena thermophila. J Eukaryot Microbiol 1993 Jan;40(1):10-3.

[65] Bell A. http://tet jsd claremont edu/about php 2010

[66] Frankel J. Tetrahymena. Encyclopedia of Life Sciences 2001.

[67] Williams NE. Surface membrane regeneration in deciliated Tetrahymena. J Cell Sci 1983 Jul;62:407-17.

[68] Williams NE, Honts JE, Kaczanowska J. The formation of basal body domains in the membrane skeleton of Tetrahymena. Development 1990 Aug;109(4):935-42.

[69] Leick V, Hellung-Larsen P. Chemosensory behaviour of Tetrahymena. Bioessays 1992 Jan;14(1):61-6.

[70] Collins T, Baker RL, Wilhelm JM, Olmsted JB. A cortical scaffold in the ciliate Tetrahymena. J Ultrastruct Res 1980 Jan;70(1):92-103.

[71] Aufderheide KJ. Mitochondrial associations with specific microtubular components of the cortex of Tetrahymena thermophila. II. Response of the mitochondrial pattern to changes in the microtubule pattern. J Cell Sci 1980 Apr;42:247-60.

[72] Nilsson JR. Intracellular distribution of lead in Tetrahymena during continuous exposure to the metal. J Cell Sci 1979 Oct;39:383-96.

[73] Muller M, Baudhuin P, de DC. Lysosomes in Tetrahymena pyriformis. I. Some properties and lysosomal localization of acid hydrolases. J Cell Physiol 1966 Oct;68(2):165-75.

[74] Muller M. Lysosomes in Tetrahymena pyriformis. II. Intracellular distribution of several acid hydrolases. Acta Biol Acad Sci Hung 1971;22(2):179-86.

[75] Muller M. Secretion of acid hydrolases and its intracellular source in Tetrahymena pyriformis. J Cell Biol 1972 Feb;52(2):478-87.

[76] Kiy T, Vosskuhler C, Rasmussen L, Tiedtke A. Three pools of lysosomal enzymes in Tetrahymena thermophila. Exp Cell Res 1993 Apr;205(2):286-92.

[77] Nilsson JR. Reversible inhibition of RNA synthesis in Tetrahymena pyriforms GL by dimethyl sulphoxide: an electron microscope autoradiographic study. Trans Am Microsc Soc 1976 Jul;95(3):403-14.

[78] Kiy T, Tiedtke A. Effects of immobilization on growth, morphology, and DNA content of the ciliated protozoon Tetrahymena thermophila. FEMS Microbiol Lett 1993 Jan 1;106(1):117-22.

[79] Sattler CA, Staehelin LA. Reconstruction of oral cavity of Tetrahymena pyriformis utilizing high voltage electron microscopy. Tissue Cell 1976;8(1):1-18.

[80] Allen RD, Wolf RW. Membrane recycling at the cytoproct of Tetrahymena. J Cell Sci 1979 Feb;35:217-27.

[81] Franke WW, Eckert WA, Krien S. Cytomembrane differentiation in a ciliate, Tetrahymena pyriformis. I. Endoplasmic reticulum and dictyosomal equivalents. Z Zellforsch Mikrosk Anat 1971;119(4):577-604.

[82] Hausmann K. Extrusive organelles in protists. Int Rev Cytol 1978;52:197-276.

[83] Williams NE, Luft JH. Use of a nitrogen mustard derivative in fixation for electron microscopy and observations on the ultrastructure of Tetrahymena. J Ultrastruct Res 1968 Nov;25(3):271-92.

[84] Satir B, Schooley C, Satir P. Membrane fusion in a model system. Mucocyst secretion in Tetrahymena. J Cell Biol 1973 Jan;56(1):153-76.

[85] Wissig SL, Satir BH. Exocytosis of mucocysts induced by Ca++-ionophore in Tetrahymena: Quantitation and effects of divalent cations. J Submicrosc Cytol 1970;12:1-16.

[86] Verbsky JW, Turkewitz AP. Proteolytic processing and Ca2+-binding activity of dense-core vesicle polypeptides in Tetrahymena. Mol Biol Cell 1998 Feb;9(2):497-511.

[87] Tiedtke A. Capsule shedding in tetrahymena. Naturwissenschaften 1976 Feb;63(2):93.

[88] Maihle NJ, Satir BH. Protein secretion in Tetrahymena thermophila. Characterization of the major proteinaceous secretory proteins. J Biol Chem 1986 Jun 5;261(16):7566-70.

[89] Chilcoat ND, Melia SM, Haddad A, Turkewitz AP. Granule lattice protein 1 (Grl1p), an acidic, calcium-binding protein in Tetrahymena thermophila dense-core secretory granules, influences granule size, shape, content organization, and release but not protein sorting or condensation. J Cell Biol 1996 Dec;135(6 Pt 2):1775-87.

[90] Sweeney R, Fan Q, Yao MC. Antisense ribosomes: rRNA as a vehicle for antisense RNAs. Proc Natl Acad Sci U S A 1996 Aug 6;93(16):8518-23.

[91] Chilcoat ND, Elde NC, Turkewitz AP. An antisense approach to phenotype-based gene cloning in Tetrahymena. Proc Natl Acad Sci U S A 2001 Jul 17;98(15):8709-13.

[92] Cowan AT, Bowman GR, Edwards KF, Emerson JJ, Turkewitz AP. Genetic, genomic, and functional analysis of the granule lattice proteins in Tetrahymena secretory granules. Mol Biol Cell 2005 Sep;16(9):4046-60.

[93] Karrer KM. Tetrahymena genetics: two nuclei are better than one. Methods Cell Biol 2000;62:127-86.

[94] RAY C, Jr. Preparation of chromosomes of Tetrahymena pyriformis for photomicrography. Stain Technol 1956 Nov;31(6):271-4.

[95] LaFountain JR, Jr., Davidson LA. An analysis of spindle ultrastructure during prometaphase and metaphase of micronuclear division in Tetrahymena. Chromosoma 1979;75(3):293-308.

[96] LaFountain JR, Jr., Davidson LA. An analysis of spindle ultrastructure during anaphase of micronuclear division in Tetrahymena. Cell Motil 1980;1(1):41-61.

[97] Allen SL. Linkage studies in variety1 of Tetrahymena pyriformis: A first case of linkage in the ciliated protozoa. Genetics 1964 Apr;49:617-27.

[98] Allen SL, Ervin PR, McLaren NC, Brand RE. The 5S ribosomal RNA gene clusters in Tetrahymena thermophila: strain differences, chromosomal localization, and loss during micronuclear ageing. Mol Gen Genet 1984;197(2):244-53.

[99] Merriam EV, Bruns PJ. Phenotypic assortment in Tetrahymena thermophila: assortment kinetics of antibiotic-resistance markers, tsA, death, and the highly amplified rDNA locus. Genetics 1988 Oct;120(2):389-95.

[100] Yao MC, Chao JL. RNA-guided DNA deletion in Tetrahymena: an RNAi-based mechanism for programmed genome rearrangements. Annu Rev Genet 2005;39:537-59.

[101] Yao MC, Choi J, Yokoyama S, Austerberry CF, Yao CH. DNA elimination in Tetrahymena: a developmental process involving extensive breakage and rejoining of DNA at defined sites. Cell 1984 Feb;36(2):433-40.

[102] Fan Q, Yao MC. A long stringent sequence signal for programmed chromosome breakage in Tetrahymena thermophila. Nucleic Acids Res 2000 Feb 15;28(4):895-900.

[103] Hamilton EP, Williamson S, Dunn S, et al. The highly conserved family of Tetrahymena thermophila chromosome breakage elements contains an invariant 10-base-pair core. Eukaryot Cell 2006 Apr;5(4):771-80.

[104] Yao MC, Yao CH, Monks B. The controlling sequence for site-specific chromosome breakage in Tetrahymena. Cell 1990 Nov 16;63(4):763-72.

[105] Fan Q, Yao M. New telomere formation coupled with site-specific chromosome breakage in Tetrahymena thermophila. Mol Cell Biol 1996 Mar;16(3):1267-74.

[106] Yu GL, Blackburn EH. Developmentally programmed healing of chromosomes by telomerase in Tetrahymena. Cell 1991 Nov 15;67(4):823-32.

[107] Coyne RS, Thiagarajan M, Jones KM, et al. Refined annotation and assembly of the Tetrahymena thermophila genome sequence through EST analysis, comparative genomic hybridization, and targeted gap closure. BMC Genomics 2008;9:562.

[108] Eisen JA, Coyne RS, Wu M, et al. Macronuclear genome sequence of the ciliate Tetrahymena thermophila, a model eukaryote. PLoS Biol 2006 Sep;4(9):e286.

[109] Kapler GM. Developmentally regulated processing and replication of the Tetrahymena rDNA minichromosome. Curr Opin Genet Dev 1993 Oct;3(5):730-5.

[110] Doerder FP, Deak JC, Lief JH. Rate of phenotypic assortment in Tetrahymena thermophila. Dev Genet 1992;13(2):126-32.

[111] Allewell NM, Oles J, Wolfe J. A physicochemical analysis of conjugation in Tetrahymena pyriformis. Exp Cell Res 1976 Feb;97(2):394-405.

[112] Kapler GM, Blackburn EH. A weak germ-line excision mutation blocks developmentally controlled amplification of the rDNA minichromosome of Tetrahymena thermophila. Genes Dev 1994 Jan;8(1):84-95.

[113] Bockau U. Tetrahymena thermophila: Ein neues Expressionssystem zur Produktion humaner Proteine. Westfälische Wilhelms-Universität Münster. Institut für allgemeine Zoologie und Genetik.; 2006.

[114] Christie KR, Weng S, Balakrishnan R, et al. Saccharomyces Genome Database (SGD) provides tools to identify and analyze sequences from Saccharomyces cerevisiae and related sequences from other organisms. Nucleic Acids Res 2004 Jan 1;32(Database issue):D311-D314.

[115] Stover NA, Krieger CJ, Binkley G, et al. Tetrahymena Genome Database (TGD): a new genomic resource for Tetrahymena thermophila research. Nucleic Acids Res 2006 Jan 1;34(Database issue):D500-D503.

[116] Tourancheau AB, Tsao N, Klobutcher LA, Pearlman RE, Adoutte A. Genetic code deviations in the ciliates: evidence for multiple and independent events. EMBO J 1995 Jul 3;14(13):3262-7.

[117] Horowitz S, Gorovsky MA. An unusual genetic code in nuclear genes of Tetrahymena. Proc Natl Acad Sci U S A 1985 Apr;82(8):2452-5.

[118] Shrimali RK, Lobanov AV, Xu XM, et al. Selenocysteine tRNA identification in the model organisms Dictyostelium discoideum and Tetrahymena thermophila. Biochem Biophys Res Commun 2005 Apr 1;329(1):147-51.

[119] Wuitschick JD, Karrer KM. Analysis of genomic G + C content, codon usage, initiator codon context and translation termination sites in Tetrahymena thermophila. J Eukaryot Microbiol 1999 May;46(3):239-47.

[120] Eichinger L, Pachebat JA, Glockner G, et al. The genome of the social amoeba Dictyostelium discoideum. Nature 2005 May 5;435(7038):43-57.

[121] Salim HM, Ring KL, Cavalcanti AR. Patterns of codon usage in two ciliates that reassign the genetic code: Tetrahymena thermophila and Paramecium tetraurelia. Protist 2008 Apr;159(2):283-98.

[122] Jaillon O, Bouhouche K, Gout JF, et al. Translational control of intron splicing in eukaryotes. Nature 2008 Jan 17;451(7176):359-62.

[123] Collins K, Gorovsky MA. Tetrahymena thermophila. Curr Biol 2005 May 10;15(9):R317-R318.

[124] Turkewitz AP, Orias E, Kapler G. Functional genomics: the coming of age for Tetrahymena thermophila. Trends Genet 2002 Jan;18(1):35-40.

[125] Cech TR. Ribozymes, the first 20 years. Biochem Soc Trans 2002 Nov;30(Pt 6):1162-6.

[126] Asai DJ, DeWall KM, Lincoln LM, Smith RK. Manipulating dynein genes in Tetrahymena. Methods Mol Biol 2001;161:269-78.

[127] Asai DJ, Wilkes DE. The dynein heavy chain family. J Eukaryot Microbiol 2004 Jan;51(1):23-9.

[128] Brownell JE, Zhou J, Ranalli T, et al. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. Cell 1996 Mar 22;84(6):843-51.

[129] Yao MC, Fuller P, Xi X. Programmed DNA deletion as an RNA-guided system of genome defense. Science 2003 Jun 6;300(5625):1581-4.

[130] Cassidy-Hanley D, Smith HR, Bruns PJ. A simple, efficient technique for freezing Tetrahymena thermophila. J Eukaryot Microbiol 1995 Sep;42(5):510-5.

[131] Kiy T, Tiedtke A. Continuous high-cell-density fermentation of the ciliated protozoon Tetrahymena in a perfused bioreactor. Appl Microbiol Biotechnol 1992 Nov;38(2):141-6.

[132] Wheatley DN, Rasmussen L, Tiedtke A. Tetrahymena: a model for growth, cell cycle and nutritional studies, with biotechnological potential. Bioessays 1994 May;16(5):367-72.

[133] Hellenbroich D, Valley U, Ryll T, et al. Cultivation of Tetrahymena thermophila in a 1.5-m3 airlift bioreactor. Appl Microbiol Biotechnol 1999 Apr;51(4):447-55.

[134] Shang Y, Song X, Bowen J, et al. A robust inducible-repressible promoter greatly facilitates gene knockouts, conditional expression, and overexpression of homologous and heterologous genes in Tetrahymena thermophila. Proc Natl Acad Sci U S A 2002 Mar 19;99(6):3734-9.

[135] Weide T, Herrmann L, Bockau U, et al. Secretion of functional human enzymes by Tetrahymena thermophila. BMC Biotechnol 2006;6:19.

[136] Cassidy-Hanley D, Bowen J, Lee JH, et al. Germline and somatic transformation of mating Tetrahymena thermophila by particle bombardment. Genetics 1997 May;146(1):135-47.

[137] Tondravi MM, Yao MC. Transformation of Tetrahymena thermophila by microinjection of ribosomal RNA genes. Proc Natl Acad Sci U S A 1986 Jun;83(12):4369-73.

[138] Brunk CF, Navas P. Transformation of Tetrahymena thermophila by electroporation and parameters effecting cell survival. Exp Cell Res 1988 Feb;174(2):525-32.

[139] Orias E, Larson D, Hu YF, et al. Replacement of the macronuclear ribosomal RNA genes of a mutant Tetrahymena using electroporation. Gene 1988 Oct 30;70(2):295-301.

[140] Gaertig J, Gorovsky MA. Efficient mass transformation of Tetrahymena thermophila by electroporation of conjugants. Proc Natl Acad Sci U S A 1992 Oct 1;89(19):9196-200.

[141] Gaertig J, Thatcher TH, Gu L, Gorovsky MA. Electroporation-mediated replacement of a positively and negatively selectable beta-tubulin gene in Tetrahymena thermophila. Proc Natl Acad Sci U S A 1994 May 10;91(10):4549-53.

[142] Weide T, Bockau U, Rave A, Herrmann L, Hartmann MW. A recombinase system facilitates cloning of expression cassettes in the ciliate Tetrahymena thermophila. BMC Microbiol 2007;7:12.

[143] Pan WJ, Blackburn EH. Tandem repeats of the 5' non-transcribed spacer of Tetrahymena rDNA function as high copy number autonomous replicons in the macronucleus but do not prevent rRNA gene dosage regulation. Nucleic Acids Res 1995 May 11;23(9):1561-9.

[144] Gaertig J, Gu L, Hai B, Gorovsky MA. High frequency vector-mediated transformation and gene replacement in Tetrahymena. Nucleic Acids Res 1994 Dec 11;22(24):5391-8.

[145] Gaertig J, Gorovsky MA. DNA-mediated transformation in Tetrahymena. Methods Cell Biol 1995;47:559-69.

[146] Mochizuki K, Gorovsky MA. Small RNAs in genome rearrangement in Tetrahymena. Curr Opin Genet Dev 2004 Apr;14(2):181-7.

[147] Herrmann L, Bockau U, Tiedtke A, Hartmann MW, Weide T. The bifunctional dihydrofolate reductase thymidylate synthase of Tetrahymena thermophila provides a tool for molecular and biotechnology applications. BMC Biotechnol 2006;6:21.

[148] Wilson CF, Anand R, Clark JT, McBride JS. Topography of epitopes on a polymorphic schizont antigen of Plasmodium falciparum determined by the binding of monoclonal antibodies in a two-site radioimmunoassay. Parasite Immunol 1987 Nov;9(6):737-46.

[149] Yu GL, Blackburn EH. Transformation of Tetrahymena thermophila with a mutated circular ribosomal DNA plasmid vector. Proc Natl Acad Sci U S A 1989 Nov;86(21):8487-91.

[150] Yao MC, Yao CH. Accurate processing and amplification of cloned germ line copies of ribosomal DNA injected into developing nuclei of Tetrahymena thermophila. Mol Cell Biol 1989 Mar;9(3):1092-9.

[151] Pan WJ, Gallagher RC, Blackburn EH. Replication of an rRNA gene origin plasmid in the Tetrahymena thermophila macronucleus is prevented by transcription through the origin from an RNA polymerase I promoter. Mol Cell Biol 1995 Jun;15(6):3372-81.

[152] Mochizuki K. High efficiency transformation of Tetrahymena using a codon-optimized neomycin resistance gene. Gene 2008 Dec 1;425(1-2):79-83.

[153] Dondero F, Cavaletto M, Ghezzi AR, La TA, Banni M, Viarengo A. Biochemical characterization and quantitative gene expression analysis of the multi-stress inducible metallothionein from Tetrahymena thermophila. Protist 2004 Jun;155(2):157-68.

[154] Piccinni E, Bertaggia D, Santovito G, Miceli C, Kraev A. Cadmium metallothionein gene of Tetrahymena pyriformis. Gene 1999 Jun 24;234(1):51-9.

[155] Gutierrez JC, Amaro F, Martin-Gonzalez A. From heavy metal-binders to biosensors: ciliate metallothioneins discussed. Bioessays 2009 Jul;31(7):805-16.

[156] Diaz S, Amaro F, Rico D, et al. Tetrahymena metallothioneins fall into two discrete subfamilies. PLoS One 2007;2(3):e291.

[157] Boldrin F, Santovito G, Formigari A, et al. MTT2, a copper-inducible metallothionein gene from Tetrahymena thermophila. Comp Biochem Physiol C Toxicol Pharmacol 2008 Mar;147(2):232-40.

[158] Hartmann M.W.W. Gewinnung von Produktionsstämmen und Entwicklung eines Verfahrens zur Produktion von Phospholipase A1 aus Tetrahymena thermophila. Westfälische Wilhelms-Universität Münster. Institut für allgemeine Zoologie und Genetik.; 1999.

[159] Broermann A. Entwicklung eines Testsystems zur Auffindung sekretionsdfizienter Mutanten bei Tetrahymena thermophila. Diplomarbeit im Studiengang Biotechnologie. Westfälische Wilhelms-Universität Münster. Institut für allgemeine Zoologie und Genetik.; 2005.

[160] Cole ES, Bruns PJ. Uniparental cytogamy: a novel method for bringing micronuclear mutations of Tetrahymena into homozygous macronuclear expression with precocious sexual maturity. Genetics 1992 Dec;132(4):1017-31.

[161] Gaertig J, Thatcher TH, McGrath KE, Callahan RC, Gorovsky MA. Perspectives on tubulin isotype function and evolution based on the observation that Tetrahymena thermophila microtubules contain a single alpha- and beta-tubulin. Cell Motil Cytoskeleton 1993;25(3):243-53.

[162] Bannon GA, Bowen JK, Yao MC, Gorovsky MA. Tetrahymena H4 genes: structure, evolution and organization in macro- and micronuclei. Nucleic Acids Res 1984 Feb 24;12(4):1961-75.

[163] Santovito G, Formigari A, Boldrin F, Piccinni E. Molecular and functional evolution of Tetrahymena metallothioneins: new insights into the gene family of Tetrahymena thermophila. Comp Biochem Physiol C Toxicol Pharmacol 2007 Jan;144(4):391-7.

[164] Kyhse-Andersen J. Electroblotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. J Biochem Biophys Methods 1984 Dec;10(3-4):203-9.

[165] Tovey ER, Baldo BA. Comparison by electroblotting of IgE-binding components in extracts of house dust mite bodies and spent mite culture. J Allergy Clin Immunol 1987 Jan;79(1):93-102.

[166] Birch JR, Racher AJ. Antibody production. Adv Drug Deliv Rev 2006 Aug 7;58(5-6):671-85.

[167] Satir B, Sale WS, Satir P. Membrane renewal after dibucaine deciliation of Tetrahymena. Freeze-fracture technique, cilia, membrane structure. Exp Cell Res 1976 Jan;97:83-91.

[168] Satir B. Dibucaine-induced synchronous mucocyst secretion in Tetrahymena. Cell Biol Int Rep 1977 Jan;1(1):69-73.

[169] Baneyx F. Recombinant protein expression in Escherichia coli. Curr Opin Biotechnol 1999 Oct;10(5):411-21.

[170] Cox MM. Cell-based protein vaccines for influenza. Curr Opin Mol Ther 2005 Feb;7(1):24-9.

[171] Brands R, Visser J, Medema J, Palache AM, van Scharrenburg GJ. Influvac: a safe Madin Darby Canine Kidney (MDCK) cell culture-based influenza vaccine. Dev Biol Stand 1999;98:93-100.

[172] Alymova IV, Kodihalli S, Govorkova EA, Fanget B, Gerdil C, Webster RG. Immunogenicity and protective efficacy in mice of influenza B virus vaccines grown in mammalian cells or embryonated chicken eggs. J Virol 1998 May;72(5):4472-7.

[173] Robertson JS, Naeve CW, Webster RG, Bootman JS, Newman R, Schild GC. Alterations in the hemagglutinin associated with adaptation of influenza B virus to growth in eggs. Virology 1985 May;143(1):166-74.

[174] Schild GC, Oxford JS, de Jong JC, Webster RG. Evidence for host-cell selection of influenza virus antigenic variants. Nature 1983 Jun 23;303(5919):706-9.

[175] Lakey DL, Treanor JJ, Betts RF, et al. Recombinant baculovirus influenza A hemagglutinin vaccines are well tolerated and immunogenic in healthy adults. J Infect Dis 1996 Oct;174(4):838-41.

[176] Powers DC, McElhaney JE, Florendo OA, Jr., et al. Humoral and cellular immune responses following vaccination with purified recombinant hemagglutinin from influenza A (H3N2) virus. J Infect Dis 1997 Feb;175(2):342-51.

[177] Powers DC, Smith GE, Anderson EL, et al. Influenza A virus vaccines containing purified recombinant H3 hemagglutinin are well tolerated and induce protective immune responses in healthy adults. J Infect Dis 1995 Jun;171(6):1595-9.

[178] Treanor JJ, Wilkinson BE, Masseoud F, et al. Safety and immunogenicity of a recombinant hemagglutinin vaccine for H5 influenza in humans. Vaccine 2001 Feb 8;19(13-14):1732-7.

[179] Treanor JJ, Betts RF, Smith GE, et al. Evaluation of a recombinant hemagglutinin expressed in insect cells as an influenza vaccine in young and elderly adults. J Infect Dis 1996 Jun;173(6):1467-70.

[180] King JC, Jr., Cox MM, Reisinger K, Hedrick J, Graham I, Patriarca P. Evaluation of the safety, reactogenicity and immunogenicity of FluBlok trivalent recombinant baculovirus-expressed hemagglutinin influenza vaccine administered intramuscularly to healthy children aged 6-59 months. Vaccine 2009 Nov 5;27(47):6589-94.

[181] Kuroda K, Geyer H, Geyer R, Doerfler W, Klenk HD. The oligosaccharides of influenza virus hemagglutinin expressed in insect cells by a baculovirus vector. Virology 1990 Feb;174(2):418-29.

[182] Cox MM, Karl AD. Production of a novel influenza vaccine using insect cells: protection against drifted strains. Influenza Other Respi Viruses 2007 Jan;1(1):35-40.

[183] Wuitschick JD, Karrer KM. Codon usage in Tetrahymena thermophila. Methods Cell Biol 2000;62:565-8.

[184] Cereghino JL, Cregg JM. Heterologous protein expression in the methylotrophic yeast Pichia pastoris. FEMS Microbiol Rev 2000 Jan;24(1):45-66.

[185] Tschopp JF, Brust PF, Cregg JM, Stillman CA, Gingeras TR. Expression of the lacZ gene from two methanol-regulated promoters in Pichia pastoris. Nucleic Acids Res 1987 May 11;15(9):3859-76.

[186] Hartner FS, Glieder A. Regulation of methanol utilisation pathway genes in yeasts. Microb Cell Fact 2006;5:39.

[187] Maya D, Quintero MJ, de la Cruz Munoz-Centeno, Chavez S. Systems for applied gene control in Saccharomyces cerevisiae. Biotechnol Lett 2008 Jun;30(6):979-87.

[188] Hamer DH. Metallothionein. Annu Rev Biochem 1986;55:913-51.

[189] Waterham HR, Digan ME, Koutz PJ, Lair SV, Cregg JM. Isolation of the Pichia pastoris glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. Gene 1997 Feb 20;186(1):37-44.

[190] Li P, Anumanthan A, Gao XG, et al. Expression of recombinant proteins in Pichia pastoris. Appl Biochem Biotechnol 2007 Aug;142(2):105-24.

[191] Hartmann M.W.W., Grenningloh M., Tiedtke A, inventors; Cilian AG, assignee. DNA sequence of the enzyme phospholipase A1 of ciliate Tetrahymena, and the use of the same. US patent 7,045,330 B2. 2006 May 16.

[192] Straus JW, Migaki G, Finch MT. An assessment of proteolytic enzymes in Tetrahymena thermophila. J Protozool 1992 Nov;39(6):655-62.

[193] Haas AL, Sabina RL. Expression, purification, and inhibition of in vitro proteolysis of human AMPD2 (isoform L) recombinant enzymes. Protein Expr Purif 2003 Feb;27(2):293-303.

[194] Cereghino GP, Cereghino JL, Ilgen C, Cregg JM. Production of recombinant proteins in fermenter cultures of the yeast Pichia pastoris. Curr Opin Biotechnol 2002 Aug;13(4):329-32.

[195] Armstrong RT, Kushnir AS, White JM. The transmembrane domain of influenza hemagglutinin exhibits a stringent length requirement to support the hemifusion to fusion transition. J Cell Biol 2000 Oct 16;151(2):425-37.

[196] Schuster M, Wasserbauer E, Aversa G, Jungbauer A. Transmembrane-sequence-dependent overexpression and secretion of glycoproteins in Saccharomyces cerevisiae. Protein Expr Purif 2001 Feb;21(1):1-7.

[197] de la Cruz Edmonds MC, Tellers M, Chan C, Salmon P, Robinson DK, Markusen J. Development of transfection and high-producer screening protocols for the CHOK1SV cell system. Mol Biotechnol 2006 Oct;34(2):179-90.

[198] Blobel G, Walter P, Chang CN, Goldman BM, Erickson AH, Lingappa VR. Translocation of proteins across membranes: the signal hypothesis and beyond. Symp Soc Exp Biol 1979;33:9-36.

[199] Chaudhuri B, Steube K, Stephan C. The pro-region of the yeast prepro-alpha-factor is essential for membrane translocation of human insulin-like growth factor 1 in vivo. Eur J Biochem 1992 Jun 15;206(3):793-800.

[200] Richard W.Scott, Russel A.Brierley, David S.Howland, inventors; Cephalon IWCPa, assignee. Secretion sequence for the production of a heterologous protein in yeast. United States patent 5,521,086. 1996 May 28.

[201] Arvan P, Castle D. Protein sorting and secretion granule formation in regulated secretory cells. Trends Cell Biol 1992 Nov;2(11):327-31.

[202] Haddad A, Bowman GR, Turkewitz AP. New class of cargo protein in Tetrahymena thermophila dense core secretory granules. Eukaryot Cell 2002 Aug;1(4):583-93.

[203] Vassileva A, Chugh DA, Swaminathan S, Khanna N. Effect of copy number on the expression levels of hepatitis B surface antigen in the methylotrophic yeast Pichia pastoris. Protein Expr Purif 2001 Feb;21(1):71-80.

[204] Kiy T, Tiedtke A. Lysosomal enzymes produced by immobilized Tetrahymena thermophila. Appl Microbiol Biotechnol 1991 Apr;35(1):14-8.

[205] Meghrous J, Mahmoud W, Jacob D, Chubet R, Cox M, Kamen AA. Development of a simple and high-yielding fed-batch process for the production of influenza vaccines. Vaccine 2009 Dec 11;28(2):309-16.

[206] Geng X, Wang L. Liquid chromatography of recombinant proteins and protein drugs. J Chromatogr B Analyt Technol Biomed Life Sci 2008 Apr 15;866(1-2):133-53.

[207] Ogata S, Muramatsu T, Kobata A. Fractionation of glycopeptides by affinity column chromatography on concanavalin A-sepharose. J Biochem 1975 Oct;78(4):687-96.

[208] Schulze IT. Effects of glycosylation on the properties and functions of influenza virus hemagglutinin. J Infect Dis 1997 Aug;176 Suppl 1:S24-S28.

[209] Wagner R, Wolff T, Herwig A, Pleschka S, Klenk HD. Interdependence of hemagglutinin glycosylation and neuraminidase as regulators of influenza virus growth: a study by reverse genetics. J Virol 2000 Jul;74(14):6316-23.

[210] Helenius A, Aebi M. Intracellular functions of N-linked glycans. Science 2001 Mar 23;291(5512):2364-9.

[211] Higashi H, Naiki M, Matuo S, Okouchi K. Antigen of "serum sickness" type of heterophile antibodies in human sera: indentification as gangliosides with N-glycolylneuraminic acid. Biochem Biophys Res Commun 1977 Nov 21;79(2):388-95.

[212] Hartgroves LC, Koudstaal W, McLeod C, et al. Rapid generation of a well-matched vaccine seed from a modern influenza A virus primary isolate without recourse to eggs. Vaccine 2010 Apr 9;28(17):2973-9.

[213] Mochalova L, Gambaryan A, Romanova J, et al. Receptor-binding properties of modern human influenza viruses primarily isolated in Vero and MDCK cells and chicken embryonated eggs. Virology 2003 Sep 1;313(2):473-80.

[214] Chen Z, Aspelund A, Jin H. Stabilizing the glycosylation pattern of influenza B hemagglutinin following adaptation to growth in eggs. Vaccine 2008 Jan 17;26(3):361-71.

[215] Hartmann M, Guberman A, Florin-Christensen M, Tiedtke A. Screening for and characterization of phospholipase A1 hypersecretory mutants of Tetrahymena thermophila. Appl Microbiol Biotechnol 2000 Sep;54(3):390-6.

[216] Schwarzer J, Rapp E, Hennig R, et al. Glycan analysis in cell culture-based influenza vaccine production: influence of host cell line and virus strain on the glycosylation pattern of viral hemagglutinin. Vaccine 2009 Jul 9;27(32):4325-36.

8. Appendix

8.1 Alignment of the influenza A virus strain New Caledonia/20/99 (H1N1) hemagglutinin

```
  M  D  I  K  A  K  L  L  V  L  L  C  T  F  T  A  T  Y  A  D  T  I  C  I  G
atg------aaagcaaaactactggtcctgttatgtacatttacagctacatatgcagacacaatatgtataggc
atggatatcaaagctaagttattagtttacttgtacttccactgctacttatgctgatactatttgtattggt

  Y  H  A  N  N  S  T  D  T  V  D  T  V  L  E  K  N  V  T  V  T  H  S  V  N
taccatgccaacaactcaaccgacactgttgacacagtacttgagaagaatgtgacagtgacacactctgtcaac
tatcacgctaataattctactgatactgttgatactgtttagaaaaaatgttactgttactcactctgttaat

  L  L  E  D  S  H  N  G  K  L  C  L  L  K  G  I  A  P  L  Q  L  G  N  C  S
ctacttgaggacagtcacaatggaaaactatgtctactaaaaggaatagccccactacaattgggtaattgcagc
ctttagaagattctcacaatggtaaattatgttacttaaaggtattgctcctttataacttggtaattgttct

  V  A  G  W  I  L  G  N  P  E  C  E  L  L  I  S  K  E  S  W  S  Y  I  V  E
gttgccggatggatcttaggaaacccagaatgcgaattactgatttccaaggaatcatggtcctacattgtagaa
gttgctggttggattcttggtaatcctgaatgtgaattacttatttctaaagaatcttggtcttatattgttgaa

  T  P  N  P  E  N  G  T  C  Y  P  G  Y  F  A  D  Y  E  E  L  R  E  Q  L  S
acaccaaatcctgagaatggaacatgttacccagggtatttcgccgactatgaggaactgagggagcaattgagt
actcctaatcctgaaaatggtacttgttatcctggttatttcgctgattatgaagaattaagagaataacttttct

  S  V  S  S  F  E  R  F  E  I  F  P  K  E  S  S  W  P  N  H  T  V  T  G  V
tcagtatcttcatttgagagattcgaaatattccccaaagaaagctcatggcccaaccacaccgtaaccggagta
tctgttcttcatttgaaagattcgaaatttttccctaaagaatcttcttggcctaatcacactgttactggtgtt

  S  A  S  C  S  H  N  G  K  S  S  F  Y  R  N  L  L  W  L  T  G  K  N  G  L
tcagcatcatgctcccataatgggaaaagcagtttttacagaaatttgctatggctgacggggaagaatggtttg
tctgcttcttgttctcataatggtaaatcttcattctatagaaatctttatggttaactggtaaaaatggtttta

  Y  P  N  L  S  K  S  Y  V  N  N  K  E  K  E  V  L  V  L  W  G  V  H  H  P
tacccaaacctgagcaagtcctatgtaaacaacaaagagaaagaagtccttgtactatggggtgttcatcacccg
tatcctaatcttttctaaatcttatgttaataataaggaaaagaagtttagttcttggggtgttcaccaccct

  P  N  I  G  N  Q  R  A  L  Y  H  T  E  N  A  Y  V  S  V  V  S  S  H  Y  S
cctaacatagggaaccaaagggccctctatcatacagaaaatgcttatgtctctgtagtgtcttcacattatagc
cctaatattggtaattaaagagctttatatcacactgaaaatgcttatgttctgttgttcttctcactattct

  R  R  F  T  P  E  I  A  K  R  P  K  V  R  D  Q  E  G  R  I  N  Y  Y  W  T
agaagattcaccccagaaatagccaaaagacccaaagtaagagatcaggaaggaagaatcaactactactggact
agaagattcactcctgaaattgctaaaagacctaaagttagagattaagaaggtagaattaactattattggact

  L  L  E  P  G  D  T  I  I  F  E  A  N  G  N  L  I  A  P  W  Y  A  F  A  L
ctgctggaacctggggatacaataatatttgaggcaaatggaaatctaatagcgccatggtatgcttttgcactg
ttattagaacctggtgatactattatttcgaagctaatggtaatcttattgctccttggtacgctttcgctttta

  S  R  G  F  G  S  G  I  I  T  S  N  A  P  M  D  E  C  D  A  K  C  Q  T  P
agtagaggctttggatcaggaatcatcacctcaaatgcaccaatggatgaatgtgatgcgaagtgtcaaacacct
tctagaggtttcggttctggtattattacttctaatgctcctatggatgaatgtgatgctaaatgttaaactcct
```

```
  Q  G  A  I  N  S  S  L  P  F  Q  N  V  H  P  V  T  I  G  E  C  P  K  Y  V
cagggagctataaacagcagtcttcctttccagaatgtacacccagtcacaataggagagtgtccaaagtatgtc
taaggtgctattaattcttcttttaccttttctaaaacgttcaccctgttactattggtgaatgtcctaaatatgtt

  R  S  A  K  L  R  M  V  T  G  L  R  N  I  P  S  I  Q  S  R  G  L  F  G  A
aggagtgcaaaattaaggatggttacaggactaaggaacatcccatccattcaatccagaggtttgtttggagcc
agatcagctaaattaagaatggttactggtttaagaaatattccttctatctaatctagaggactttcggtgct

  I  A  G  F  I  E  G  G  W  T  G  M  V  D  G  W  Y  G  Y  H  H  Q  N  E  Q
attgccggtttcattgaaggggggtggactggaatggtagatgggtggtatggttatcatcatcagaatgagcaa
attgctggtttcattgagggaggatggacaggtatggttgatggttggtacggttatcaccaccaaaatgaataa

  G  S  G  Y  A  A  D  Q  K  S  T  Q  N  A  I  N  G  I  T  N  K  V  N  S  V
ggatctggctatgctgcagatcaaaaaagtacacaaaatgccattaacgggattacaaacaaggtgaattctgta
ggttctggttatgctgctgattaaaagtctactcaaaacgctattaatggtattactaacaaagttaattctgtt

  I  E  K  M  N  T  Q  F  T  A  V  G  K  E  F  N  K  L  E  R  R  M  E  N  L
attgagaaaatgaacactcaattcacagctgtgggcaaagaattcaacaaattggaaagaaggatggaaaactta
attgaaaaaatgaatacttaattcactgctgttggtaaagaattcaataagttagaaagaagaatggaaaatctt

  N  K  K  V  D  D  G  F  L  D  I  W  T  Y  N  A  E  L  L  V  L  L  E  N  E
aataaaaaagttgatgatgggtttctagacatttggacatataatgcagaattgttggttctactggaaaatgaa
aataagaaagttgatgatggtttcttagatatttggacttataatgctgaattattagttctttagaaaatgaa

  R  T  L  D  F  H  D  S  N  V  K  N  L  Y  E  K  V  K  S  Q  L  K  N  N  A
aggactttggatttccatgactccaatgtgaagaatctgtatgagaaagtaaaaagccaattaaagaataatgcc
agaactttagatttccacgattctaatgttaaaaatcttttatgaaaagttaaatcttaacttaagaataatgct

  K  E  I  G  N  G  C  F  E  F  Y  H  K  C  N  N  E  C  M  E  S  V  K  N  G
aaagaaataggaaacgggtgttttgaattctatcacaagtgtaacaatgaatgcatggagagtgtgaaaaatgga
aaagaaattggtaatggttgtttcgaattctatcacaaatgtaataatgaatgtatggaatctgttaagaatggt

  T  Y  D  Y  P  K  Y  S  E  E  S  K  L  N  R  E  K  I  D  G  V  K  L  E  S
acttatgactatccaaaatattccgaagaatcaaagttaaacagggagaaaattgatggagtgaaattggaatca
acttatgattatcctaaatattctgaagaatctaagttaaatagagaaaaattgatggtgttaaattagaatct

  M  G  V  Y  Q  I  L  A  I  Y  S  T  V  A  S  S  L  V  L  L  V  S  L  G  A
atgggagtctatcagattctggcgatctactcaactgtcgccagttccctggttcttttggtctccctgggggca
atgggagttactaaattcttgctatctattctactgttgcttcttcttagtttattagttcttaggtgct

  I  S  F  W  M  C  S  N  G  S  L  Q  C  R  I  C  I  Q  R  S
atcagcttctggatgtgttccaatgggtctttgcagtgtagaatatgcatctga------
atttcttctggatgtgttctaatggttctttataatgtagaatttgtatttgaagatct
```

The upper line shows the non-optimized original sequence of the hemagglutinin of the Influenza A strain New Caledonia/20/99 (H1N1). The lower line shows the codon-optimized sequence (lower case letters). Differences are highlighted in yellow. The ER-leader sequence is depicted in blue letters, and the transmembrane domain in green letters as amino acid sequence (in capital letters). The added EcoRV restriction site (gatatc) and the BglII restriction site (agatct) are shown in red letters. The stop codon is highlighted in red, and the start codon is highlighted in purple.

8.2 Alignment of the C-terminally truncated influenza A virus strain New Caledonia/20/99 (H1N1) hemagglutinin (NC531)

```
  M  D  I  K  A  K  L  L  V  L  L  C  T  F  T  A  T  Y  A  D  T  I  C  I  G
atg------aaagcaaaactactggtcctgttatgtacatttacagctacatatgcagacacaatatgtataggc
atggatatcaaagctaagttattagtttacttttgtacttccactgctacttatgctgatactatttgtattggt

  Y  H  A  N  N  S  T  D  T  V  D  T  V  L  E  K  N  V  T  V  T  H  S  V  N
taccatgccaacaactcaaccgacactgttgacacagtacttgagaagaatgtgacagtgacacactctgtcaac
tatcacgctaataattctactgatactgttgatactgttttagaaaaaaatgttactgttactcactctgttaat

  L  L  E  D  S  H  N  G  K  L  C  L  L  K  G  I  A  P  L  Q  L  G  N  C  S
ctacttgaggacagtcacaatggaaaactatgtctactaaaaggaatagccccactacaattgggtaattgcagc
ctttagaagattctcacaatggtaaattatgttacttaaaggtattgctccttatatacttggtaattgttct

  V  A  G  W  I  L  G  N  P  E  C  E  L  L  I  S  K  E  S  W  S  Y  I  V  E
gttgccggatggatcttaggaaacccagaatgcgaattactgatttccaaggaatcatggtcctacattgtagaa
gttgctggttggattcttggtaatcctgaatgtgaattacttatttctaaagaatcttggtcttatattgttgaa

  T  P  N  P  E  N  G  T  C  Y  P  G  Y  F  A  D  Y  E  E  L  R  E  Q  L  S
acaccaaatcctgagaatggaacatgttacccagggtatttcgccgactatgaggaactgagggagcaattgagt
actcctaatcctgaaaatggtacttgttatcctggttatttcgctgattatgaagaattaagagaataactttct

  S  V  S  S  F  E  R  F  E  I  F  P  K  E  S  S  W  P  N  H  T  V  T  G  V
tcagtatcttcatttgagagattcgaaatattccccaaagaaagctcatggcccaaccacaccgtaaccggagta
tctgttcttcatttgaaagattcgaaatttttccctaaagaatcttcttggcctaatcacactgttactggtgtt

  S  A  S  C  S  H  N  G  K  S  S  F  Y  R  N  L  L  W  L  T  G  K  N  G  L
tcagcatcatgctcccataatgggaaaagcagtttttacagaaatttgctatggctgacggggaagaatggtttg
tctgcttcttgttctcataatggtaaatcttcattctatagaaatctttatggttaactggtaaaaatggtttta

  Y  P  N  L  S  K  S  Y  V  N  N  K  E  K  E  V  L  V  L  W  G  V  H  H  P
tacccaaacctgagcaagtcctatgtaaacaacaaagagaaagaagtccttgtactatggggtgttcatcacccg
tatcctaatctttctaaatcttatgttaataataaggaaaaagaagtttagttcttttggggtgttcaccaccct

  P  N  I  G  N  Q  R  A  L  Y  H  T  E  N  A  Y  V  S  V  V  S  S  H  Y  S
cctaacatagggaaccaaagggccctctatcatacagaaaatgcttatgtctctgtagtgtcttcacattatagc
cctaatattggtaatttaaagagcttttatatcacactgaaaatgcttatgttttctgttgttcttctcactattct

  R  R  F  T  P  E  I  A  K  R  P  K  V  R  D  Q  E  G  R  I  N  Y  Y  W  T
agaagattcaccccagaaatagccaaaagacccaaagtaagagatcaggaaggaagaatcaactactactggact
agaagattcactcctgaaattgctaaaagacctaaagttagagattaagaaggtagaattaactattattggact

  L  L  E  P  G  D  T  I  I  F  E  A  N  G  N  L  I  A  P  W  Y  A  F  A  L
ctgctggaacctggggatacaataatatttgaggcaaatggaaatctaatagcgccatggtatgcttttgcactg
ttattagaacctggtgatactattatttcgaagctaatggtaatcttattgctccttggtacgctttcgctttta

  S  R  G  F  G  S  G  I  I  T  S  N  A  P  M  D  E  C  D  A  K  C  Q  T  P
agtagaggctttggatcaggaatcatcacctcaaatgcaccaatggatgaatgtgatgcgaagtgtcaaacacct
tctagaggtttcggttctggtattattacttctaatgctcctatggatgaatgtgatgctaaatgttaaactcct

  Q  G  A  I  N  S  S  L  P  F  Q  N  V  H  P  V  T  I  G  E  C  P  K  Y  V
cagggagctataaacagcagtcttcctttccagaatgtacacccagtcacaataggagagtgtccaaagtatgtc
taaggtgctattaattcttcttaccttctaaaacgttcaccctgttactattggtgaatgtcctaaatatgtt

  R  S  A  K  L  R  M  V  T  G  L  R  N  I  P  S  I  Q  S  R  G  L  F  G  A
aggagtgcaaaattaaggatggttacaggactaaggaacatcccatccattcaatccagaggtttgtttggagcc
agatcagctaaattaagaatggttactggtttaagaaatattccttctatctaatctagaggactttcggtgct
```

```
  I  A  G  F  I  E  G  G  W  T  G  M  V  D  G  W  Y  G  Y  H  H  Q  N  E  Q
attgccggtttcattgaaggggggtggactggaatggtagatgggtggtatggttatcatcatcagaatgagcaa
attgctggtttcattgagggaggatggacaggtatggttgatggttggtacggttatcaccaccaaaatgaataa

  G  S  G  Y  A  A  D  Q  K  S  T  Q  N  A  I  N  G  I  T  N  K  V  N  S  V
ggatctggctatgctgcagatcaaaaaagtacacaaaatgccattaacgggattacaaacaaggtgaattctgta
ggttctggttatgctgctgattaaaagtctactcaaaacgctattaatggtattactaacaaagttaattctgtt

  I  E  K  M  N  T  Q  F  T  A  V  G  K  E  F  N  K  L  E  R  R  M  E  N  L
attgagaaaatgaacactcaattcacagctgtgggcaaagaattcaacaaattggaaagaaggatggaaaactta
attgaaaaaatgaatacttaattcactgctgttggtaaagaattcaataagttagaaagaagaatggaaaatctt

  N  K  K  V  D  D  G  F  L  D  I  W  T  Y  N  A  E  L  L  V  L  L  E  N  E
aataaaaaagttgatgatgggtttctagacatttggacatataatgcagaattgttggttctactggaaaatgaa
aataagaaagttgatgatggtttcttagatatttggacttataatgctgaattattagttctttagaaaatgaa

  R  T  L  D  F  H  D  S  N  V  K  N  L  Y  E  K  V  K  S  Q  L  K  N  N  A
aggactttggatttccatgactccaatgtgaagaatctgtatgagaaagtaaaaagccaattaaagaataatgcc
agaactttagatttccacgattctaatgttaaaaatcttatgaaaagttaaatcttaacttaagaataatgct

  K  E  I  G  N  G  C  F  E  F  Y  H  K  C  N  N  E  C  M  E  S  V  K  N  G
aaagaaataggaaacgggtgttttgaattctatcacaagtgtaacaatgaatgcatggagagtgtgaaaaatgga
aaagaaattggtaatggttgtttcgaattctatcacaaatgtaataatgaatgtatggaatctgttaagaatggt

  T  Y  D  Y  P  K  Y  S  E  E  S  K  L  N  R  E  K  I  D  G  V  K  L  E  S
acttatgactatccaaaatattccgaagaatcaaagttaaacagggagaaaattgatggagtgaaattggaatca
acttatgattatcctaaatattctgaagaatctaagttaaatagagaaaaattgatggtgttaaattagaatct

  M  G  V  Y  Q  I  L  Q  R  S  A  I  Y  S  T  V  A  S  S  L  V  L  L  V  S
atgggagtctatcagattctg---------gcgatctactcaactgtcgccagttccctggttcttttggtctcc
atgggagttactaaattcttTGAagatct----------------------------------------------

  L  G  A  I  S  F  W  M  C  S  N  G  S  L  Q  C  R  I  C  I  Q
ctgggggcaatcagcttctggatgtgttccaatgggtctttgcagtgtagaatatgcatctga
------------------------------------------------------------
```

The upper line shows the non-optimized original sequence of the hemagglutinin of the Influenza A strain New Caledonia/20/99 (H1N1). The lower line shows the codon-optimized sequence (lower case letters). Differences are highlighted in yellow. The ER-leader sequence is depicted in blue letters, and the transmembrane domain in green letters as amino acid sequence (in capital letters). The added EcoRV restriction site (gatatc) and the BglII restriction site (agatct) are shown in red letters. The stop codon is highlighted in red, and the start codon is highlighted in purple. The truncation of the HA is demonstrated.

8.3 Alignment of the C-terminally truncated influenza A virus strain New Caledonia/20/99 (H1N1) hemagglutinin (NC344)

```
  M  D  I  K  A  K  L  L  V  L  L  C  T  F  T  A  T  Y  A  D  T  I  C  I  G
atg------aaagcaaaactactggtcctgttatgtacatttacagctacatatgcagacacaatatgtataggc
atggatatcaaagctaagttattagtttacttgtacttccactgctacttatgctgatactatttgtattggt

  Y  H  A  N  N  S  T  D  T  V  D  T  V  L  E  K  N  V  T  V  T  H  S  V  N
taccatgccaacaactcaaccgacactgttgacacagtacttgagaagaatgtgacagtgacacactctgtcaac
tatcacgctaataattctactgatactgttgatactgttttagaaaaaaatgttactgttactcactctgttaat

  L  L  E  D  S  H  N  G  K  L  C  L  L  K  G  I  A  P  L  Q  L  G  N  C  S
ctacttgaggacagtcacaatggaaaactatgtctactaaaaggaatagccccactacaattgggtaattgcagc
ctttagaagattctcacaatggtaaattatgttacttaaaggtattgctccttatatacttggtaattgttct

  V  A  G  W  I  L  G  N  P  E  C  E  L  L  I  S  K  E  S  W  S  Y  I  V  E
gttgccggatggatcttaggaaacccagaatgcgaattactgatttccaaggaatcatggtcctacattgtagaa
gttgctggttggattcttggtaatcctgaatgtgaattacttatttctaaagaatcttggtcttatattgttgaa

  T  P  N  P  E  N  G  T  C  Y  P  G  Y  F  A  D  Y  E  E  L  R  E  Q  L  S
acaccaaatcctgagaatggaacatgttacccagggtatttcgccgactatgaggaactgagggagcaattgagt
actcctaatcctgaaaatggtacttgttatcctggttatttcgctgattatgaagaattaagagaataactttct

  S  V  S  S  F  E  R  F  E  I  F  P  K  E  S  S  W  P  N  H  T  V  T  G  V
tcagtatcttcatttgagagattcgaaatattccccaaagaaagctcatggcccaaccacaccgtaaccggagta
tctgttcttcatttgaaagattcgaaatttttccctaaagaatcttcttggcctaatcacactgttactggtgtt

  S  A  S  C  S  H  N  G  K  S  S  F  Y  R  N  L  L  W  L  T  G  K  N  G  L
tcagcatcatgctcccataatgggaaaagcagtttttacagaaatttgctatggctgacggggaagaatggtttg
tctgcttcttgttctcataatggtaaatcttcattctatagaaatctttatggttaactggtaaaaatggtttta

  Y  P  N  L  S  K  S  Y  V  N  N  K  E  K  E  V  L  V  L  W  G  V  H  H  P
tacccaaacctgagcaagtcctatgtaaacaacaaagagaaagaagtccttgtactatggggtgttcatcacccg
tatcctaatctttctaaatcttatgttaataataaggaaaaagaagtttagttcttttggggtgttcaccaccct

  P  N  I  G  N  Q  R  A  L  Y  H  T  E  N  A  Y  V  S  V  V  S  S  H  Y  S
cctaacatagggaaccaaagggccctctatcatacagaaaatgcttatgtctctgtagtgtcttcacattatagc
cctaatattggtaatttaaagagcttttatatcacactgaaaatgcttatgttttctgttgttcttctcactattct

  R  R  F  T  P  E  I  A  K  R  P  K  V  R  D  Q  E  G  R  I  N  Y  Y  W  T
agaagattcaccccagaaatagccaaaagacccaaagtaagagatcaggaaggaagaatcaactactactggact
agaagattcactcctgaaattgctaaaagacctaaagttagagattaagaaggtagaattaactattattggact

  L  L  E  P  G  D  T  I  I  F  E  A  N  G  N  L  I  A  P  W  Y  A  F  A  L
ctgctggaacctggggatacaataatatttgaggcaaatggaaatctaatagcgccatggtatgcttttgcactg
ttattagaacctggtgatactattatttcgaagctaatggtaatcttattgctccttggtacgctttcgctttta

  S  R  G  F  G  S  G  I  I  T  S  N  A  P  M  D  E  C  D  A  K  C  Q  T  P
agtagaggctttggatcaggaatcatcacctcaaatgcaccaatggatgaatgtgatgcgaagtgtcaaacacct
tctagaggtttcggttctggtattattacttctaatgctcctatggatgaatgtgatgctaaatgttaaactcct

  Q  G  A  I  N  S  S  L  P  F  Q  N  V  H  P  V  T  I  G  E  C  P  K  Y  V
cagggagctataaacagcagtcttcctttccagaatgtacacccagtcacaataggagagtgtccaaagtatgtc
taaggtgctattaattcttcttaccttctaaaacgttcaccctgttactattggtgaatgtcctaaatatgtt

  R  S  A  K  L  R  M  V  T  G  L  R  N  I  P  S  I  Q  S  R  Q  R  S  G  L
aggagtgcaaaattaaggatggttacaggactaaggaacatcccatccattcaatccaga---------ggtttg
agatcagctaaattaagaatggttactggtttaagaaatattccttctatctaatctagatgaagatct------
```

```
  F  G  A  I  A  G  F  I  E  G  G  W  T  G  M  V  D  G  W  Y  G  Y  H  H  Q
attgccggtttcattgaaggggggtggactggaatggtagatgggtggtatggttatcatcatcagaatgagcaa
------------------------------------------------------------------------------
  N  E  Q  G  S  G  Y  A  A  D  Q  K  S  T  Q  N  A  I  N  G  I  T  N  K  V
tttggagccggatctggctatgctgcagatcaaaaaagtacacaaaatgccattaacgggattacaaacaaggtg
------------------------------------------------------------------------------

  N  S  V  I  E  K  M  N  T  Q  F  T  A  V  G  K  E  F  N  K  L  E  R  R  M
aattctgtaattgagaaaatgaacactcaattcacagctgtgggcaaagaattcaacaaattggaaagaaggatg
------------------------------------------------------------------------------

  E  N  L  N  K  K  V  D  D  G  F  L  D  I  W  T  Y  N  A  E  L  L  V  L  L
gaaaacttaaataaaaaagttgatgatgggtttctagacatttggacatataatgcagaattgttggttctactg
------------------------------------------------------------------------------

  E  N  E  R  T  L  D  F  H  D  S  N  V  K  N  L  Y  E  K  V  K  S  Q  L  K
gaaaatgaaaggactttggatttccatgactccaatgtgaagaatctgtatgagaaagtaaaaagccaattaaag
------------------------------------------------------------------------------

  N  N  A  K  E  I  G  N  G  C  F  E  F  Y  H  K  C  N  N  E  C  M  E  S  V
aataatgccaaagaaataggaaacgggtgttttgaattctatcacaagtgtaacaatgaatgcatggagagtgtg
------------------------------------------------------------------------------

  K  N  G  T  Y  D  Y  P  K  Y  S  E  E  S  K  L  N  R  E  K  I  D  G  V  K
aaaaatggaacttatgactatccaaaatattccgaagaatcaaagttaaacagggagaaaattgatggagtgaaa
------------------------------------------------------------------------------

  L  E  S  M  G  V  Y  Q  I  L  A  I  Y  S  T  V  A  S  S  L  V  L  L  V  S
ttggaatcaatgggagtctatcagattctggcgatctactcaactgtcgccagttccctggttcttttggtctcc
------------------------------------------------------------------------------

  L  G  A  I  S  F  W  M  C  S  N  G  S  L  Q  C  R  I  C  I  Q
ctgggggcaatcagcttctggatgtgttccaatgggtctttgcagtgtagaatatgcatctga
------------------------------------------------------------
```

The upper line shows the non-optimized original sequence of the hemagglutinin of the Influenza A strain New Caledonia/20/99 (H1N1). The lower line shows the codon-optimized sequence (lower case letters). Differences are highlighted in yellow. The ER-leader sequence is depicted in blue letters, and the transmembrane domain in green letters as amino acid sequence (in capital letters). The added EcoRV restriction site (gatatc) and the BglII restriction site (agatct) are shown in red letters. The stop codon is highlighted in red, and the start codon is highlighted in purple. The truncation of the HA is demonstrated.

8.4 Alignment of the influenza A virus strain Uruguay/716/2007 (H3N2) hemagglutinin

```
  M  D  I  K  T  I  I  A  L  S  Y  I  L  C  L  V  F  T  Q  K  L  P  G  N  D
atg------aagactatcattgctttgagctacattctatgtctggttttcactcaaaaacttcccggaaatgac
atggatatcaagactattattgctcttcttacattcttgccttgttttcacttaaaagttacctggtaatgat

  N  S  T  A  T  L  C  L  G  H  H  A  V  P  N  G  T  I  V  K  T  I  T  N  D
aacagcacggcaacgctgtgccttgggcaccatgcagtaccaaacggaacgatagtgaaaacaatcacgaatgac
aattctactgctactttatgccttggtcaccacgctgttcctaatggtactattgttaagactattactaacgat

  Q  I  E  V  T  N  A  T  E  L  V  Q  S  S  S  T  G  E  I  C  D  S  P  H  Q
caaattgaagttactaatgctactgagctggttcagagttcctcaacaggtgaaatatgcgacagtcctcatcag
caaattgaagttactaatgctactgaacttgtttaatcttcttctactggtgaaatttgcgattcctcactaa

  I  L  D  G  E  N  C  T  L  I  D  A  L  L  G  D  P  Q  C  D  G  F  Q  N  K
atccttgatggagaaaactgcacactaatagatgctctattgggagaccctcagtgtgatggcttccaaaataag
attcttgatggtgaaaattgcactcttattgatgcttacttggtgatccttaatgtgatggtttccaaaacaag

  K  W  D  L  F  V  E  R  S  K  A  Y  S  N  C  Y  P  Y  D  V  P  D  Y  A  S
aaatgggacctttttgttgaacgcagcaaagcctacagcaactgttacccttatgatgtgccggattatgcctcc
aagtgggatctttttcgttgaaagatcaaaggcttactctaattgctacccttacgatgttcctgattacgcttct

  L  R  S  L  V  A  S  S  G  T  L  E  F  N  N  E  S  F  N  W  T  G  V  T  Q
cttaggtcactagttgcctcatccggcacactggagtttaacaatgaaagcttcaattggactggagtcactcaa
cttagaagtcttgttgcttcttctggtacttcttgaattcaataatgaatcttttcaattggactggtgttactcaa

  N  G  T  S  S  A  C  I  R  R  S  N  N  S  F  F  S  R  L  N  W  L  T  H  L
aacggaacaagctctgcttgcataaggagatctaataacagtttctttagtagattgaattggttgacccactta
aacggtacttcttctgcttgcattagaagatcaaataattctttcttctctagacttaattggcttactcacctt

  K  F  K  Y  P  A  L  N  V  T  M  P  N  N  E  K  F  D  K  L  Y  I  W  G  V
aaattcaaatacccagcattgaacgtgactatgccaaacaatgaaaaatttgacaaattgtacatttgggggtt
aagttcaagtaccctgctcttaatgttactatgcctaataatgaaaagttcgataagcttacatttggggtgtt

  H  H  P  G  T  D  N  D  Q  I  F  L  Y  A  Q  A  S  G  R  I  T  V  S  T  K
caccacccgggtacggacaatgaccaaatcttcctgtatgctcaagcatcaggaagaatcacagtctctaccaaa
caccaccctggtactgataatgatcaaatttccttacgcttaagcttctggtagaattactgttctactaag

  R  S  Q  Q  T  V  I  P  N  I  G  S  R  P  R  V  R  N  I  P  S  R  I  S  I
agaagccaacaaactgtaatcccgaatatcggatctagacccagagtaaggaatatccccagcagaataagcatc
agatcataacaaactgttattcctaatattggttctagacctagagttagaaatattccttctagaatttctatc

  Y  W  T  I  V  K  P  G  D  I  L  L  I  N  S  T  G  N  L  I  A  P  R  G  Y
tattggacaatagtaaaaccgggagacatacttttgattaacagcacagggaatctaattgctcctaggggttac
tactggactattgttaaacctggtgatattcttcttattaactctactggtaatcttattgctcctagaggatac

  F  K  I  R  S  G  K  S  S  I  M  R  S  D  A  P  I  G  K  C  N  S  E  C  I
ttcaaaatacgaagtgggaaaagctcaataatgagatcagatgcacccattggcaaatgcaattctgaatgcatc
ttcaaaattagaagtggtaaatcttctattatgagatcagatgctcctattggtaaatgcaattctgaatgcatt

  T  P  N  G  S  I  P  N  D  K  P  F  Q  N  V  N  R  I  T  Y  G  A  C  P  R
actccaaacggaagcattcccaatgacaaaccattccaaaatgtaaacaggatcacatacggggcctgtcccaga
actcctaatggttctattcctaatgataagccttccaaaacgttaatagaattacttacggtgcttgccctaga

  Y  V  K  Q  N  T  L  K  L  A  T  G  M  R  N  V  P  E  K  Q  T  R  G  I  F
tatgttaagcaaaacactctgaaattggcaacagggatgcgaaatgtaccagagaaacaaactagaggcatattt
tacgttaagcaaaacactcttaagcttgctactggaatgagaaatgttcctgaaaagtaaactagaggaatttc
```

```
  G  A  I  A  G  F  I  E  N  G  W  E  G  M  V  D  G  W  Y  G  F  R  H  Q  N
ggcgcaatcgcgggtttcatagaaaatggttgggagggaatggtggatggttggtacggtttcaggcatcaaaat
ggtgctattgctggtttcattgaaaatggttgggaaggaatggttgatggttggtacggtttcagacaccaaaac

  S  E  G  I  G  Q  A  A  D  L  K  S  T  Q  A  A  I  D  Q  I  N  G  K  L  N
tctgagggaataggacaagcagcagatctcaaaagcactcaagcagcaatcgatcaaatcaatgggaagctgaat
tctgaaggtattggttaagctgctgatcttaagtctacttaagctgctattgatcaaattaatggaaagcttaac

  R  L  I  G  K  T  N  E  K  F  H  Q  I  E  K  E  F  S  E  V  E  G  R  I  Q
aggttgatcgggaaaaccaacgagaaattccatcagattgaaaaagaattctcagaagtcgaagggagaattcag
agacttattggtaaaactaacgaaaagttccactaaattgaaaaagaattctctgaagttgaaggtagaatctag

  D  L  E  K  Y  V  E  D  T  K  I  D  L  W  S  Y  N  A  E  L  L  V  A  L  E
gaccttgagaaatatgttgaggacaccaaaatagatctctggtcatacaacgcggagcttcttgttgccctggag
gatcttgaaaaatacgttgaagatactaagattgatcttggtcttacaatgctgaacttcttgttgctcttgaa

  N  Q  H  T  I  D  L  T  D  S  E  M  N  K  L  F  E  K  T  K  K  Q  L  R  E
aaccaacatacaattgatctaactgactcagaaatgaacaaactgtttgaaaaaacaaagaagcaactgagggaa
aattaacacactattgatcttactgattctgaaatgaataagctttcgaaaagactaagaagtaacttagagaa

  N  A  E  D  M  G  N  G  C  F  K  I  Y  H  K  C  D  N  A  C  I  G  S  I  R
aatgctgaggatatgggcaatggttgtttcaaaatataccacaaatgtgacaatgcctgcataggatcaatcaga
aacgctgaagatatgggtaatggttgcttcaagatataccacaagtgcgataatgcttgcattggttctattaga

  N  G  T  Y  D  H  D  V  Y  R  D  E  A  L  N  N  R  F  Q  I  K  G  V  E  L
aatggaacttatgaccacgatgtatacagagatgaagcattaaacaaccggttccagatcaagggcgttgagctg
aatggtacttacgatcacgatgttacagagatgaagctcttaacaatagattctaaattaagggtgttgaacttt

  K  S  G  Y  K  D  W  I  L  W  I  S  F  A  I  S  C  F  L  L  C  V  A  L  L
aagtcaggatacaaagattggatcctatggatttcctttgccatatcatgttttttgctttgtgttgctttgttg
aagtctggttacaaggattggattcttttggatttctttcgctattcttgcttccttctttgcgttgctctttta

  G  F  I  Q  R  S
gggttcatcatg------
ggtttcatttgaagatct
```

The upper line shows the non-optimized original sequence of the hemagglutinin of the Influenza A strain Uruguay/716/2007 (H3N2). The lower line shows the codon-optimized sequence (lower case letters). Differences are highlighted in yellow. The ER-leader sequence is depicted in blue letters, and the transmembrane domain in green letters as amino acid sequence (in capital letters). The added EcoRV restriction site (gatatc) and the BglII restriction site (agatct) are shown in red letters. The stop codon is highlighted in red, and the start codon is highlighted in purple.

8.5 Alignment of the influenza B virus strain Florida/4/2006 hemagglutinin

```
  M  D  I  K  A  I  I  V  L  L  M  V  V  T  S  N  A  D  R  I  C  T  G  I  T
atg------aaggcaataattgtactactcatggtagtaacatccaatgcagatcgaatctgcactggaataaca
atggatatcaaggctattattgttcttcttatggttgttacttctaatgctgatagaatttgcactggtattact

  S  S  N  S  P  H  V  V  K  T  A  T  Q  G  E  V  N  V  T  G  V  I  P  L  T
tcttcaaactcacctcatgtggtcaaaacagccactcaaggggaggtcaatgtgactggtgtgataccactaaca
tcatctaattctcctcacgttgttaagactgctacttaaggagaagttaacgttactggtgttattcctcttact

  T  T  P  T  K  S  Y  F  A  N  L  K  G  T  R  T  R  G  K  L  C  P  D  C  L
acaacaccaacaaaatcttattttgcaaatctcaaaggaacaaggaccagagggaaactatgcccagactgtctc
actactcctactaagtcttacttcgctaatcttaagggtactagaactagaggaaaacttgccctgattgcctt

  N  C  T  D  L  D  V  A  L  G  R  P  M  C  V  G  T  T  P  S  A  K  A  S  I
aactgcacagatctggatgtggctttgggcagaccaatgtgtgtggggaccacaccttcggcgaaagcttcaata
aattgcactgatcttgatgttgctcttggtagacctatgtgtgttggtactactccttctgctaaggcttctatt

  L  H  E  V  K  P  V  T  S  G  C  F  P  I  M  H  D  R  T  K  I  R  Q  L  P
ctccacgaagtcaaacctgttacatccgggtgctttcctataatgcacgacagaacaaaaatcaggcaactaccc
cttcacgaagttaagcctgttacttctggttgcttccctattatgcacgatagaactaagattagataacttcct

  N  L  L  R  G  Y  E  N  I  R  L  S  T  Q  N  V  I  D  A  E  K  A  P  G  G
aatcttctcagaggatatgaaaatatcaggctatcaacccaaaacgtcatcgatgcggaaaaggcaccaggagga
aatcttcttagaggttacgaaaacattagacttctactcaaaacgttattgatgctgaaaaggctcctggtggt

  P  Y  R  L  G  T  S  G  S  C  P  N  A  T  S  K  S  G  F  F  A  T  M  A  W
ccctacagacttggaacctcaggatcttgccctaacgctaccagtaagagcggatttttcgcaacaatggcttgg
ccttacagacttggtacttctggttcttgtcctaatgctacttctaagtctggttttcttcgctactatggcttgg

  A  V  P  K  D  N  N  K  N  A  T  N  P  L  T  V  E  V  P  Y  I  C  T  E  G
gctgtcccaaaggacaacaacaaaaatgcaacgaacccactaacagtagaagtaccatacatttgtacagaaggg
gctgttcctaaggataacaataagaatgctactaatcctcttactgttgaagttccttacatttgcactgaagga

  E  D  Q  I  T  V  W  G  F  H  S  D  D  K  T  Q  M  K  N  L  Y  G  D  S  N
gaagaccaaatcactgtttgggggttccattcagatgacaaaacccaaatgaagaacctctatggagactcaaat
gaagatcaaattactgtttggggttccactctgatgataagacttaaatgaagaatcttacggtgattctaat

  P  Q  K  F  T  S  S  A  N  G  V  T  T  H  Y  V  S  Q  I  G  S  F  P  D  Q
cctcaaaagttcacctcatctgctaatggagtaaccacacactatgtttctcagattggcagcttcccagatcaa
ccttaaaagttcacttcttctgctaatggtgttactactcactacgtttcttaaattggttcttttcccttgattaa

  T  E  D  G  G  L  P  Q  S  G  R  I  V  V  D  Y  M  M  Q  K  P  G  K  T  G
acagaagacggaggactaccacaaagcggcaggattgttgttgattacatgatgcaaaaacctgggaaaacagga
actgaagatggtggtttaccttaatctggtagaattgttgttgattacatgatgtaaaagcctggtaaaactggt

  T  I  V  Y  Q  R  G  V  L  L  P  Q  K  V  W  C  A  S  G  R  S  K  V  I  K
acaattgtctaccaaagaggtgttttgttgcctcaaaaggtgtggtgcgcgagtggcaggagcaaagtaataaaa
actattgttactaaagaggtgttcttcttccttaaaaagttggtgcgcttctggtagatcaaaagttattaag

  G  S  L  P  L  I  G  E  A  D  C  L  H  E  K  Y  G  G  L  N  K  S  K  P  Y
gggtccttgcctttaattggtgaagcagattgccttcatgaaaaatacggtggattaaacaaaagcaagccttac
ggttctcttcctcttattggtgaagctgattgcttacacgaaaagtacggtggtcttaataagtctaagccttac

  Y  T  G  E  H  A  K  A  I  G  N  C  P  I  W  V  K  T  P  L  K  L  A  N  G
tacacaggagaacatgcaaaagccataggaaattgcccaatatgggtgaaaacacctttgaagcttgccaatgga
tacactggtgaacacgctaaggctattggtaattgccctatttgggttaagactcctttaaagcttgctaatggt
```

```
  T  K  Y  R  P  P  A  K  L  L  K  E  R  G  F  F  G  A  I  A  G  F  L  E  G
accaaatatagacctcctgcaaaactattaaaggaaaggggtttcttcggagctattgctggtttcctagaagga
actaagtacagacctcctgctaagttacttaaagaaaagaggattcttttggtgctattgctggtttccttgaagga

  G  W  E  G  M  I  A  G  W  H  G  Y  T  S  H  G  A  H  G  V  A  V  A  A  D
ggatgggaaggaatgattgcaggctggcacggatacacatctcacggagcacatggagtggcagtggcggcggac
ggttgggaaggaatgattgctggttggcacggttatacttctcatggtgctcacggtgttgctgttgctgctgat

  L  K  S  T  Q  E  A  I  N  K  I  T  K  N  L  N  S  L  S  E  L  E  V  K  N
cttaagagtacgcaagaagctataaacaagataacaaaaaatctcaattctttgagtgagctagaagtaaagaat
cttaagtctacttaagaagctattaacaagattactaagaatcttaattctctttctgaacttgaagttaaaaat

  L  Q  R  L  S  G  A  M  D  E  L  H  N  E  I  L  E  L  D  E  K  V  D  D  L
cttcaaagactaagtggtgccatggatgaactccacaacgaaatactcgagctggatgagaaagtggatgatctc
ctttaaagacttttctggtgctatggatgaacttcacaatgaaattcttgaacttgatgaaaaagttgatgatctt

  R  A  D  T  I  S  S  Q  I  E  L  A  V  L  L  S  N  E  G  I  I  N  S  E  D
agagctgacactataagctcgcaaatagaacttgcagtcttgctttccaacgaaggaataataaacagtgaagat
agagctgatactatttcttcttaaattgaacttgctgttcttctttctaatgaaggtatcattaactctgaagat

  E  H  L  L  A  L  E  R  K  L  K  K  M  L  G  P  S  A  V  E  I  G  N  G  C
gagcatctattggcacttgagagaaaactaaagaaaatgctgggtccctctgctgtagagataggaaatggatgc
gaacaccttcttgctcttgaaagaaagcttaagaagatgcttggtccttctgctgttgaaattggtaatggttgc

  F  E  T  K  H  K  C  N  Q  T  C  L  D  R  I  A  A  G  T  F  N  A  G  E  F
ttcgaaaccaaacacaagtgcaaccagacctgcttagacaggatagctgctggcacctttaatgcaggagaattt
ttcgaaactaagcacaagtgcaattaaacttgccttgatagaattgctgctggtactttcaatgctggtgaattc

  S  L  P  T  F  D  S  L  N  I  T  A  A  S  L  N  D  D  G  L  D  N  H  T  Q
tctctccccacttttgattcactgaacattactgctgcatctttaaatgatgatggattggataaccatactata
tctcttcctactttcgattctcttaacattactgctgcttctcttaatgatgatggtcttgataatcacacttga

  R  S
------
agatct
```

The upper line shows the non-optimized original sequence of the hemagglutinin of the Influenza B strain Florida/4/2006. The lower line shows the codon-optimized sequence (lower case letters). Differences are highlighted in yellow. The ER-leader sequence is depicted in blue letters, and the transmembrane domain in green letters as amino acid sequence (in capital letters). The added EcoRV restriction site (gatatc) and the BglII restriction site (agatct) are shown in red letters. The stop codon is highlighted in red, and the start codon is highlighted in purple.

8.6 Sequence of the Grl3 gene with 5´and 3´flanking region

aaggtgctgctatgggagagaaaatatgttttttaatcaaataaagaatcgttttaattttgctgccttttat
ttcttcatctaaaattttgaatgcataaacattctggtaacatctaatgcaaaatcaataaatattatttaatgc
tcataaaatttcattcatttttatcaagtgattgtatgcattattaagtatttattttatatttattattaacaat
ataaggcaatcatttttattttagcttttaaacataatctttctccaatataacgtttaaattgttgaagaggtat
ggaaatttaagtaggtatccttaaacaaagaacatatttcatgctcataaatttttatcacagtttagtattttc
atctttgtaataacaacaatacctaccatttttgtatttaaaagcatatccgggttaaggggtaatttgaatcca
tttttgaaaaacaaatttatatttgaagttttgagtgctcctaagtgcgttattgaatatttttttggacatgat
aaaatgggatgatgctattgagagggcttcaattataataatgaaagataaaagcgtcaattatctaaatagaat
aagattcgtttaagaaaatttgattaaggaaaatgaatggaagcatacaaaatacaaggaattagatttctacgt
tttaacttaatatctgtttcggaaaacggctaaaaaacaaagaggtttaatagaataggatttttgttatcaac
tggctttctcccccgactttccttttgtttttctttattaattaaagcaatcttcattatttatttcaaaagag
aatgaaaaaaataaattaattaataagcaaaaactcatttttttatataaacaaaaaagaaaaataaattaaaaaa
aacgaaaaaaaaaaacaaaagaaaagtaatattcgattccttgcgatgaagaatttagctatcgttttggctact
ttgtgcatttttgctcaagccacaagtgtcttcgagacccctgctttccttgaagtaagctatttcttaatctaa
taaagagatcagaccaaatatgagataaagcattttatcttcccacctactttcactttctcctcacctgataat
ttaattattaataacttctcatggagcaacgagaagttacaactaagttaaatttgatgtttaatcaatcgaata
cacaaaaggagtatcgcgttgcataaaaaaagaaatatatggaggatacagctcttcacattcaaaaattttttt
atttaagatttccactcttacccccaaagaatataaaaagaagatcgcttatcctcttaagattaaacaaacaaa
caaacaaacatacagaatgaggttttctataatttaagtgtgaaaagaaatgcgaaaagtaaaaacaactaatca
aataaacgtcaatcaaaaccaactgcatttaattgaactagctttaattttgctaaattaaagatgattcattca
ttaatccctcagttttttaacttctttaattttacaattaaagctttttattcttttcacataaacgttttgg
aaatcaatcaaattgcaaccaacaaactctttccttaatggtatatcaggtgcttgtttgagatttgttttaagc
taattgaacctttattcatctagcctttaacacagaatatacccaccatattatttttttaagctcactaattt
attgaaactctgaagaaggtgtttattcatttaaaataaatacaaccataaagttgaatgaatatgctttaatta
attaattaattttcttctgtctgtctgactctcttttttaaaatatttaaaatgtttaatacttttattaacatta
aatttttttattaatagattaagagcaacccccttcggtcacaccgttgcttctttggtttaattaaacttaggtg
ctggataaagtgctggtagattagatgctattgctgaagccttgaacactattgaagcttaactcgaaaatacca
gagatcacaatgatgctgaaatccaaagacaaagaggctggtgcagtgatcaagaagctaccatctaagctaaca
ttgattaagctgaaagtgatttgagcaactactaaaacgaataaacttaaagaaactaagccgttgctgacctca
cccaaaacttaaacaacgaataataatccctcgctgaaaactaaaataacttagccaacgcctaataggaattag
atgatgaaaactcttcttatgctgaatcctctaaggactacgctgatgctattgctgcttgtgaataagctctca
agctcttagccaccttgtaaactaacccctccggtttcatctaatctaaggctagattcggtaacgttgttactc
tcctccaaaagcaccttgccaacaagtcttccaacttcgttcaacccatcttgaacgtcttgactgaaatggcta
gctccaccaacgaagttgactaatccagtcttgctaaggttgtttctctcattaacgatttattagaagaattaa
gaaactaaagtgctgctaatgactaaagacactaataagttgttgatagcctcacctccaacattgcaaatctcg
aacaacttattgataattctaacaaccttatttcataatactaaggttaaatctaagaaaatgaagacagattag
cctaacttgctggtcttatccaacaaacctaagctatccttgactaagccaactaaaatcttagctaagtttaag
attaatgcgctggatacgatcaagaatatgcttctttcaagaacgaagtgtaattaatctatttattcattcatt
tctccatattttcatttatttattttactcctacttataatgaataatttatttaattatatttatattaatagt

gatcaataactcgctaccttataagccctcaaggaatatttcaagagcaaggttgttcctgctgttgaaaacatt
aacgattctgcttatgctgatgaattaactgttgatgtttgatttctctaatatacaagcaaaatttaaggttat
tcctataatcgatattacttcaaatagctttgagtatttatgatttgaagaagaatattaaaatatatgtttacg
aaacaacaattcatatcatggtctatgttgtacttctttcttacttttttatgaatcaatatcattacttttttac
aaatcatatctttttaaattccattaatatttgtttgtctattcgaaatttgataattaatacattaaatttcat
acatatatataaacacacaatcattcatttttgatttattaattcattcatttttttgtttatttatttaaaaa
tatattcattcacatttttgaaattcatttaattaaaatgcaaaattaaatcccaaaatcttttatctttttat
taataaaataaattactcaaacagttaagccaaaatttttatcaatagataatttagttaataaaacaataattg
aatacaaaatgtattttaatttttttgggta

The 5´flanking region of the Grl3 gene is shown in green, the 3´flanking region in red. The Grl3 gene is shown in black letters.

I, the undersigned, hereby declare that this dissertation entitled, "*Tetrahymena thermophila*: An expression platform for the production of viral antigens." is my own work, and that all the sources I have used or quoted have been indicated or acknowledged by means of completed references.

.....

Christine Sachse