PRODUCTION AND CHARACTERIZATION OF HUMAN CARBONIC ANHYDRASE IX IN INSECT CELLS

Master´s thesis

BioMediTech, University of Tampere

December 2014

Mikaela Lindfors-Huttunen

PRO GRADU –TUTKIELMA

Paikka:	TAMPEREEN YLIOPISTO, BioMediTech
Tekijä:	LINDFORS-HUTTUNEN, MIKAELA HENRIETTA
Otsikko:	Hiilihappoanhydraasi IX:n tuotto ja karakterisointi hyönteissoluissa
Sivumäärä:	63
Ohjaajat:	Professori Seppo Parkkila, dosentti Mika Hilvo
Tarkastajat:	Professori Seppo Parkkila, professori Markku Kulomaa
Aika:	Joulukuu 2014

Tiivistelmä

Hiilihappoanhydraasi IX (CA IX) on solukalvoentsyymi, joka katalysoi hiilidioksidin hydraatiota. Se auttaa ylläpitämään syöpäsolujen kasvulle ja elinvoimaisuudelle edullisia pH-olosuhteita. CA IX on merkkiproteiini hypoksialle, minkä taas on huomattu olevan huonon selviytymisennusteen indikaattori tietyissä syöpätyypeissä. Vaikka CA IX:ää on analysoitu intensiivisesti viime vuosikymmenien aikana, niin osa sen biokemiallisista ominaisuuksista oli vielä selvittämättä silloin, kun tämä tutkimus tehtiin. Tästä syystä tämän työn tarkoituksena oli tuottaa ja puhdistaa CA IX -rekombinanttiproteiinineja hyönteissoluissa käyttäen bakuloviruksekspressiosysteemiä. Tämä oli ensimmäisiä tutkimuksia, joissa hiilihappoanhydraaseja tuotettiin hyönteissoluissa. Tarkoitus oli käyttää tuotettuja CA IX -proteiineja aktiivisuusmäärityksiin ja muihin tutkimuksiin biokemiallisten ominaisuuksien määrittämiseksi. CA IX:n katalyyttinen domeeni ja proteoglykaanidomeeni tuotettiin erikseen, puhdistettiin ja analysoitiin. Lisäksi tutkittiin myös CA IX:ää, jossa oli kummatkin domeenit samassa rekombinanttiproteiinissa. Käytetyn hyönteissolutuottosysteemin todettiin olevan toimiva tapa tuottaa CA IX:ää. Havaittiin, että proteoglykaanidomeeni nostaa CA IX entsyymin aktiivisuutta. Lisäksi aktiivisuusarvot vahvistivat aikaisemmat tutkimukset, joiden perusteella CA IX:n katalyyttisen domeenin aktiivisuus on keskitasoa. Kuitenkin havaittiin, että $ZnCl_2$ lisäysreaktioon lisäsi huomattavasti CA IX:n katalyyttistä aktiivisuutta ja sen todettiin olevan CA IX:n erityisominaisuus. Aktiivisuusarvot, jotka mitattiin kokopitkälle CA IX:n muodolle $ZnCl_2$:n läsnäollessa, olivat korkeimmat, mitä oli koskaan mitattu hiilihappoanhydraaseilla.

MASTER´S THESIS

Place:	UNIVERSITY OF TAMPERE, BioMediTech
Author:	LINDFORS-HUTTUNEN, MIKAELA HENRIETTA
Title:	Production and characterization of human carbonic anhydrase IX in insect cells
Pages:	63
Supervisors:	Professor Seppo Parkkila, Docent Mika Hilvo
Reviewers:	Professor Seppo Parkkila, Professor Markku Kulomaa
Time:	December 2014

Abstract

Carbonic anhydrase IX (CA IX) is a cell surface enzyme, which catalyzes CO_2 hydration. It has a function to maintain favorable pH conditions for cancer cell growth and survival. Expression of CA IX is a marker of hypoxia, which has been noticed to be an indicator for poor prognosis in specific cancers. CA IX has been analyzed intensively during the last decades but some of its biochemical properties were not revealed until this study was conducted. Hence, the recombinant protein forms of CA IX were produced and purified in this study using the baculovirus insect cell expression system. The aim was to produce CA IX recombinant proteins for further studies to analyze the activity and other biochemical properties of these enzymes. The recombinant proteins, which had either the catalytic domain or the proteoglycan domain of CA IX were separately produced, purified and analyzed. Additionally a CA IX form having both domains in the same recombinant protein was also studied. The used baculovirus insect cell expression system was found to be a feasible way to produce recombinant CA IX. It was found that the proteoglycan domain has a notable influence on enzyme activity of CA IX. In addition, activity values confirmed the previously reported results that the catalytic domain of CA IX has average activity as compared to the other CA isozymes. However, the addition of $ZnCl_2$ to the reaction was found to increase the catalytic activity of CA IX, and this property was concluded to be a unique feature for CA IX. In the presence of $ZnCl_2$ the catalytic activity of the full-length recombinant protein having both the catalytic and proteoglycan domains was higher than measured before for any other CAs.

Acknowledgements

My sincere gratitude to the head of the Tissue biology group and the supervisor for the thesis prof. Seppo Parkkila, for accepting and providing me with the opportunity to work on such an intresting topic. The whole learning experience with working in the group was instructive. I am also very gratefull for docent Mika Hilvo for long-term and accommodating supervision during the study for this master´s thesis. I would also like to thank Aulikki Lehmus and Jukka Lehtonen for helping and teaching me in practical issues. I would also like to thank other group members during that time Alise Hyrskyluoto, Peiwen Pan, Heini Kallio, Alejandra Rodriguez Martinez and Anna-Kaisa Parkkila as well as late Henri Nordlund from Markku Kulomaa´s research group, and Claudiu T. Supuran and Alessio Innocenti from the University of Florence.

Contents

Abbreviations

AE	anion exchanger
AcMNPV	autographa californica multiplenucleopolyhedrovirus
α-CA	α-carbonic anhydrase
BEVS	baculovirus expression vector system
CA	catalytic domain
CA I	carbonic anhydrase I
CA II	carbonic anhydrase II
CA IV	carbonic anhydrase IV
CA IX	carbonic anhydrase IX
CA9	carbonic anhydrase 9
CA XII	carbonic anhydrase XII
CA XIV	carbonic anhydrase XIV
CA XV	carbonic anhydrase XV
CHO	chinese hamster ovary cells
CpG sites	regions of DNA where a cytosine nucleotide occurs next to a guanine nuccleotide in the linear sequence
EDTA	ethylene diamine tetra acetic acid
G250	antibody
HeLa	cervical cancer cells
HIF	hypoxia-inducible factor
HRE	hypoxia-response element
IC	intra cytoplasmic tail
IE	immediate early promoters
IPLB-SF-21	spodoptera frugiperda cell line
IPTG	isopropyl β-d-1-thiogalactopyranoside
lack	structural gene of lac operon
LB	Luria-Bertani plate
MNPV	multicapsid nuclear polyhedrosis viruses
M75	antibody
PET	positron-emission tomography
PG	proteoglycan-like domain
PI3K	phosphatidylinositol 3-kinase
PVDF	polyvinydine fluoride membrane
PR	protected regions
RT-PCR	real-time polymer chain reaction
RMCE	recombinase-mediated cassette exchange
SES-PCR	stepwise elongation of sequence-PCR
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
Sf9	spodoptera frugiperda 9
SP	signal peptide
SP1/3	transcription factor
S2	drosofila Schneider cells
TBST	mixture of tris-buffered saline and tween 20
TM	transmembrane region
Tn7	transposon
VHL	von Hippel-Lindau protein
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

1 Introduction

Carbonic anhydrase IX (CA IX) is a cell surface enzyme, which belongs to a larger α-CA gene family and is strongly associated to malignant tumors (Pastorekova, Zavada 2004, Syrjanen, Luukkaala et al. 2014). Expression of CA IX is a marker of hypoxia, which has been noticed to be an indicator for poor prognosis in specific cancers (Potter, Harris 2003, McDonald, Winum et al. 2012). The function of CA IX has been linked to maintain favorable pH conditions for cancer cell growth and survival. It creates acidification at cell surface and disrupts cell-to-cell interactions increasing tumor propagation (Svastova, Zilka et al. 2003, Swietach, Hulikova et al. 2010). Due to these properties and because the expression of CA IX is mainly limited to tumor cells (Pastorekova, Zavada 2004), CA IX is an important target molecule in cancer therapy and diagnostics (Pastorekova, Parkkila et al. 2006, McDonald, Winum et al. 2012). Clinical trials in diagnostic purposes have yielded positive results and additionally promising results have been reported in cancer therapy studies conducted with monoclonal antibodies and small molecule inhibitors (McDonald, Winum et al. 2012). In order to analyze the properties and functions of CA IX in detail, it was necessary to produce CA IX as a recombinant protein for these studies.

Recombinant proteins can be produced by various expression systems, such as chemical synthesis or bacterial, fungal, insect and mammalian cells (Kollewe, Vilcinskas 2013). Bacteria are usually the first option considered when the longer peptides or complete proteins are aimed to be produced. Bacterial cells are used because high amounts of recombinant proteins can be generated in relatively short time and little effort (Structural Genomics, China Structural Genomics et al. 2008). However, in this study a eukaryotic expression system was chosen for recombinant CA IX production. This is because bacterial cells (*E. Coli*) had been already used in CA IX recombinant protein production and because CA IX has post-translational modifications, which are not produced correctly in bacterial cells.

Insect cell expression systems represent a rational compromise between the bacterial and the mammalian expression systems (Becker-Pauly, Stöcker 2011). In insect cells, signal peptides are cleaved as in mammalian cells, disulfide bonds can be formed and

proteolytic processing is possible because of protein-converting enzymes (Jacobs, Callewaert 2009). In addition, larger amounts of the recombinant proteins are possible to be produced especially when insect cells are cultured in suspension cultures (Jarvis 2009). The aim was to produce high amounts of recombinant proteins in this study because of aimed crystallization experiments and other studies.

Even though CA IX has been studied widely, some of its biochemical properties were still unknown at the time when this study was conducted. At that time recombinant CA protein production had not been tested in insect cells. Therefore, the aim of this study was to produce and purify recombinant proteins, which could be used in activity studies and in further analyses, such as inhibitory and oligomerisation studies as well as crystallization experiments to define the 3D structure of CA IX.

2 Review of the literature

2.1 α-Carbonic anhydrases

Carbonic anhydrases are ubiquitous metallo-enzymes in the nature, which usually contain a zinc atom as a metal cofactor. They are found in animals, plants and also in microbes. CAs exist in different forms, with different structures, in different molecular weights and also their activities vary. These enzymes are involved in basic cellular processes, such as photosynthesis, respiration, transportation of ions, calcification, and regulation of acid–base balance. Additionally, CAs are the fastest known enzymes. There are five different CA families (α , β , γ , δ and ζ) (Kanth, Pack 2013). Human CA IX, which is the target enzyme in this study, belongs to the α-CA gene family (Pastorekova, Zavada 2004).

α-carbonic anhydrases participate in several functions in human metabolism, such as respiration, bone resorption, production of gastric acid and different body fluids, renal and testicular acidification, and other biological processes (Pastorekova, Zavada 2004). In the regulation of acid base balance, α-CAs use zinc-activated hydroxide mechanism to catalyze the conversion of carbon dioxide and water to bicarbonate and proton in a reversible reaction ($CO_2 + H_2O \leftrightarrow HCO_3^- + H^+$) (Supuran, Scozzafava 2007). α-CAs are expressed in several locations in the cell. CA I-III, CA VII and CA XIII are expressed in the cytosol. CA IV, CA IX, CA XII and CA XIV are located on the cell surface. CA VA and CA VB are proteins found in mitochondria and CA VI is a secreted protein. All these CA isozymes are catalytically active in mammals. The decreased activity of certain isozymes has been associated with several diseases including glaucoma, osteopetrosis, heart oedema, renal failure, and neurological and neuromuscular disorders, etc. (Pastorekova, Zavada 2004). The connection with cancer was, however, not noticed before the discovery of CA IX (Pastorekova, Zavadova et al. 1992).

2.2 Carbonic anhydrase IX

2.2.1 History

CA IX was first found in HeLa-cells and at that time it was named as MN-protein (Pastorekova, Zavadova et al. 1992). Later, when the cDNA of the MN-protein was managed to be cloned, it was noticed that the MN-protein had strong homology with α-

CA family (Pastorek, Pastorekova et al. 1994). Because of this strong homology and enzyme activity, MN-protein was named to CA IX. However, Oosterwijk *et al* had noticed even earlier that the antibody G250 recognized an antigen in kidney cancer cells (Oosterwijk, Ruiter et al. 1986). Years later, this particular antigen was identified to be CA IX (Grabmaier, Vissers et al. 2000).

2.2.2 Molecular properties

2.2.2.1 Gene transcription

CA IX is located in chromosome 9p12-p13 (Nakagawa, Uemura et al. 1998). The coding sequence of the gene is split to 11 exons and the size of the protein is 10.9 kb (Opavsky, Pastorekova et al. 1996). The mRNA is transcribed as a size of 1,5kb and on reducing conditions SDS-PAGE, CA IX migrates as 54kDa and 58 kDa double bands. (Pastorek, Pastorekova et al. 1994).

There are no found mutations in the *CA9* cDNA, which could possible cause different functions between normal and cancer cells (Pastorekova, Parkkila et al. 1997). However, by using RT-PCR (real-time polymerase chain reaction), an alternative mRNA splicing variant of *CA9*, that does not include exons 8-9, has been found. This alternative form of the mRNA is expressed in very low concentrations in several different cell lines, and it is coding a shorter and less active form of the mRNA of the CA IX protein. The physiological role of this alternatively spliced variant has not yet been identified, but it is indicated that the splicing variant may functionally intervene with the full-length CA IX. This could be relevant especially under the conditions of a mild hypoxia (Barathova, Takacova et al. 2008).

2.2.2.2 Structure

CA IX is a glycoprotein that is located at the cell surface. CA IX migrates as a 153-kDa band on non-reducing conditions in SDS-PAGE, which was the reason why CA IX was assumed to be a trimeric protein (Pastorekova, Zavadova et al. 1992). This trimeric structure is stabilized by the formation of intermolecular disulfide bonds mediated by cysteine residues. There are three histidine residues on the bottom of the catalytic center of CA IX and they bind a one zinc ion. CA IX has a signal peptide, a catalytic domain, a proteoglycan-like domain, a transmembrane region, and a cytoplasmic tail

(Pastorekova, Zavada 2004). The organization of the domains of CA IX is demonstrated in a Figure 1.

Figure 1. *The domain organization of CA IX: SP (a signal peptide), PG (a proteoglycan-like domain), CA (a catalytic domain), TM (a transmembrane region) and IC (a cytoplasmic tail). The figure is modified from original reference* (Pastorekova, Zavada 2004)

2.2.3 Expression and connection to prognosis

CA IX is expressed only in a few normal tissues. It has been identified for example, in gastric mucosa and gall bladder by immunohistochemical staining (Pastorekova, Parkkila et al. 1997). Smaller amounts have also been found in intestinal epithelia, pancreatic duct, and testis (Ivanov, Liao et al. 2001).

A special characteristic property of the CA IX expression is that it is overexpressed especially in cancer tissues, whereas its expression in normal tissues is usually very low. In addition, the expression of CA IX in malignant tissues is connected to poor prognosis (Potter, Harris 2003, Maseide, Kandel et al. 2004, Driessen, Landuyt et al. 2006, McDonald, Winum et al. 2012). According to the literature review conducted by Christian Potter and Adrian L Harris, the expression of CA IX correlated with poor prognosis in breast, lung, cervical, head and a neck cancer (Potter, Harris 2003). Also in stomach and oesophagus cancer, the expression of CA IX has been linked to poor prognosis (Driessen, Landuyt 2006). Additionally, tissue microarray studies have supported the correlation between the expression of CA IX and poor survival in breast, lung, ovarian and bladder cancers as well as in astrocytomas (McDonald, Winum et al. 2012).

CA IX expression does not, however, always correlate with poor prognosis. According to Nakamura *et al,* the expression of hypoxia inducible factor-1 (HIF-1α) correlated significantly with poor survival of stomach cancer patients, whereas correlation with the CA IX expression was not found (Nakamura, Kitajima et al. 2011). The weak expression of CA IX in kidney cancer was connected to poor survival and strong expression was connected to better prognosis (Bui, Visapaa et al. 2004). Also

according to the literature review made by Stillebroer *et al,,* the strong homogenic expression of CA IX in kidney cancer is related to a better response to systemic treatment and survival (Stillebroer, Mulders et al. 2010).

2.2.4 Regulation of gene expression

2.2.4.1 Effect of hypoxia and acidosis on regulation of CA IX expression

Hypoxia is particularly related to cancer tissues. It induces transcription factors, which further activate several different genes. These activations will lead to a cell response that usually causes growth and proliferation of the cancer cells. For this reason, hypoxia has an important role in tumorigenesis, tumor development and metastasis (Brahimi-Horn, Pouyssegur 2007).

CA IX is usually expressed strongly in cancer tissues in hypoxic conditions. In a hypoxic environment, a hypoxia responsive element (HRE) has an effect on the expression of CA IX through HIF-pathway (Wykoff, Beasley et al. 2000). The HRE activates gene transcription of CA IX in the hypoxic environment. The HRE is activated by a transcription factor HIF1 (hypoxia inducible factor-1) that consists of two sub-units, HIF1 α and HIF1 β. HIF1 α stabilizes in hypoxic conditions, and after that it drifts into a nucleus. In the nucleus, HIF1 β is expressed constantly and after HIF1 α has drifted into the nucleus HIF1 α and HIF1 β together will form an active transcription factor HIF1 (Pastorekova, Ratcliffe et al. 2008).

In normal conditions, the HRE-gene expression is prevented when a von Hippel Linday (VHL) tumor suppressor ubiquitin-protein binds to the hydroxylated HIF1 α sub-unit and forms an ubiquitin system. This will lead to fragmentation of the ubiquitin complex and consequently, drifting of the HIF1 α into the nucleus will be prevented (Semenza 2001). The loss of function in VHL in renal cell carcinomas and cerebellar hemangioblastomas cancer causes constant HIF1 α -activation, which finally leads to the overexpression of CA IX (Semenza 2001, Supuran 2008).

Protected regions PR1 - PR5 affect also the regulation of CA IX expression. In the circumstances when HIF1 is active, the SP1/SP3 transcription factors bind on PR1- and PR5-protected regions. This phenomenon has been found to be critical for the CA IX

transcription. This influence for CA IX transcription has been indicated by using SP1 inhibitor (Mithramycin A), which has been noticed to prevent expression of CA IX (Kaluzova, Pastorekova et al. 2001, Kaluz, Kaluzova et al. 2003). Activation protein 1 has also an important role in activation of CA IX transcription and it binds to the position PR2 (Laderoute 2005). The PR3 region binds also proteins that have been noticed to have positive effect on activation of CA IX promoter, but the functions of these proteins are still unknown. Region of PR4 have been conversely perceived to act as a silencer in transcription of CA IX (Kaluz, Kaluzova et al. 1999).

Acidosis, which is prevalent in a microenvironment of high cell density tissues, has been noticed to have an effect to regulation of CA IX expression. Acidosis as well as PI-3K (phosphatidylinositol 3-kinase) has a stabilizing influence on HIF-1α, which enhances drifting of the HIF-1α into the nucleus (Kaluz, Kaluzova et al. 2009). Regulation of CA IX expression is visualized in Figure 2.

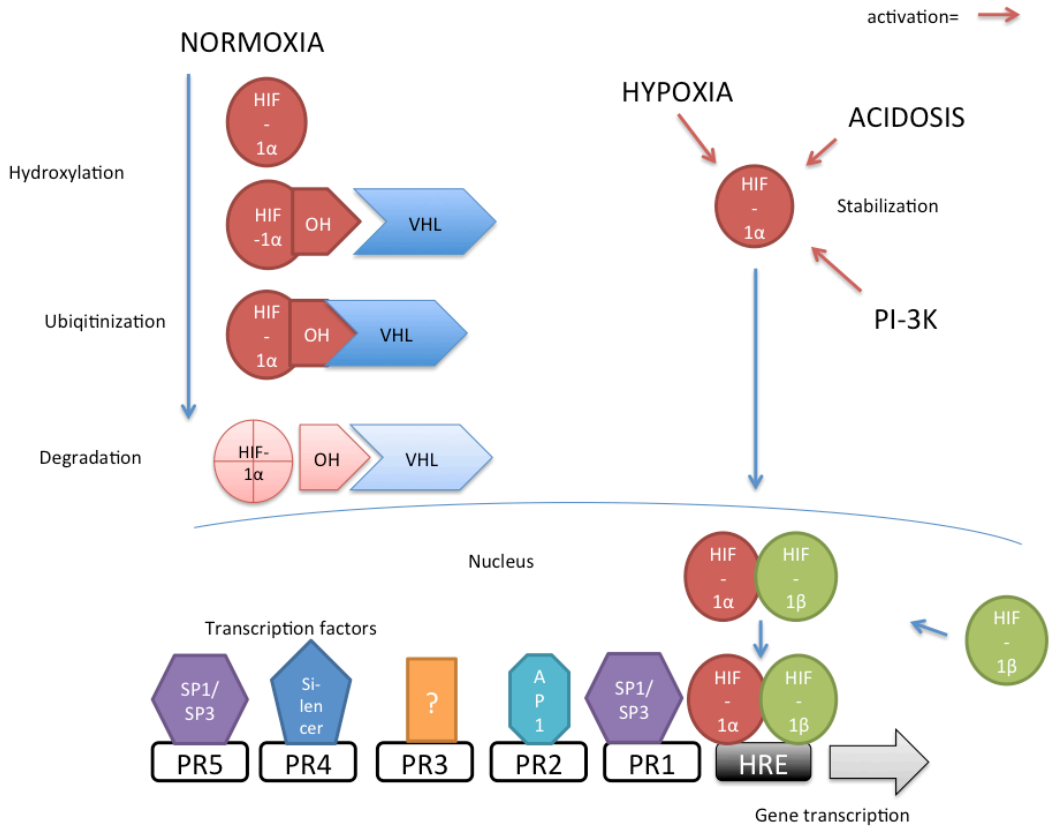


Figure 2. *Regulation of CA IX expression. The figure is modified from original references* (Kaluz, Kaluzova et al. 2009, De Simone, Supuran 2010)

2.2.4.2 Effects of methylation status on regulation of the CA IX expression

In the region of *CA9* gene promoter, methylation of CpG sites has been detected to have an important function in the regulation of gene expression. Especially, methylation status at -74 and -6 CpG sites have been correlated with CA IX expression in kidney cancer cells. Ashida *et al* noticed that all studied CA IX positive cell lines were hypo-methylated (decreased methylation) and all CA IX negative cell lines were hyper-methylated (increased methylation) (Ashida, Nishimori et al. 2002).

Also Grabmaier *et al* showed that CA IX expression and hypo-methylation of the CA IX promoter correlates in the kidney cancer. In addition, they noticed that hypo-methylation did not occur in primary kidney cancer cells or in normal kidney tissues. Additionally they identified that re-cultured primary kidney cancer cells had increased hypo methylation status in a CA IX gene. Because hypo-methylation was only detected in cultured kidney cells, except primary kidney cancer cells, Grabmaier *et al* assumed that hypo methylation of CA IX promoter was induced by *in vitro*-conditions in tissue cultures (Grabmaier, de Weijert et al. 2002).

Jakubickova *et al* investigated the methylation of the CA IX promoter using HeLa cervical carcinoma cells. The study showed that the methylation of CA IX promoter decreased the expression of CA IX in dense cell cultures but it did not occur in hypoxia or in sparse culturing conditions. Therefore it was concluded that the methylation of CA IX promoter may not be an essential factor in HeLa cervical carcinoma cells in gene regulation of CA IX, but it could be an additional mechanism in the regulation of CA IX promoter in dense cell cultures. (Jakubickova, Biesova et al. 2005).

As mentioned before, hypoxic environment has been detected to have a strong influence on the CA IX expression. However, CA IX expression and hypoxia are not always correlating and the regulation mechanisms of CA IX are still partially unknown. For example, Nakamura *et al* have showed that the regulation of CA IX is more depended on methylation status than the stimulation caused by hypoxia. According to immunohistochemical studies, the epigenetic changes in CA IX differ in a diffuse-type stomach cancer and an intestine cancer. In the diffuse-type stomach cancer, the increased methylation rates have been identified and in the intestine cancer the decreased methylation rates have been observed (Nakamura, Kitajima et al. 2011).

2.2.5 Function

2.2.5.1 Catalytic activity

Similarly to other α-CAs, one function of CA IX is to act in the zinc-activated hydroxide mechanism to catalyze carbon dioxide in a reversible reaction, forming hydrogen carbonate. (Supuran, Scozzafava 2007). In the active site of CA IX there is a zinc ion coordinated by three histidine residues, which has important function for the catalytic activity of CA IX (Pastorekova, Zavada 2004).

The hydroxide ion is bound to the zinc ion on the active (basic) form of an enzyme. When the carbon dioxide bounds to a hydrophobic pocket, it will react in a reversible reaction with a zink coordinated hydroxide ion, forming a bicarbonate-ion. When a water molecule replaces the bicarbonate-ion, the result will be the formation of an inactive (acidic) form of the enzyme. In order to transform back to the catalytically active form, a proton transfer reaction from the active site to the surroundings will occur. This reaction is essential for the reaction rate, and it is presumably assisted by histidine residue of the active site (Supuran, Scozzafava 2007). These catalytic reactions have been presented in reaction formulas a) and b). In the reaction formula a) and b), E refers to enzyme (Supuran, Scozzafava 2007).

a) $EZn^{2+}\text{-}\ OH^{-} + CO_2 \Leftrightarrow EZn^{2+}\text{-}\ HCO_3^{-} \Leftrightarrow {}^{H}{}_{2}{}^{O}\ EZn^{2+}\text{-}\ OH_2 + HCO_3^{-}$

b) $EZn^{2+}\text{ -}\ OH_2 \Leftrightarrow EZn^{2+}\text{ -}\ HO^{-} + H^{+}$

2.2.5.2 Effect on pH control

CA IX is usually expressed in highly differentiated and metabolically active cells, such as cancer cells. Therefore, CA IX has been suggested to have an essential physiological function in the regulation of pH-balance in tumor tissues (Swietach, Hulikova et al. 2010). CA IX actively increases the acidification of the cell surface and it neutralizes the intra cellular pH, protecting cancer cells from hypoxia and acidosis (Swietach, Hulikova et al. 2010, Svastova, Hulikova et al. 2004).

CA IX is assumed to co-operate with anion exchangers in the pH-balance regulation. When CA IX converts carbon dioxide to hydrogen carbonate, anion exchangers in the

cell surface will use the hydrogen carbonate. In that case the hydrogen carbonate is transported through these anion exchangers to the cytosol and in the cytosol CA II will transform the hydrogen carbonate back to carbon dioxide, which then diffuses out of the cell. (Pastorekova, Zavada 2004).

Morgan *et al* have also studied the interactions of CA IX with anion exchangers. They assumed that Cl^-/HCO^-_3 -exchangers (anion exchangers AE1, AE2 ja AE3) interact physically as well as functionally with CA IX. The assumption was that CA IX and anion exchangers could form together a bicarbonate transport metabolon system (Morgan, Pastorekova et al. 2007). Even though this direct binding of CA IX and anion exchangers has been questioned, co-operation of CA IX with anion exchangers is considered to be possible (Swietach, Hulikova et al. 2010, Boron 2010). Also the study by Svastova *et al* supports an assumption of the interaction of CA IX with anion exchangers and as well as the formation of the metabolon system of CA IX with anion exchangers (Svastova, Witarski et al. 2012).

The glucose metabolism of healthy normal cells usually uses a mitochondrial pathway, whereas cancer cells prefer a glycolytic route, which leads to high lactic acid concentrations (Brahimi-Horn, Pouyssegur 2007). As mentioned previously, CA IX has been noticed to function actively, neutralizing the pH of the cytosol in these acidic conditions. Swietach *et al* have suggested that CA IX may use three different mechanisms when participating in the pH balance regulation. In oxidation conditions, catalysis may occur using a mitochondrial model and in glycolytic environment (cancer cells), depending on expression and activity of the anion exchangers, using either an HCO^-_3 titration model or an H^+ extrusion model (Swietach, Hulikova et al. 2010). CA IX catalyzes carbon dioxide to hydrogen carbonate in the mitochondrial model and in the titration model. In these cases, extra cellular pH will decrease and intra cellular pH will increase. Instead, in the extrusion model the extra cellular pH will increase and the intra cellular will decrease. The HCO^-_3 titration model corresponds the hypoxic environment of the tumor cells when dehydration of CA IX will neutralize intra-cellular pH, enhancing tumor growth. (Swietach, Hulikova et al. 2010, Svastova, Hulikova et al. 2004). Svastova *et al* have also showed experimentally how the protons of CA IX may

cause acidosis in hypoxic conditions in the microenvironment of tumors. (Svastova, Hulikova et al. 2004).

2.2.5.3 Effect on adhesion, migration and proliferation of the cells

CA IX has a proteoglycan-like domain, which indicates that it may be related to cell adhesion (Zavada, Zavadova et al. 2000). Typical properties of cancer tissues include strong cell division and migration. It has been noticed that CA IX increases the de-adhesion of cells and thus induces tumor progression (Svastova, Zilka et al. 2003).

Svastova *et al* have indicated that CA IX may negatively affect E-cadherin-mediated cell adhesion over β-catenin. The E-cadherin is an important adhesion molecule whose disappearance or instability is linked to tumor progression. Because of this connection, the disruptive effect of CA IX to E-cadherin-mediated cell-adhesion and strong expression of CA IX in the malignant tissues can be assumed to be connected to each other (Svastova, Zilka et al. 2003).

According to Hulikova *et al,* also the intracellular tail has been detected to have an influence on the de-adhesion of CA IX expressing cells. Hulikova *et al* found that mutations in the basic amino acids of the intracellular tail did not disrupt the localization of CA IX on cell membrane, but decreased CA IX-mediated de-adhesion of the cells (Hulikova, Zatovicova et al. 2009).

2.2.6 Clinical significance

2.2.6.1 In diagnosis

CA IX is a potential marker of cancer, because it is very strongly expressed in cancer tissues, but the expression in normal tissues is low (Potter, Harris 2003). Antibodies M75 and G250 recognize the CA IX protein. The M75 is a very specific antibody for the PG-domain of CA IX and it is generally used to identify CA IX in immunohistochemical studies. However, for the radiolabeling and for the immune therapy an alternative antibody, chimeric antibody cG250, has been developed (McDonald, Winum et al. 2012).

Divgin *et al* analyzed 25 patients who were suspected to have kidney cancer, using iodine-124-labelled antibody chimeric G250 Positron-Emission Tomography (124) I-cG250 immuno-PET). Sixteen of the studied patients were assured to have kidney cancer, and in fifteen of those patients, cancer was also identified in (124) I-cG250 PET-scanning. (124) I-cG250 binding was not detected in any of the patients who did not have kidney cancer. The study showed that (124) I-cG250 PET could be a feasible way to diagnose malignant tumors and a potential method in decreasing invasive surgeries (Divgi, Pandit-Taskar et al. 2007). The expression in kidney cancer has also been observed in a preclinical study (mouse model) using Immuno-Pet ((89) Zr- cG250 immuno-PET) where the signal was clearly positive in kidney tumors (Stillebroer, Mulders et al. 2010).

2.2.6.2 In cancer therapy

CA IX is a possible target also in the cancer therapy. One method is to use CA IX-specific monoclonal antibodies (mAbs) in immune therapy. It is possible to achieve therapeutic influence *via* several different mechanisms by using these monoclonal antibodies. For instance, a direct binding of these monoclonal antibodies may cause a toxic cell response, which could have an inhibitory effect on the tumor. Alternatively, antibodies can be used to direct various therapeutic molecules such as cytotoxines or radionuclides on the cell surface (McDonald, Winum et al. 2012).

Another method in CA IX-mediated cancer therapy is to use CA IX-specific small molecule inhibitors. Several different molecules and compounds are known to effectively inhibit CAs. Especially sulfonamides, sulfomates and coumarines have been noticed to have an inhibitory effect on malignant tumors. The sulfonamides inhibit the function of CA IX by binding to the zinc ion on the active site of CA IX, whereas coumarine and thiocoumarine bind to a gap of the active site, thus inhibiting the enzyme activity. Several studies have been conducted to analyze the effects of the inhibitors. Results have been mostly positive when studying the effects of inhibitors to the tumor progression (McDonald, Winum et al. 2012).

2.3 Recombinant protein production

2.3.1 Chemical synthesis

It is possible to produce short linear peptides using a purely chemical synthesis. Proteins produced by chemical synthesis are also easier to purify than those produced by other expression systems, due to the less complex production conditions. The chemical synthesis has, however, limitations, such as restricted amino acid length, and difficulties to produce peptides containing high proportion of challenging amino acids, such as arginine, cysteine, methionine and tryptophan. Also post-translational modifications are a challenging issue when using chemical synthesis (Kollewe, Vilcinskas 2013). Besides the challenges in the synthesis of peptides containing post-translational modifications, the peptides containing disulfide bonds have been synthesized (Reinwarth, Nasu et al. 2012).

2.3.2 Bacterial cells

Bacteria are usually the first option to be considered when longer peptides or complete proteins are produced, because high amounts of the recombinant proteins can be generated in relatively short time and little effort (Structural Genomics, China Structural Genomics et al. 2008). Most commonly used bacteria are gram-negative *Escherichia coli (E.coli)* and gram-positive *Bacillus subtilis (B.subtilis)*. *Bacillus subtilis* is mostly used in industrial applications (Westers, Westers et al. 2004, Rosano, Ceccarelli 2014). However, in spite of the extensive knowledge in the molecular biology of bacteria, the recombinant protein production in bacteria may also be challenging (Sivashanmugam, Murray et al. 2009). For instance, some issues may appear in the expression of the genes with unique structural features, translational efficiency, and stability of mRNA. Additional challenges may appear especially with protein folding and other post-translational modifications. There might also be problems with protein degradation by host cell proteases, and with protein being toxic to the host cells (Sivashanmugam, Murray et al. 2009). Some of these problems can be solved by suitable design of expression systems, such as by selecting a correct host strain, plasmid, and promoter. The ideal promoter for the expression of recombinant proteins in *E. coli* directs efficient transcription and protein production, which is tightly regulated to minimize metabolic burdens and toxic effects. In addition, the process of mRNA degradation provides a major control point of gene expression. Also codon optimization may be essential, if the

sequence of the produced protein contains non-typical codons for *E.coli* (Jana, Deb 2005).

If the structure or function of the recombinant protein depends on disulfide bonds, proteolytic cleavage, or other post-translational modifications, bacteria may not be the ideal production system. The targeting of recombinant proteins into the periplasmic space can encourage the formation of disulfide bonds, but amounts of the produced recombinant protein tends to be significantly lower in contrast to cytoplasmic expression (Mergulhao, Summers et al. 2005). Under these circumstances, a eukaryotic expression system could be a more functional option.

2.3.3 Yeast cells

Yeasts have been popular hosts for recombinant protein production in industrial applications, because these expression systems are economical, they can rapidly reach high cell densities and high protein concentrations, and they do not contain pathogens or viral inclusions (Celik, Calik 2012). In addition, yeast cells have an ability to perform better eukaryotic post-translational modifications when compared to bacteria (Mattanovich, Branduardi et al. 2012). However, proteins with antifungal activity are difficult to produce in yeast cells and the glycosylation in yeast differs from human cells. This may lead to hyperglycosylation, which may reduce enzyme activity or make the protein immunogenic (Jayaraj, Smooker 2009).

There are plenty of yeast expression hosts, such as *Saccharomyces cerevisiae*, *Pichia pastoris*, *Hansenula polymorpha*, *Kluyveromyces lactis*, *Schizosaccharomyces pombe*, *Yarrowia lipolytica,* and *Arxula adeninivorans* (Celik, Calik 2012). Two of the most used yeast strains are *S. cerevisiae* and the methylotrophic yeast *P. Pastoris* (Demain, Vaishnav 2009). *Saccharomyces cerevisiae* is the first and best characterized yeast expression system. A wide variety of recombinant proteins produced in *S. cerevisiae* have been listed in a review made by Demain and Vaishnav. The recombinant proteins manufactured in *S. cerevisiae* are, for instance, insulin, hepatitis B surface antigen, urate oxidase, glucagon, granulocyte macrophage colony stimulating factor, hirudin, and platelet-derived growth factor (Demain, Vaishnav 2009). However, numerous cases of plasmid instability, low protein yields, and the hyperglycosylation of the proteins have limited the use of *S. Cerevisiae* (Celik, Calik 2012).

An advantage of *P.Pastoris*, when compared to other yeasts, in manufacturing recombinant proteins is its ability to secrete proteins. Success has been achieved in genetically engineered *P. Pastoris* secretory pathway in order to produce human type N-glycosylated proteins (Choi, Bobrowicz et al. 2003). *P. Pastoris* has several advantages over *S. Cerevisiae* as a cloning host, such as higher protein productivity as well as lack of hyperglycosylation. It can be cultured in methanol solutions, and it is economical to set up and maintain. In addition, integration of multiple copies of foreign DNA into chromosomal DNA yields stable transformants. There are, however, some disadvantages when using *P. Pastoris* as a host for heterologous expression. A number of proteins require chaperones for proper folding, and *P. Pastoris* is unable to produce such proteins (Demain, Vaishnav 2009).

2.3.4 Insect cells

Insect cell expression systems are a considerable compromise between bacterial and mammalian expression systems (Becker-Pauly, Stöcker 2011). In insect cells, the signal peptides are cleaved as in mammalian cells, disulfide bonds are formed in the endoplasmic reticulum, and protein-converting enzymes are available for proteolytic processing (Jacobs, Callewaert 2009).

The insect cell lines used in recombinant protein production usually achieve higher cell densities than mammalian cells. Thus smaller culture volumes are needed. All in all, the insect cell cultures are less demanding than mammalian cells. With insect cell cultures, shake or spinner flasks can be used, and there is no need for a CO_2 atmosphere because insect cell culture media are buffered with phosphate (Moraes, Jorge et al. 2012, Jarvis 2009). However, the sterility of the cell culture conditions is equally important. In contrast to mammalian cell cultures, the biosafety level is not increased when a heterologous gene is introduced by baculovirus infection. Although the insect cell expression systems are an effective and functional way to produce recombinant proteins, the bacterial expression system may be a preferable option if the recombinant protein does not require post-translational modifications (Kollewe, Vilcinskas 2013).

There are several insect cell expression systems available, such as stably transfected drosophila cells, stably-transfected lepidopteran cell lines, site-specific gene integration

for stable expression, and baculovirus expression vector system (BEVS) (Kollewe, Vilcinskas 2013). The most commonly used vector system for recombinant protein expression in insect cells is the baculovirus (Demain, Vaishnav 2009).

2.3.5 Mammalian cells

Mammalian cells are well suited for demanding proteins when, for instance, glycan structures or other post-translational modifications are critical for the protein structure (Kim, Kim et al. 2012). Chinese hamster ovary cells (CHO) are particularly widely used for the production of recombinant proteins (Zhu 2012). The fast way to produce recombinant proteins in mammalian cells is transient transfection (Baldi, Hacker et al. 2007). Transient transfection is, however, more suitable for analytical experiments than in large-scale production, because it is not sufficiently productive. Because of the low yields of the recombinant proteins with transient transfection, especially in CHO cells, an alternative strategy is to use a stable transfection. The stable transfection allows the selection of individual high-yielding cells to increase overall productivity (Ye, Alvin et al. 2010). In addition, CHO cells have been also indicated as safe hosts, and thus it is easier to obtain approval to the market from regulatory agencies such as the FDA (Kim, Kim et al. 2012). In mammalian cells, viruses can also be used for transfection. Then the produced protein yields may be higher, but when viruses are used, the biosafety level is higher which will increase expenses of the protein production (Kollewe, Vilcinskas 2013).

2.4 Insect cell expression systems

2.4.1 Stably-transfected drosophila cells

Drosophila Schneider cells were isolated from late-stage Drosophila melanogaster embryos over 30 years ago (Schneider 1972). Drosophila Schneider cells are still used. However, to date only Schneider's cell lines 2 and 3 are used for heterologous gene expression. Drosophila S2 cells are suitable for high-level, low-cost production of eukaryotic proteins. It is possible to integrate multiple copies of expression plasmids allowing isolation of highly productive and stable polyclonal cell lines. S2 cell cultures may achieve high cell densities, which can also decrease laboratory expenses (Moraes, Jorge et al. 2012). Additionally they can be maintained at room temperature in serum-free media without CO_2 supply. Endogenous Drosophila proteins generally do not

interact with mammalian proteins. Therefore S2 cells can provide a "null background" for functional studies of the proteins (Invitrogen). Drosophila S2 cells can also produce recombinant proteins with post-translational modifications, such as glycan structures (Becker-Pauly, Stöcker 2011). However, in some cases the molecular weight of the protein has been different compared to the native mammalian protein. Some differences have also been detected in glycosylations, but despite differences in the structure, the activity of the protein has been even higher (Li, Tsing et al. 1996, Chang, Yang et al. 2005)

There are several factors that affect the growth and gene expression of S2 cells: particularly the selected cell line, cell passage, inoculum concentration, culture medium, temperature, dissolved oxygen concentration, pH, hydrodynamic forces, and toxic metabolites (Moraes, Jorge et al. 2012). In spite of the fact that S2 cell lines are difficult to grow at low densities, which somehow limits the process to select the desired cell clone, good performances in terms of the desired heterologous protein production have been reported (Lemos, Santos et al. 2009, Dos Santos, Lemos et al. 2009, Ventini, Astray et al. 2010, Yang, Song et al. 2012).

2.4.2 Stably-transfected lepidopteran cell lines

Genetically transformed lepidopteran insect cell lines have biotechnological applications as constitutive recombinant protein production platforms, and as improved hosts for baculovirus-mediated recombinant protein production (Lin, Jarvis 2013). Although *Sf9* and High Five^{TM} cells are generally associated with the baculovirus expression vector system, they are also used for stable transfection (Gouveia, Kandzia et al. 2010). Constitutive expression is generated by means of Immediate Early (IE) promoters from Multicapsid Nuclear Polyhedrosis Viruses (MNPV) (Lin, Jarvis 2013). Stably transfected cells are selected using either a resistance gene cassette linked on the expression vector or unlinked on a separate plasmid. The protein amounts obtained are typically in the range of 0.2- 50 mg L^{-1} by using the commercially available InsectSelectTM System, by Life Technologies (Gouveia, Kandzia et al. 2010).

2.4.3 Site-specific gene integration for stable expression

Site-specific DNA integration allows heterologous gene expression without extensive clone screening (Fernandes, Vidigal et al. 2012). The site-specific DNA integration

technology has been used in cultured insect cells (*sf9*) and also in cultured mammalian cells. Additionally it is also used in organisms, such as *drosophila* and mice (Fernandes, Vidigal et al. 2012).

A tagging cassette harboring a reporter gene is randomly introduced into the cell genome after screening different transfection protocols. After single-cell cloning, reporter gene expression is used to screen for high-producer clones. By using the Recombinase-Mediated Cassette Exchange (RMCE), any gene of interest can then be integrated into the first cell line genome at the same site. The cells are transfected with two plasmids, with the plasmid carrying the transgene flanked by compatible recombinase recognition sites and a plasmid providing the recombinase coding sequence. This recombinase coding sequence allows the transgene to be exchanged for an integrated reporter gene (Turan, Zehe et al. 2013).

2.5 Baculovirus expression vector systems (BEVS)

Baculovirus–insect cell expression systems have the capacity to produce recombinant proteins at high levels by simultaneously providing protein-processing capabilities with high similarity to eukaryotic protein processing (Jarvis 2009). Additionally BEVS is suitable for production of multi-protein complexes. BEVS can also provide specialized proteins for enhanced processing (e.g., chaperones) (Sokolenko, George et al. 2012).

Through genetic manipulation baculoviruses, and in particular *autographa californica* multiple nucleopolyhedrovirus (AcMNPV), have been engineered to be a biotechnological tool that is able to transduce insect and also mammalian cells (Sokolenko, George et al. 2012). Gene transfer in the BEVS is carried out by a baculovirus infection, followed by episomal replication and expression, which remove the need to select integrated transgenes. In addition, strong viral promoters will help achieve high yields in recombinant protein production (Kollewe, Vilcinskas 2013).

The first recombinant baculovirus vectors were produced using a homologous recombination. The double crossover recombination needed in this method was a relatively rare event. Additionally recognizing polyhedron-negative baculoviral plaques was a serious problem because it had to be done with a trained eye by visually

identifying polyhedron-negative plaques by examining the assay plate under a dissecting microscope (Jarvis 2009). Since this first method to produce recombinant baculoviruses, more improved techniques have been developed. Improvements and modifications were made to transfer plasmids into the baculovirus genome. Commercialized examples of these techniques used in baculovirus expression system are, for instance, Bac-to-Bac™ system by Life Technologies, flashBAC™ by Oxford Expression Technologies, BaculoDirect™ system by Harwood and colleagues at Invitrogen and DiamondBac™ (Jarvis 2009).

Because a Bac-to-Bac™ system was used in this study, the method is described in more detail. In the Bac-to-Bac® Baculovirus Expression System, a gene of interest is first cloned into a donor plasmid (pFastBac™). After the cloning, the correct sequence is checked by a restriction analysis, PCR or sequencing, or a combination of previous techniques. The confirmed donor plasmid is then transformed into competent cells (*E. Coli* DH10Bac™), which contain bacmid DNA, and a helper plasmid. After transposition, the bacmid is propagated in *E. Coli* DH10Bac™ (Luckow, Lee et al. 1993). The bacmid propagates in *E. coli* DH10Bac™ cells as a large plasmid that confers resistance to kanamycin and can complement a *lacZ* deletion present on the chromosome to form colonies that are blue (Lac+) in the presence of a chromogenic substrate such as Bluo-gal or X-gal and the inducer, IPTG. Transposing a mini Tn7 element from a pFastBacTM donor plasmid to the mini-attTn7 attachment site on the bacmid generates recombinant bacmids. A helper plasmid provides Tn7 transposition functions allowing the transfer of gene of interest into the bacmid DNA and it also confers resistance to tetracycline. The resulting large molecular weight recombinant bacmid DNA can then be extracted and after PCR or sequencing confirmation of insertion, it may be transfected into insect cells to generate recombinant baculovirus. Once the virus is harvested and the titer checked by plaque assay, virus stocks can be grown and protein production assayed. As the system uses suspension cells, protein production can be scaled up as required (Bac-to-Bac® Baculovirus Expression System manual, Invitrogen).

Another important part of the baculovirus expression system are the used insect cell lines. Most commonly used hosts for the baculovirus transfer vector are two lepidopteran insect cell lines *Spodoptera frugiperda 9* (*Sf9*) and High Five™ insect cells

(Jarvis 2009). The Order Lepidoptera moths and butterflies are the hosts for many viruses in the family of Baculoviridae, including AcMNPV (Jarvis 2009). The first Lepidoptera insect cell cultures were established by Grace in 1962 (Grace 1962). Currently approximately 250 insect cell lines have been developed (Lynn 2007). *Sf9* is a subclone of IPLB-SF-21 and it was described by Summers and Smith in 1987 (Summers, Smith 1987). This cell line has been isolated from pupal ovarian tissue of *Spodoptera frugiperda* (fall armyworm) (Vaughn, Goodwin et al. 1977). The High Five™ (BTI Tn 5B-1) cell line has been originally isolated from adult ovarian tissue of *Trichoplusia ni* (cabbage looper) in 1992 (Wickham, Davis et al. 1992).

Sf9 and High Five™ cells can be cultured in either adherent or suspension cultures. When the cells are cultured in suspension, the production can be scaled up as mentioned before. Adherent cultures are usually used for plaque assay and for quantify recombinant baculovirus expression vectors (Jarvis 2009). The optimal cell culture temperature is lower (28°C) for insect cells than for mammalian cell cultures (37°C). In addition, EDTA or trypsin is not needed because *Sf9* and High Five™ cells are not very adherent cells. Both cell lines can be cultured in either growth media with serum or in serum-free media. However, *Sf9* and High Five™ cells have to be adapted to a specific growth medium. It is advisable to do the adaptation by slowly weaning the cells out of the serum-containing medium into the serum-free medium. Sometimes the slow weaning is also needed to move insect cell cultures from one specific type of serum-free medium into another (Jarvis 2009).

As mentioned earlier, in insect cells, recombinant proteins can be folded, they can be chemically modified and assembled into soluble end products. However, despite these properties, insect cell production systems have also limitations (Jacobs, Callewaert 2009). For instance, both *Sf9* and High Five™ cells are inable to process N-linked carbohydrate side chains (N-glycans) to the same extent as in mammalian cells (Shi, Jarvis 2007). While native mammalian glycoproteins often have complex N-glycan structures with terminal sialic acids, insect cell derived recombinant glycoproteins usually have much more simple side chains. These side chains are known as paucimannosidic N-glycans. In addition, insect-derived N-glycans may cause allergenic reactions. This may be a problem when the baculovirus insect cell expression system is

used for producing pharmaceutical applications, particularly in therapeutic applications (Shi, Jarvis 2007).

Several strategies have been developed to solve this problem. For example, development of culture conditions that might support the production of recombinant glycoproteins with more authentic, mammalian-like N-glycans, and identifying alternative, natural baculovirus-host combinations that might be able to produce recombinant glycoproteins with more authentic N-glycans (Shi, Jarvis 2007). In addition, baculovirus vectors or their hosts have been genetically engineered to encode and express mammalian genes involved in the latter steps of N-glycan processing (Hollister, Jarvis 2001).

3 Aim of the study

The aim of this study was to produce and purify recombinant CA IX enzymes using Bac-To-Bac baculovirus expression vector system in insect cells. Purifications were performed by using a His tag purification system and an Avidin tag purification system. Two different purification systems were used in order to find out which was more feasible for CA IX protein. The His tag was included either in C-terminus or in N-terminus of the recombinant protein, in order to find out for which one of the His tags the thrombin cleavage would be more successful, and to find out if the His tag location has an effect on the enzyme activity. The work included the production and purification of CA IX catalytic domain and the proteoglycan domain. The catalytic domain and PG-domain were produced separately, because the crystallization of the proteins was predicted to work better in two parts. The analysis of the enzyme activity of the catalytic domain was also studied. Enzymes and results produced in this thesis were aimed to be used in further studies that revealed biochemical properties of CA IX, such as oligomerisation, glycosylation and the 3D structure.

4 Materials and methods

4.1 Recombinant protein production using Baculovirus expression system (BVES)

Recombinant baculoviruses were manufactured according to the instructions of the Bac-To-Bac® Baculovirus expression system (Invitrogen, Carlsbad, CA).

4.1.1 Generation of the recombinant *CA9* cDNA constructs

Three cDNA constructs encoding the active domain of the human *CA9* and one encoding the proteoglycan domain of the human *CA9* were made using the stepwise elongation of sequence-PCR (SES-PCR). DNA and amino acid sequences of the produced proteins in the study are presented in Appendix 1. CA IX sequence was based on UniProt entry Q16790. The aim was to produce soluble recombinant proteins CAIX-Chis, CAIX-NHis and CAIX-PG that could be purified directly from the insect cell medium. CAIX-Avd was aimed to be purified from the cells after sonication. Therefore the recombinant proteins CAIX-C-his, CAIX-N-His and CAIX-PG contained a *CA9* native signal sequence, of which codons had been optimized for the *Spodoptera frugiperda 9 (sf9)*. CAIX-C-His, CAIX-N-His and CAIX-Avd included only the CA domain and CAIX-PG included only the PG-domain. CAIX-Avd contained the Avidin signal sequence and the Avidin linker. The transmembrane domain and the intracytoplasmic tail were removed from all the recombinant sequences in order to make secretion possible. Sequences encoded also eight histidine residues (CAIX-C-His, CAIX-N-His and CAIX-PG) or alternatively Avidin tag (CAIX-Avd) for protein purifications. A thrombin cleavage site was added for polyhistidine- and Avidin tag removal. BglII and HindIII restriction enzyme sites were included for cloning the 5´and 3´ends, respectively. cDNA constructs are presented in Figure 3**.**

The PCR reactions were performed by using a thermal cycler (Gene Amp PCR system 9700, Applied Biosystems, Foster City, CA). The enzyme used in these reactions was Phusion polymerase (Finnzymes, Espoo, Finland), and it was used according to manufacture´s instructions. 10 % dimethyl sulfoxide (DMSO) was used as an additive in PCR reactions C, D, E, G and J (Table 1), as in these reactions primer dimers due to long primer sequences interfered with the PCR reaction. Primer sequences used in SES-PCR reactions are presented in Table 2. In the first PCR reaction, the template was the

full-length human *CA9* cDNA in the pOTB7 vector (IMAGE Clone 4865275, Geneservice Ltd, Cambridge, UK). PCR program for the first reaction (CAIX-PCR) included 98°C denaturation step for 2 minutes, which was followed by 33 cycles of denaturation at 98°C for 10 s, then annealing at 65°C for 30 s and extension at 72°C for one minute, followed by final extension at 72°C for 8 min. With other reactions denaturation step was 3 minutes and extension time was 30 seconds followed by 7 minutes final extension. Used temperatures were same with the exception of annealing temperatures mentioned in Table 1. All PCR reactions are presented in Table 1.

The PCR-products were analyzed using 1 % agarose gel electrophoresis containing 0,1 µg/ml ethidium bromide and 100 bp DNA ladder (New England BioLabs, Beverly, MA). The gel running time was 35 minutes and voltage was 110V (Biometra, Göttingen, Germany). Correct PCR product was cut from the gel and purifications of the DNA were performed using GFX PCR DNA and Gel Band Purification Kit (Amersham Biosciences, Poole UK) according to manufactures instructions. The purified DNA was used as a template in further reactions as described in Table 1.

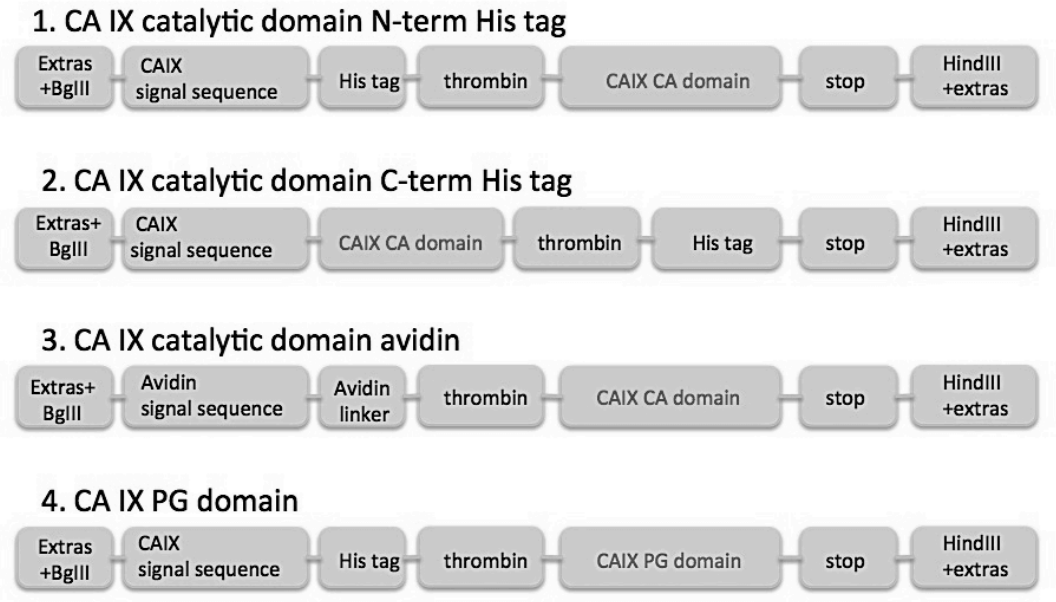


Figure 3. *Schematic figure of the used CA IX cDNA constructs.*

Table 1. SES-PCR reactions.

Protein	Reaction	Template	Primers	Annealing T (°C)	DMSO (5µl)
CAIX	CAIX-PCR	Vector	F1+ R1	65	-
CAIX-C-His	A	Vector	F2+ R2	50	-
	B	CAIX-PCR	F3+ R3	50	-
	C	B	F3+ R4	50	+
	D	A+C	F2+ R4	50	+
CAIX-N-His	E	Vector	F2+ R5	50	+
	F	CAIX-PCR	F4+ R6	50	-
	G	E+F	F2+ R6	45	+
CAIX-Avd	H	CAIX-PCR	F5+ R6	50	-
	I	H+ AVD	F6+ R6	45	-
CAIX-PG	J	Vector	F2+ R7	45	+

Table 2. Primer sequences used in the SES-PCR reactions (F=forward, R=reverse).

Primers	Sequences (5´→ 3´)
F1	CTGCTGCTTCTGATGCCTGTCC
F2	GGCCAGATCTATGGCTCCCCTGTGCCCCTCCCCCTGGCTGCCCCTGCTGATCCCCGCTCCC
F3	CACCCCCAGAGTCATTGGCGCTATGGAGGC
F4	CCACAGGGACAAAGAAGGGGATGACCAGAGTCATTGGCGCTATGGAGGCGACCCGCC
F5	GGTGGCTCCCTGGTGCCCCGTGGTTCCCACAGGGACAAAGAAGGGGATGACCAGAGT
F6	GGCCAGATCTATGGTGCACGCAACCTCCCCG
R1	CCTCTGGCTGGCTTCTCACATTCT
R2	ATAGCGCCAATGACTCTGGGGGTGCACGGGCATCAGCAGCAGCAGGGA
R3	ATGGTGGTGGGAACCACGGGGCACCAGAGGGAAGGAGGCCTCAATCACTCGCCCATT
R4	TTCGCCAAGCTTTTAGTGGTGATGGTGGTGATGGTGGTGGGAACCACGGGGCACCAG
R5	TGGTCATCCCCTTCTTTGTCCCTGTGGGAACCACGGGGCACCAGGTGGTGATGGTGG
R6	TTCGCCAAGCTTTTAGAAGGAGGCCTCAATCACTCGCCCATT
R7	TTCGCCAAGCTTTTAAGGTAGATCCTCTAACTTCAGGGAGCC

4.1.2 Generation of the recombinant plasmid

Enzymes BglII and HindIII were used to digest final PCR products and BamHI and HindIII were used to digest pFastBac1 vector. Ligation of the PCR products into the pFastBac1 vector was carried out using T4 ligase (Invitrogen). After ligation, the recombinant plasmids were transformed into TOP 10 cells according to manufacturer's instructions (Invitrogen). Plasmids were purified using the Qiagen Spin Miniprep Kit (Hilden, Germany).

Following the transformation and the purification, the plasmids were sequenced to confirm the correct coding sequence of the recombinant proteins. Sequencing was performed in both directions by means of ABI PRISM Genetic Analyser 9100

instrument (Applied Biosystems) and using the ABI PRISM Big Dye Terminator Cycle Sequencing ready Reactions Kit version 3.1 (Applied Biosystems). The reaction mixture contained 2 µl Bigdye mix, 2 µl sequencing buffer, 1 µl primer (1,6 pmol) 1 µl plasmid and 4 µl H_2O and the PCR program was performed according to the instructions of the manufacturer. The ethanol precipitation was carried out in the following way: 80 µl of 76 % EtOH was added into sample tubes. This was followed by vortexing and incubation at room temperature for 15 minutes. After the incubation, the samples were briefly centrifuged and the supernatants were removed. Then 250 µl of 70 % EtOH was added into the sample tubes and samples were mixed by vortexing and centrifuged for 10 minutes. After centrifugation, the supernatants were removed and samples were allowed to dry over night at 4°C. Next day the samples were re-suspensed with 13 µl Hidi and then mixed by vortexing. This was followed by fast spinning and incubation at 95°C for two minutes. Then the samples were put on ice and after that vortexing and spinning were conducted and finally the samples were kept on ice until placing them to the analyzer.

4.1.3 Generation of the recombinant bacmid

Site-directed transposition was used to generate recombinant bacmids containing the cDNAs coding for CA IX. In transposition, the DH10Bac™ *E. coli* cells were used as a host for pFastBac1 vector. DH10Bac™ cells contained a baculovirus shuttle vector (bacmid) and a helper plasmid. The helper plasmid contained Tn7 transposition functions that are needed for transposition with donor vector. The following method was used for transformation: a streak of DH10Bac bacteria was re-suspended to 100 µl of 100 mM $CaCl_2$ and then incubated on ice for 15 min. After the incubation, 100 ng of plasmids were added to the cell suspension and incubation on ice continued 30 min. Heat shock was carried out at 37°C for 2 min and after heat shock 450 µl of SOC medium (Invitrogen) was added. The cells were grown for 4 hours in an orbital shaker at 37°C and then collected by centrifuging 1 minute at 6000g at 20°C. After centrifugation, the cells were re-suspended to 100 µl SOC medium and then applied on Luria-Bertani (LB) plates that were prepared according to the Bac-to-Bac® instructions. Also the recombinant bacmid selection and transfection of the insect cells were performed according to the manufacturer's instructions.

4.1.4 Production and purifications of the recombinant enzymes

The *Sf9* insect cells (*Spodoptera frugiperda 9*, Invitrogen) were cultured in HyQ SFX-Insect serum free cell culture medium (Hyclone, Logan, UT) at 27°C in an orbital shaker. After 72 hours from infection, the culture medium was collected for protein purifications by centrifuging cells at 20°C, 5 minutes and 2000 G.

The purifications for recombinant enzymes CAIX-C-His, CAIX-N-His and CAIX-PG were carried out using Probond™ Purification System (Invitrogen). The purifications were performed in the following way: 100 ml insect cell culture medium, 1000 ml of native binding buffer and 7 ml of the nickel-chelating resin were combined and then mixed on a magnetic stirrer at 25°C for 1 hour. During the time on the magnetic stirrer, His tagged proteins had possibility to bind to the resin. After mixing the resin and the insect cell culture medium on magnetic stirrer, the solution was transferred into the column. The resin was washed with 500 ml of binding buffer and the proteins were eluted with elution buffer containing 250 mM imidazole. The native binding, washing and elution buffers were prepared according to manufacturer´s instructions.

Purifications for recombinant enzyme CAIX-Avd were performed using the Avidin-biotin purification system. Avidin is a chicken egg-white glycoprotein, which can bind biotin molecules and therefore it is widely used for affinity-based separations (Airenne, Laitinen et al. 1999). Purifications for CAIX-Avd were carried out as follows. The cell pellet (from 50 ml cell culture) was re-suspended to 10 ml Hillo I buffer (50 mM Tris–HCl, pH 8/1% Triton X-100/2mM EDTA/150 mM NaCl) and then the suspension was sonicated for 4 minutes (pulse 5 second on and 1 off, amplitude 20 %) and centrifuged 15000 G at 4°C for 15 minutes. During the centrifugation 2-iminobiotin-sepharose resin was prepared by washing (centrifugation 500 G for 5 minutes) it two times with 50 ml pH 11 buffer (50 mM Na_2CO_3, 1 M NaCl). After the 15 minutes centrifugation, 600 µl 2-iminobiotin-sepharose resin was added into the supernatant and 15 ml tube was filled up with pH 11 buffer (binding buffer). After gentle shaking on a rotator at 4°C for 1 hour, the resin was collected by centrifugation 500 G at 4°C for 5 minutes. Washing of the resin was carried out for two times with 15 ml binding buffer by centrifugation and the protein was eluted using pH 4 elution buffer (50 mM $C_2H_3NaO_2$, 0,1M NaCl).

The purified CAIX-C-His and CAIX-N-His (CAIX-PG and CAIX-Avd were not managed to be purified) were exchanged into 50 mM Tris-HCl, pH 7.5 (Sigma-Aldrich, St. Louis, MO) by means of Amicon Ultra 10 kDa cut-off centrifugal filter devices (Millipore, Carrigtwohill, Ireland). Because of His tag removal, purified recombinant proteins were treated with 60 µl of 50 % thrombin slurry (Thrombin CleanCleave KIT, Sigma-Aldrich) per 1mg of protein with slow mixing on a rotator at 25°C for 1 hours. A filtration was used to remove the thrombin slurry from the protein solution and a centrifugal filter device was used to wash and remove the His tag. Even though long incubation times and high enzyme concentrations were tested, the His tag could not be removed from the CAIX-C-His.

4.2 Characterization of the recombinant enzymes

SDS-PAGE analyses were used to characterize recombinant enzymes. SDS-PAGE gels were treated with the Colloidal blue staining Kit (Invitrogen) or Western blotting. SDS-PAGE gels were made from 5 % stacking and 10 % separating polyacrylamide gel and running of the gel was performed under reducing conditions according to previous literature (Laemmli 1970). Precision Plus Protein™ Dual Color Standard (Bio-Rad, Hercules, CA) was used for determining the molecular weight of proteins.

Western blot analyses were performed using a Novex XCell II blot module (Invitrogen). The separated recombinant proteins were transferred electrophoretically from the gel to polyvinydine fluoride (PVDF) membrane (Macherey-Nagel, Düren, Germany) at 36 V and for 1 hour and 15 minutes. The proteins were blocked for 25 minutes with cow colostral whey (Biotop Oy, Oulu, Finland) diluted to 1:10 in TBST buffer (20 mM Tris, 500 mM NaCl, 0.3 % Tween-20, pH 7.5). Then the proteins were incubated with primary antibody in TBST buffer for 60 minutes (His-mouse diluted 1:1000 for his tag and Avd-rabbit diluted 1:1000 for Avd tag). The antibody M75 diluted 1:100 was used to detect CAIX-PG (Pastorekova, Zavadova et al. 1992). After incubation, the membranes were washed 5 times for 5 minutes in TBST buffer. The secondary antibody (Horseradish peroxidase-labeled sheep anti-mouse IgG for M75) was diluted 1:25000 in TBST buffer and the membranes were incubated in secondary antibody for 60 minutes. After the incubation in the secondary antibody, the membranes were washed four times

for five minutes in TBST buffer. The ECL detection was completed according to manufacturer's instructions (Amersham Biosciences).

4.3 Protein concentration assays

Concentrations of the purified recombinant proteins were measured using BCA Protein Assay Reagent Kit (Pierce, Rockfort, IL) and DC Protein Assay (Bio-Rad, Hercules, CA). The protocols were performed according to manufacturer´s instructions. Concentrations were measured using three different dilutions and X replicates. The mean values of these results were used as the final concentrations.

4.4 Protein activity analyses

Activity analyses were performed using an applied photophysics stopped-flow instrument in the laboratory of Dr. Claudiu T. Supuran (Bioinorganic Chemistry Laboratory, University of Florence, Italy). These analyses are reported in an article by Hilvo *et al* (Hilvo, Baranauskiene et al. 2008).

5 Results

5.1 Production of recombinant proteins

CA IX recombinant proteins were produced using Bac-To-Bac® Baculovirus expression system (Invitrogen, Carlsbad, CA) and *Sf9* insect cells.

5.1.1 PCR-products of the constructs

PCR products encoding the recombinant proteins were separated and analyzed using agarose gel electrophoresis. The DNA ladder had 12 bands ranging from 100 to 1517 base pairs. The 500 and 1000 base pair bands had increased intensity to indicate reference point. The lengths of the final DNA constructs were 940 bp (CAIX-C-His), 970 bp (CAIX-N-His), 1300 bp (CAIX-Avd) and 420 bp (CAIX-PG). Several PCR-reactions and DNA purifications were done before the final PCR-products, but those results are not shown. Complete digested PCR-products of the constructs are presented in Figures 4-6.

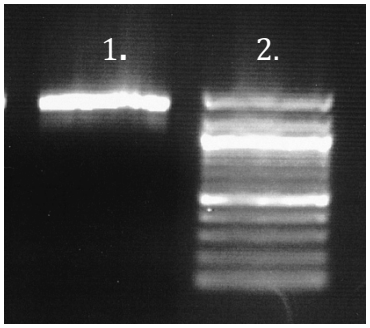


Figure 4. *1. (CAIX-Avd) 2. (ladder)*

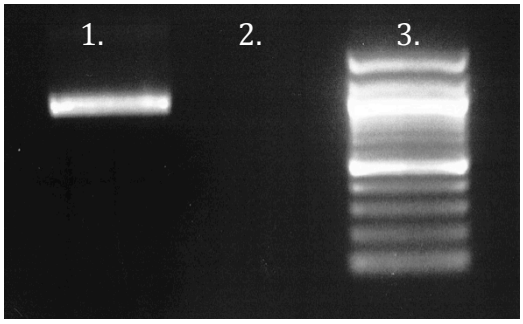


Figure 5. *1. (CAIX-C-His) 2. (-) 3. (ladder)*

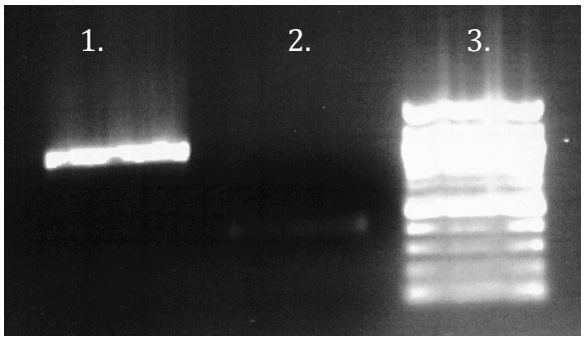


Figure 6. *1. (CAIX-N-His) 2. (CAIX-PG) 3. (ladder)*

5.1.2 Sequencing of the cDNA constructs

Correct coding sequences of the recombinant proteins were confirmed by sequencing. Sequencing was performed in both directions. If the sequencing was successful in one or the other direction, the sequence was interpreted to be correct. Results of the sequencing are presented in appendix 2.

5.1.3 Western Blot after production in insect cells

Western Blot analysis was used to confirm that the insect cell medium contained correct produced recombinant proteins. Western Blot results are presented in Figures 7 and 8. Recombinant proteins CAIX-N-His, CAIX-C-His were detected from the insect cell medium using a primary antibody for the His tag (Figure 7). CAIX-Avd was detected using a primary antibody for the Avidin and the CAIX-PG were detected using a primary antibody for the PG-domain. CAIX-Avd and CAIX-PG were detected from insect cells after sonication (Figure 8). The standard that was used contained ten bands of molecular weights (10–250 kD) including eight blue-stained bands and two pink reference bands (25 and 75 kD). Molecular weights of the produced proteins were CAIX-C-His (34,8 kD), CAIX-N-His (35,9 kD), CAIX-Avd (48,1 kD) and CAIX-PG (15,5 kD). All produced proteins showed signals of correct molecular weights in Western Blot analyses.

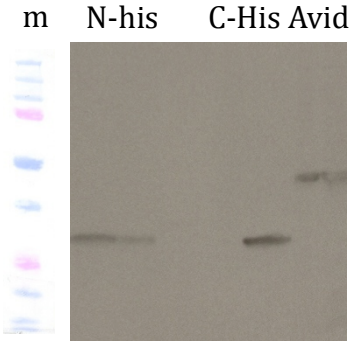


Figure 7. *Western Blot analysis to the CAIX-N-His and the CAIX-C-His from culture medium of sf9 cells and from CAIX-Avd after sonication of sf9 cells.*

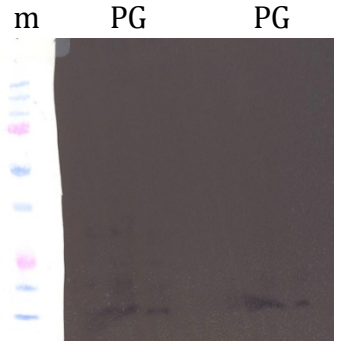


Figure 8. *Western Blot from the CAIX-PG-domain after sonication of the sf9 cells.*

5.2 Purification of recombinant proteins

Protein purifications were carried out using either affinity chromatography Propond purification system or Avidin tag purification system. SDS-PAGE analyses were used to characterize purified recombinant enzymes. Purifications were successful with the CAIX-C-His and the CAIX-N-His. Protein fractions of the CAIX-C-his and the CAIX-N-His were additionally extremely pure protein fractions as can be seen in Figures 9a and 9b. CAIX-Avd purification was not successful. The produced protein CAIX-Avd was found in the resin (black arrow) but not in the eluted fractions Figure 9c. This result indicated that the protein was not eluted. Because the purifications were successful with other recombinant proteins containing the CA IX catalytic domain, it was decided to proceed the studies with CAIX-C-His and CAIX-N-His. The experiments with CAIX-Avd were not continued. Results of the characterizations are presented in Figures 9a, 9b and 9c.

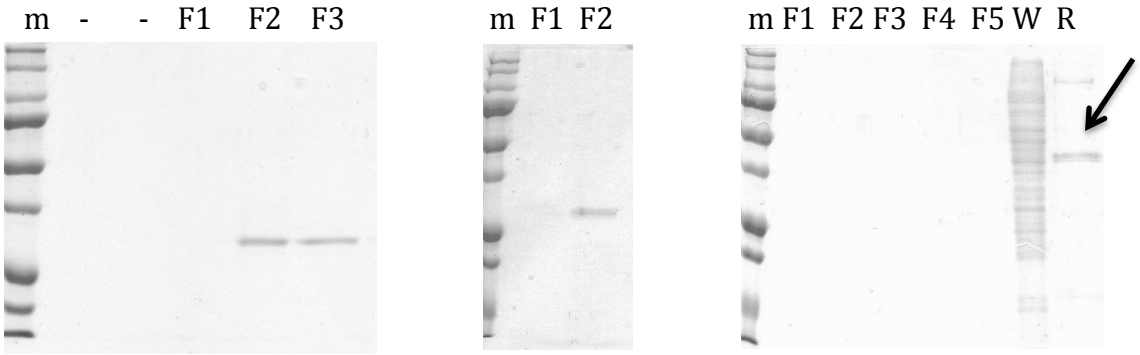


Figure 9. *SDS-PAGE gel figures* ***a*** *(CAIX-C-His),* ***b*** *(CAIX-N-His) and* ***c*** *(CAIX-Avd) from protein purifications by using coomassie staining. m=marker F=fraction W= wash R= resin*

5.3 His tag removal after protein purifications

Recombinant proteins were characterized using SDS-PAGE and Western Blot analyses before and after the His tag removal. After the protein purifications, eluted protein fractions were pooled and the buffer was exchanged to 50mM Tris-HCl. pH 7.5. The protein solution was then treated with thrombin slurry to remove the His tag. The His tag removal was successful with CAIX-N-His but not with CAIX-C-His. The protein size was clearly smaller after thrombin treatment which indicated a successful His tag removal. These SDS-PAGE results are presented in Figure 10. The His tag removal was also confirmed using Western Blot analysis. The antibody used in Western Blot was an

anti-His antibody, which showed that after thrombin treatment the His tag was successfully removed (Figure 11).

Figure 10. *CAIX-N-His before (bt) and after (at) thrombin treatment.*

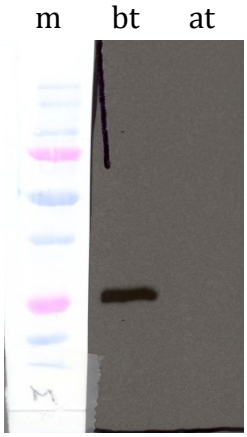


Figure 11. *CAIX-N-His before (bt) and after (at) thrombin treatment.*

5.4 Concentration of final protein preparations

Concentrations of the recombinant proteins were measured using DC™ Protein Assay from Bio-Rad and Pierce™ BCA Protein Assay. Protein concentrations of 358 ug/ml (CAIX-C-His) and 618 ug/ml (CAIX-N-His) were moderate and sufficient for further analyses, but not for the crystallization experiments for 3D structure determination. The results of the concentrations are presented in Table 3. The amounts of the protein solutions were 900 µl of CAIX-C-His and 930 µl of CAIX-N-His.

Table 3. Concentrations of the recombinant proteins CAIX-C-His and the CAIX-N-His.

	DC	BCA	DC + BCA
Average CAIX-C-His	339 ug/ml	370 ug/ml	358 ug/ml
Average CAIX-N-His	707 ug/ml	529 ug/ml	618 ug/ml

5.5 Catalytic activity of the recombinant proteins

Activity analyses were carried out using an applied photophysics stopped-flow instrument in the laboratory of Dr. Claudiu T. Supuran (Bioinorganic Chemistry Laboratory, University of Florence, Italy). Enzyme activities were measured with and without 50 μ M $ZnCl_2$. Activity values were also measured in presence of the $ZnCl_2$, because it was unknown if the purified fraction of recombinant protein contained

enough Zn^{2+} to occupy the active sites of enzyme molecules. Results are presented in Table 4, which includes also previously published activity values measured with catalytic domain of CA IX produced in *E.Coli* (Wingo, Tu et al. 2001). CAIX-C-His showed identical values with the CA IX catalytic domain produced in *E.Coli*. Full length CA + PG form showed similar activity values with CA II that has been considered to be the most active CA, according to previous literature (Khalifah 1971). Addition of the $ZnCl_2$ into the reaction increased activity values. Catalytic activity increased with CAIX-C-his form by tenfold, and with whole length CA + PG form by twentyfold in the presence of the $ZnCl_2$. Comparative experiments were also done with other CA isozymes, I, II, IV, XII, XIV and CA XV, in the presence of $ZnCl_2$ (Table 5). CA XV was produced under the same experimental conditions used for CA IX. CA XV did not show same increase in activity values. These results showed that the insect cell expression system offered enough Zn2+ ions to the active site of the CA molecules. Additionally it indicated that the addition of $ZnCl_2$ did not cause an artificial increase in the CO_2 hydration (Hilvo, Baranauskiene et al. 2008).

Table 4. Activity values of CAIX: CAIX-C-His, CAIX E Coli and CA+PG form with and without 50 μ M $ZnCl_2$ (Hilvo, Baranauskiene et al. 2008)

CAIX	K_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m ($M^{-1}S^{-1}$)
CAIX, *E Coli*	3.8×10^5	6.9	5.5×10^7
CAIX-C-His, *sf9*	3.8×10^5	7.0	5.4×10^7
CAIX-C-His, *sf9*, $ZnCl_2$	4.2×10^6	7.5	5.6×10^8
CAIX, CA+PG, *sf9*	1.1×10^6	7.5	1.5×10^8
CAIX, CA+PG, *sf9*, $ZnCl_2$	2.5×10^7	7.3	3.4×10^9

Table 5. Effect of 50 μ M $ZnCl_2$ on the CO_2 hydration activity of various CA isozymes (Hilvo, Baranauskiene et al. 2008).

Isoenzyme	CA activity (%)
hCA I	73 ± 6
hCA II	61 ± 4
hCA IV	63 ± 5
hCA IX	2266 ± 131
hCA XII	58 ± 4
hCA XIV	57 ± 3
mCA XV	60 ± 5

6 Discussion

CA IX has been analyzed intensively during the last decades, but some of its biochemical properties were uncharacterized when this study was conducted. In addition, the production of recombinant protein in insect cells was not reported in the previous literature for any CA. Thus, the goal of this study was to produce and purify recombinant CA IX proteins in insect cells. The aim was to use these CA IX recombinant proteins to analyze the activity, inhibition and oligomerisation of CA IX. The activity analyses were a part of this study, but the oligomerisation, stability, glycosylation, and inhibitory analyses are not reported here as they were conducted on separate further analyses (Hilvo, Baranauskiene et al. 2008, Innocenti, Hilvo et al. 2008).

Another aim was to conduct crystallization experiments to define the 3D structure of CA IX. The crystallization of CA IX had been tried in several research groups after production in bacterial and mammalian expression systems, but these attempts had failed. Therefore, the catalytic domain (CAIX-C-His, CAIX-N-His and CAIX-Avd) and PG-domain (CAIX-PG) were produced separately in this study, as the leading hypothesis was that the crystallization might be more feasible with several different target molecules. The insect cell line used for the production was *Sf9,* and the used expression system was baculovirus expression vector system (BVES). Also the High Five cells were planned to be used for the production, but during the time of this study it was decided not to use these cells because they did not grow well enough. The use of insect cell expression system for the production of recombinant CA proteins was found to be a functional way to produce highly pure CA IX enzyme. Later the results of this analysis were included in a report published by Hilvo *et al* (Hilvo, Baranauskiene et al. 2008). Hilvo *et al* used, however, also full length form of CA IX (PG+CA) containing both CA IX catalytic domain and CA IX PG-domain. The activity data of PG+CA form is also reported in this study.

Purifications of produced recombinant proteins were performed by means of affinity chromatography using Probond purification system and Avidin tag purification system. Two different purification systems were used in order to find out which one is more

feasible for CA IX protein. For the separations of the recombinant proteins, Probond purification system had an 8 x His tag and Avidin purification system had an Avidin tag. The His and the Avidin tags used in this study were aimed to be removed by thrombin cleavage. This removal was unsuccessful with C-terminal His tag (CAIX-C-His) but it was, however, successful with N-terminal His tag (CAIX-N-His). The reason why the thrombin cleavage did not work with C-terminal His tag is unclear. It may relate to enzyme structure, which may have prevented thrombin from accessing the cleavage site. The removal of His tag was not successful, even though long incubation times and different enzyme concentrations were tested. Besides the observation that C-terminal His tag was not removable, it may be concluded that it does not affect the catalytic activity of CA IX enzyme, firstly because the activity data was identical between CAIX-N-His (without His-tag) and CAIX-C-His enzymes (Hilvo, Baranauskiene et al. 2008), and secondly because the activity values were similar with the catalytic domain expressed in *E.Coli*, where the glutathione *S*-transferase tag was successfully removed (Wingo, Tu et al. 2001).

The production of recombinant protein CAIX-Avd was not as successful, and therefore, the comparison between the purification systems could not be done. Similarly to CAIX-Avd, the production of recombinant protein CAIX-PG was not as successful as was expected. The amounts of produced CAIX-PG protein were very low and the protein remained inside the cell for unexplained reasons. Both of the proteins CAIX-Avd and CAIX-PG were detected inside the cell using the Western Blot analysis. CAIX-Avd was aimed to be detected inside the cell. The cells were broken by sonication, but even after the sonications, the purifications were not successful. It is possible that these proteins formed aggregates inside the cell and therefore the secretion did not occur. There are several reasons why the produced recombinant protein may form aggregates. Heterologous proteins may reach non-physiological concentrations, which may promote aggregation. Overexpression of heterologous genes is stressful and may cause the saturation of the cellular folding machinery. Rapid intracellular protein accumulation may increase the probability of aggregation. Accordingly, the formation of aggregates is likely to occur during the production of recombinant proteins. The way of induction may also affect the solubility of the recombinant product (Palomares, Estrada-Moncada et al. 2004). Amounts of the purified recombinant proteins (CAIX-C-His, CAIX-N-His) produced in this study were moderate but suitable for the aimed activity analyses,

oligomerisation, stability, glycosylation and inhibitory studies, especially because the protein purity was high (Hilvo, Baranauskiene et al. 2008, Innocenti, Hilvo et al. 2008). However, due to low amount of purified proteins, the planned structural studies could not be conducted (Table 3).

Activity analyses were performed with CAIX-C-His, CA-N-His (data not shown) and also with full length PG+CA produced by Hilvo *et al* (Hilvo, Baranauskiene et al. 2008). Measured activity values were also compared with previously published results, where the catalytic activity of CA IX was measured with recombinant CA protein produced in *E. Coli* (Wingo, Tu et al. 2001). This recombinant CA protein produced in *E. Coli* contained only the catalytic domain of CA IX. Additionally, the catalytic activity was measured in the presence of 50 μM $ZnCl_2$. The analyses were performed with the presence of $ZnCl_2$, because it was unknown whether the recombinant protein produced in insect cells contained enough Zn^{2+} to saturate the active sites of all enzyme molecules. The correlation between the activity of the catalytic domain produced in the baculovirus-insect cell and the bacterial expression systems was observed. However, the addition of $ZnCl_2$ to the CAIX-C-His enzyme produced in insect cell increased its catalytic performance by approximately tenfold. The catalytic activity of PG + CA was increased by twentyfold. In fact, the activity values measured with full length PG+CA form showed catalytic activity, which was higher than reported in previous literature for this enzyme (Wingo, Tu et al. 2001). The PG + CA form showed the same *k*cat/*Km* value as that measured for CA II without the addition of $ZnCl_2$ or other metal additives (Khalifah 1971). This result was very novel because CA II has been considered to be the most active CA isozyme (Hilvo, Baranauskiene et al. 2008).

To confirm that the insect cell expression system offered enough Zn^{2+} ions, comparative analyses were also performed with other CA isozyme, CA XV, in the presence of $ZnCl_2$ (Table 5). CA XV was produced using similar production system and experimental conditions. $ZnCl_2$ addition to CA XV inhibited catalytic activity of the enzyme. In control experiments with isozymes I, II, IV, XII, and XIV, the addition of $ZnCl_2$ to the assay decreased the CO_2 hydration activity. These results showed that the insect cell expression system offered enough Zn^{2+} ions for catalysis. Additionally it also indicated

that the addition of $ZnCl_2$ did not cause an artificial increase in the CO_2 hydration (Hilvo, Baranauskiene et al. 2008).

The PG-domain was observed to have a notable influence on enzyme activity of CA IX. The difference between the PG + CA and CA forms was that the PG + CA form consisted also the PG domain. Because the activity values showed a significant difference between the catalytic domain and the full length CA+PG form, it may be concluded that the PG domain has an increasing effect in the CO_2 hydration activity of the CA domain. The mechanism behind that effect remained unexplained.

Before this study was conducted, CA IX was assumed to have a trimeric structure (Pastorekova, Zavada 2004). However, the later oligomerisation and stability studies made with CA-form (CAIX-C-his) and PG+CA form showed that CA IX probably has a dimeric structure that is stabilized by intermolecular disulfide bond(s). This important observation was later reported by Hilvo *et al* (Hilvo, Baranauskiene et al. 2008), and it was crucial in finishing the crystallization of CA IX successfully. Indeed, also the later crystallization experiments made with catalytic domain of CA IX (mutated Cys-41 to Ser) proved the assumption that CA IX is a dimeric protein (Alterio, Hilvo et al. 2009).

7 Conclusions

This study belonged to novel experiments, which proved that the Baculovirus insect cell expression system is a functional way to produce CA IX. Also His tag purification system was found to be a fuctional way to purify CA IX. Amounts of the purified recombinant proteins were relatively low, but they were suitable for the activity studies. The planned structural studies could not, however, be made due to low amount of purified proteins. According to the results of the activity analyses, the CA IX catalytic domain has identical activity values with the catalytic domain produced in *E.Coli*. Comparison of the activity of the CA IX catalytic domain and the full length CA IX form (PG+CA) showed that the enzyme with PG domain has remarkably higher activity values. Thus, it was concluded that the proteoglycan domain has a notable influence on enzyme activity. The addition of $ZnCl_2$ to the enzyme fraction was found to increase the catalytic activity. This property was concluded to be a unique feature for CA IX. In fact, the catalytic activity of the PG + CA form in the presence of $ZnCl_2$ was higher than what was measured before for any other CAs. The results of this study were used in further analyses, which revealed for the first time that CA IX has a dimeric structure, which is stabilized by intermolecular disulfide bond(s).

8 References

AIRENNE, K.J., LAITINEN, O.H., ALENIUS, H., MIKKOLA, J., KALKKINEN, N., ARIF, S.A., YEANG, H.Y., PALOSUO, T. and KULOMAA, M.S., 1999. Avidin is a promising tag for fusion proteins produced in baculovirus-infected insect cells. *Protein Expression & Purification,* **17**(1), pp. 139-145.

ALTERIO, V., HILVO, M., DI FIORE, A., SUPURAN, C.T., PAN, P., PARKKILA, S., SCALONI, A., PASTOREK, J., PASTOREKOVA, S., PEDONE, C., SCOZZAFAVA, A., MONTI, S.M. and DE SIMONE, G., 2009. Crystal structure of the catalytic domain of the tumor-associated human carbonic anhydrase IX. *Proceedings of the National Academy of Sciences of the United States of America,* **106**(38), pp. 16233-16238.

ASHIDA, S., NISHIMORI, I., TANIMURA, M., ONISHI, S. and SHUIN, T., 2002. Effects of von Hippel-Lindau gene mutation and methylation status on expression of transmembrane carbonic anhydrases in renal cell carcinoma. *Journal of Cancer Research & Clinical Oncology,* **128**(10), pp. 561-568.

BALDI, L., HACKER, D.L., ADAM, M. and WURM, F.M., 2007. Recombinant protein production by large-scale transient gene expression in mammalian cells: state of the art and future perspectives. *Biotechnology Letters,* **29**(5), pp. 677-684.

BARATHOVA, M., TAKACOVA, M., HOLOTNAKOVA, T., GIBADULINOVA, A., OHRADANOVA, A., ZATOVICOVA, M., HULIKOVA, A., KOPACEK, J., PARKKILA, S., SUPURAN, C.T., PASTOREKOVA, S. and PASTOREK, J., 2008. Alternative splicing variant of the hypoxia marker carbonic anhydrase IX expressed independently of hypoxia and tumour phenotype. *British journal of cancer,* **98**(1), pp. 129-136.

BECKER-PAULY, C. and STÖCKER, W., 2011. Insect Cells for Heterologous Production of Recombinant Proteins. *Insect Biotechnology,* 2, pp. 197-209.

BORON, W.F., 2010. Evaluating the role of carbonic anhydrases in the transport of HCO3--related species. *Biochimica et biophysica acta,* **1804**(2), pp. 410-421.

BRAHIMI-HORN, M.C. and POUYSSEGUR, J., 2007. Hypoxia in cancer cell metabolism and pH regulation. *Essays in biochemistry,* **43**, pp. 165-178.

BUI, M.H.T., VISAPAA, H., SELIGSON, D., KIM, H., HAN, K., HUANG, Y., HORVATH, S., STANBRIDGE, E.J., PALOTIE, A., FIGLIN, R.A. and BELLDEGRUN, A.S., 2004. PROGNOSTIC VALUE OF CARBONIC ANHYDRASE IX AND KI67 AS PREDICTORS OF SURVIVAL FOR RENAL CLEAR CELL CARCINOMA. *Journal of Urology,* **171**(6, Part 1 of 2), pp. 2461-2466.

CELIK, E. and CALIK, P., 2012. Production of recombinant proteins by yeast cells. *Biotechnology Advances,* **30**(5), pp. 1108-1118.

CHANG, K.H., YANG, J.M., CHUN, H.O. and CHUNG, I.S., 2005. Enhanced activity of recombinant beta-secretase from Drosophila melanogaster S2 cells transformed with cDNAs encoding human beta1,4-galactosyltransferase and Galbeta1,4-GlcNAc alpha2,6-sialyltransferase. *Journal of Biotechnology,* **116**(4), pp. 359-367.

CHOI, B.K., BOBROWICZ, P., DAVIDSON, R.C., HAMILTON, S.R., KUNG, D.H., LI, H., MIELE, R.G., NETT, J.H., WILDT, S. and GERNGROSS, T.U., 2003. Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast Pichia pastoris. *Proceedings of the National Academy of Sciences of the United States of America,* **100**(9), pp. 5022-5027.

DE SIMONE, G. and SUPURAN, C.T., 2010. Carbonic anhydrase IX: Biochemical and crystallographic characterization of a novel antitumor target. *Biochimica et biophysica acta,* **1804**(2), pp. 404-409.

DEMAIN, A.L. and VAISHNAV, P., 2009. Production of recombinant proteins by microbes and higher organisms. *Biotechnology Advances,* **27**(3), pp. 297-306.

DIVGI, C.R., PANDIT-TASKAR, N., JUNGBLUTH, A.A., REUTER, V.E., GONEN, M., RUAN, S., PIERRE, C., NAGEL, A., PRYMA, D.A., HUMM, J., LARSON, S.M., OLD, L.J. and RUSSO, P., 2007. Preoperative characterisation of clear-cell renal carcinoma using iodine-124-labelled antibody chimeric G250 (124I-cG250) and PET in patients with renal masses: a phase I trial. *The lancet oncology,* **8**(4), pp. 304-310.

DOS SANTOS, A.S., LEMOS, M.A., PEREIRA, C.A. and JORGE, S.A., 2009. Rabies virus glycoprotein expression in Drosophila S2 cells: influence of re-selection on protein expression. *Biotechnology Journal,* **4**(11), pp. 1578-1581.

DRIESSEN, A., LANDUYT, W., PASTOREKOVA, S., MOONS, J.N., GOETHALS, L., HAUSTERMANS, K., NAFTEUX, P., PENNINCKX, F., GEBOES, K., LERUT, T. and ECTORS, N., 2006. Expression of Carbonic Anhydrase IX (CA IX), a Hypoxia-Related Protein, Rather Than Vascular-Endothelial Growth Factor (VEGF), a Pro-Angiogenic Factor, Correlates With an Extremely Poor Prognosis in Esophageal and Gastric Adenocarcinomas. *Annals of Surgery,* **243**(3), pp. 334-340.

FERNANDES, F., VIDIGAL, J., DIAS, M.M., PRATHER, K.L., COROADINHA, A.S., TEIXEIRA, A.P. and ALVES, P.M., 2012. Flipase-mediated cassette exchange in Sf9 insect cells for stable gene expression. *Biotechnology and bioengineering,* **109**(11), pp. 2836-2844.

GOUVEIA, R., KANDZIA, S., CONRADT, H.S. and COSTA, J., 2010. Production and N-glycosylation of recombinant human cell adhesion molecule L1 from insect cells using the stable expression system. Effect of dimethyl sulfoxide. *Journal of Biotechnology,* **145**(2), pp. 130-138.

GRABMAIER, K., DE WEIJERT, M., UEMURA, H., SCHALKEN, J. and OOSTERWIJK, E., 2002. Renal cell carcinoma-associated G250 methylation and expression: in vivo and in vitro studies. *Urology,* **60**(2), pp. 357-362.

GRABMAIER, K., VISSERS, J.L., DE WEIJERT, M.C., OOSTERWIJK-WAKKA, J.C., VAN BOKHOVEN, A., BRAKENHOFF, R.H., NOESSNER, E., MULDERS, P.A., MERKX, G., FIGDOR, C.G., ADEMA, G.J. and OOSTERWIJK, E., 2000. Molecular cloning and immunogenicity of renal cell carcinoma-associated antigen G250. *International Journal of Cancer,* **85**(6), pp. 865-870.

GRACE, T.D., 1962. Establishment of four strains of cells from insect tissues grown in vitro. *Nature,* **195**, pp. 788-789.

HILVO, M., BARANAUSKIENE, L., SALZANO, A.M., SCALONI, A., MATULIS, D., INNOCENTI, A., SCOZZAFAVA, A., MONTI, S.M., DI FIORE, A., DE SIMONE, G.,

LINDFORS, M., JANIS, J., VALJAKKA, J., PASTOREKOVA, S., PASTOREK, J., KULOMAA, M.S., NORDLUND, H.R., SUPURAN, C.T. and PARKKILA, S., 2008. Biochemical characterization of CA IX, one of the most active carbonic anhydrase isozymes. *Journal of Biological Chemistry,* **283**(41), pp. 27799-27809.

HOLLISTER, J.R. and JARVIS, D.L., 2001. Engineering lepidopteran insect cells for sialoglycoprotein production by genetic transformation with mammalian beta 1,4-galactosyltransferase and alpha 2,6-sialyltransferase genes. *Glycobiology,* **11**(1), pp. 1-9.

HULIKOVA, A., ZATOVICOVA, M., SVASTOVA, E., DITTE, P., BRASSEUR, R., KETTMANN, R., SUPURAN, C.T., KOPACEK, J., PASTOREK, J. and PASTOREKOVA, S., 2009. Intact intracellular tail is critical for proper functioning of the tumor-associated, hypoxia-regulated carbonic anhydrase IX. *FEBS letters,* **583**(22), pp. 3563-3568.

INNOCENTI, A., HILVO, M., SCOZZAFAVA, A., LINDFORS, M., NORDLUND, H.R., KULOMAA, M.S., PARKKILA, S. and SUPURAN, C.T., 2008. Carbonic anhydrase inhibitors: the very weak inhibitors dithiothreitol, beta-mercaptoethanol, tris(carboxyethyl)phosphine and threitol interfere with the binding of sulfonamides to isozymes II and IX. *Bioorganic & medicinal chemistry letters,* **18**(6), pp. 1898-1903.

IVANOV, S., LIAO, S.Y., IVANOVA, A., DANILKOVITCH-MIAGKOVA, A., TARASOVA, N., WEIRICH, G., MERRILL, M.J., PROESCHOLDT, M.A., OLDFIELD, E.H., LEE, J., ZAVADA, J., WAHEED, A., SLY, W., LERMAN, M.I. and STANBRIDGE, E.J., 2001. Expression of hypoxia-inducible cell-surface transmembrane carbonic anhydrases in human cancer. *American Journal of Pathology,* **158**(3), pp. 905-919.

JACOBS, P.P. and CALLEWAERT, N., 2009. N-glycosylation engineering of biopharmaceutical expression systems. *Current Molecular Medicine,* **9**(7), pp. 774-800.

JAKUBICKOVA, L., BIESOVA, Z., PASTOREKOVA, S., KETTMANN, R. and PASTOREK, J., 2005. Methylation of the CA9 promoter can modulate expression of the tumor-associated carbonic anhydrase IX in dense carcinoma cell lines. *International journal of oncology,* **26**(4), pp. 1121-1127.

JANA, S. and DEB, J.K., 2005. Strategies for efficient production of heterologous proteins in Escherichia coli. *Applied Microbiology & Biotechnology,* **67**(3), pp. 289-298.

JARVIS, D.L., 2009. Baculovirus-insect cell expression systems. *Methods in enzymology,* **463**, pp. 191-222.

JAYARAJ, R. and SMOOKER, P.M., 2009. So you Need a Protein - A Guide to the Production of Recombinant Proteins . *The Open Veterinary Science Journal,* **3**, pp. 28-34.

KALUZ, S., KALUZOVA, M., LIAO, S.Y., LERMAN, M. and STANBRIDGE, E.J., 2009. Transcriptional control of the tumor- and hypoxia-marker carbonic anhydrase 9: A one transcription factor (HIF-1) show? *Biochimica et biophysica acta,* **1795**(2), pp. 162-172.

KALUZ, S., KALUZOVA, M., OPAVSKY, R., PASTOREKOVA, S., GIBADULINOVA, A., DEQUIEDT, F., KETTMANN, R. and PASTOREK, J., 1999. Transcriptional regulation of the MN/CA 9 gene coding for the tumor-associated carbonic anhydrase IX. Identification and characterization of a proximal silencer element. *Journal of Biological Chemistry,* **274**(46), pp. 32588-32595.

KALUZ, S., KALUZOVA, M. and STANBRIDGE, E.J., 2003. Expression of the hypoxia marker carbonic anhydrase IX is critically dependent on SP1 activity. Identification of a novel type of hypoxia-responsive enhancer. *Cancer research,* **63**(5), pp. 917-922.

KALUZOVA, M., PASTOREKOVA, S., SVASTOVA, E., PASTOREK, J., STANBRIDGE, E.J. and KALUZ, S., 2001. Characterization of the MN/CA 9 promoter proximal region: a role for specificity protein (SP) and activator protein 1 (AP1) factors. *Biochemical Journal,* **359**(Pt 3), pp. 669-677.

KANTH, B.K. and PACK, S.P., 2013. Carbonic Anhydrase as CO2 capturing agent: its
Classes and Catalytic Mechanisms. *Proceedings of the 2013 International Conference on Biology, Medical Physics, Medical Chemistry, Biochemistry and Biomedical Engineering,* , pp. 57-61.

KHALIFAH, R.G., 1971. The carbon dioxide hydration activity of carbonic anhydrase. I. Stop-flow kinetic studies on the native human isoenzymes B and C. *The Journal of biological chemistry,* **246**(8), pp. 2561-2573.

KIM, J.Y., KIM, Y.G. and LEE, G.M., 2012. CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Applied Microbiology & Biotechnology,* **93**(3), pp. 917-930.

KOLLEWE, C. and VILCINSKAS, A., 2013. Production of recombinant protein in insect cells. *American Journal of Biochemistry and Biotechnology,* **9(3)**, pp. 255-271.

LADEROUTE, K.R., 2005. The interaction between HIF-1 and AP-1 transcription factors in response to low oxygen. *Seminars in cell & developmental biology,* **16**(4-5), pp. 502-513.

LEMOS, M.A., SANTOS, A.S., ASTRAY, R.M., PEREIRA, C.A. and JORGE, S.A., 2009. Rabies virus glycoprotein expression in Drosophila S2 cells. I: design of expression/selection vectors, subpopulations selection and influence of sodium butyrate and culture medium on protein expression. *Journal of Biotechnology,* **143**(2), pp. 103-110.

LI, B., TSING, S., KOSAKA, A.H., NGUYEN, B., OSEN, E.G., BACH, C., CHAN, H. and BARNETT, J., 1996. Expression of human dopamine beta-hydroxylase in Drosophila Schneider 2 cells. *Biochemical Journal,* **313**(Pt 1), pp. 57-64.

LIN, C.H. and JARVIS, D.L., 2013. Utility of temporally distinct baculovirus promoters for constitutive and baculovirus-inducible transgene expression in transformed insect cells. *Journal of Biotechnology,* **165**(1), pp. 11-17.

LUCKOW, V.A., LEE, S.C., BARRY, G.F. and OLINS, P.O., 1993. Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in Escherichia coli. *Journal of virology,* **67**(8), pp. 4566-4579.

LYNN, D.E., 2007. Available lepidopteran insect cell lines. *Methods in molecular biology (Clifton, N.J.),* **388**, pp. 117-138.

MASEIDE, K., KANDEL, R.A., BELL, R.S., CATTON, C.N., O'SULLIVAN, B., WUNDER, J.S., PINTILIE, M., HEDLEY, D. and HILL, R.P., 2004. Carbonic anhydrase IX as a marker for poor prognosis in soft tissue sarcoma. *Clinical Cancer Research,* **10**(13), pp. 4464-4471.

MATTANOVICH, D., BRANDUARDI, P., DATO, L., GASSER, B., SAUER, M. and PORRO, D., 2012. Recombinant protein production in yeasts. *Methods in molecular biology (Clifton, N.J.),* **824**, pp. 329-358.

MCDONALD, P.C., WINUM, J.Y., SUPURAN, C.T. and DEDHAR, S., 2012. Recent developments in targeting carbonic anhydrase IX for cancer therapeutics. *Oncotarget,* **3**(1), pp. 84-97.

MERGULHAO, F.J., SUMMERS, D.K. and MONTEIRO, G.A., 2005. Recombinant protein secretion in Escherichia coli. *Biotechnology Advances,* **23**(3), pp. 177-202.

MORAES, A.M., JORGE, S.A., ASTRAY, R.M., SUAZO, C.A., CALDERON RIQUELME, C.E., AUGUSTO, E.F., TONSO, A., PAMBOUKIAN, M.M., PICCOLI, R.A., BARRAL, M.F. and PEREIRA, C.A., 2012. Drosophila melanogaster S2 cells for expression of heterologous genes: From gene cloning to bioprocess development. *Biotechnology Advances,* **30**(3), pp. 613-628.

MORGAN, P.E., PASTOREKOVA, S., STUART-TILLEY, A.K., ALPER, S.L. and CASEY, J.R., 2007. Interactions of transmembrane carbonic anhydrase, CAIX, with bicarbonate transporters. *American Journal of Physiology - Cell Physiology,* **293**(2), pp. 738-748.

NAKAGAWA, Y., UEMURA, H., HIRAO, Y., YOSHIDA, K., SAGA, S. and YOSHIKAWA, K., 1998. Radiation hybrid mapping of the human MN/CA9 locus to chromosome band 9p12-p13. *Genomics,* **53**(1), pp. 118-119.

NAKAMURA, J., KITAJIMA, Y., KAI, K., HASHIGUCHI, K., HIRAKI, M., NOSHIRO, H. and MIYAZAKI, K., 2011. Expression of hypoxic marker CA IX is regulated by site-specific DNA methylation and is associated with the histology of gastric cancer. *American Journal of Pathology,* **178**(2), pp. 515-524.

OOSTERWIJK, E., RUITER, D.J., HOEDEMAEKER, P.J., PAUWELS, E.K., JONAS, U., ZWARTENDIJK, J. and WARNAAR, S.O., 1986. Monoclonal antibody G 250 recognizes a determinant present in renal-cell carcinoma and absent from normal kidney. *International Journal of Cancer,* **38**(4), pp. 489-494.

OPAVSKY, R., PASTOREKOVA, S., ZELNIK, V., GIBADULINOVA, A., STANBRIDGE, E.J., ZAVADA, J., KETTMANN, R. and PASTOREK, J., 1996. Human MN/CA9 gene, a novel member of the carbonic anhydrase family: structure and exon to protein domain relationships. *Genomics,* **33**(3), pp. 480-487.

PALOMARES, L.A., ESTRADA-MONCADA, S. and RAMÍREZ, O.T., 2004. Production of Recombinant Proteins: Challenges and Solutions. In: P. BALBÁS and A. LORENCE, eds, *Recombinant Gene Expression*. Humana Press, pp. 15-51.

PASTOREK, J., PASTOREKOVA, S., CALLEBAUT, I., MORNON, J.P., ZELNIK, V., OPAVSKY, R., ZAT'OVICOVA, M., LIAO, S., PORTETELLE, D. and STANBRIDGE, E.J., 1994. Cloning and characterization of MN, a human tumor-associated protein with a domain homologous to carbonic anhydrase and a putative helix-loop-helix DNA binding segment. *Oncogene,* **9**(10), pp. 2877-2888.

PASTOREKOVA, S. and ZAVADA, J., 2004. Carbonic anhydrase IX (CA IX) as a potential
target for cancer therapy. *Cancer Therapy,* **2**, pp. 245-264.

PASTOREKOVA, S., PARKKILA, S., PARKKILA, A.K., OPAVSKY, R., ZELNIK, V., SAARNIO, J. and PASTOREK, J., 1997. Carbonic anhydrase IX, MN/CA IX: analysis of stomach complementary DNA sequence and expression in human and rat alimentary tracts. *Gastroenterology,* **112**(2), pp. 398-408.

PASTOREKOVA, S., PARKKILA, S. and ZAVADA, J., 2006. Tumor-associated carbonic anhydrases and their clinical significance. *Advances in Clinical Chemistry,* **42**, pp. 167-216.

PASTOREKOVA, S., ZAVADOVA, Z., KOSTAL, M., BABUSIKOVA, O. and ZAVADA, J., 1992. A novel quasi-viral agent, MaTu, is a two-component system. *Virology,* **187**(2), pp. 620-626.

PASTOREKOVA, S., RATCLIFFE, P.J. and PASTOREK, J., 2008. Molecular mechanisms of carbonic anhydrase IX-mediated pH regulation under hypoxia. *BJU international,* **101**(Supplement 4), pp. 8-15.

POTTER, C.P. and HARRIS, A.L., 2003. Diagnostic, prognostic and therapeutic implications of carbonic anhydrases in cancer. *British journal of cancer,* **89**(1), pp. 2-7.

REINWARTH, M., NASU, D., KOLMAR, H. and AVRUTINA, O., 2012. Chemical synthesis, backbone cyclization and oxidative folding of cystine-knot peptides: promising scaffolds for applications in drug design. *Molecules (Basel, Switzerland),* **17**(11), pp. 12533-12552.

ROSANO, G.L. and CECCARELLI, E.A., 2014. Recombinant protein expression in : advances and challenges. *Frontiers in microbiology,* **5**, pp. 172.

SCHNEIDER, I., 1972. Cell lines derived from late embryonic stages of Drosophila melanogaster. *Journal of embryology and experimental morphology,* **27**(2), pp. 353-365.

SEMENZA, G.L., 2001. Hypoxia-Inducible Factor 1: Control of Oxygen Homeostasis in Health and Disease. *Pediatric research,* **49**(5), pp. 614-617.

SHI, X. and JARVIS, D.L., 2007. Protein N-glycosylation in the baculovirus-insect cell system. *Current Drug Targets,* **8**(10), pp. 1116-1125.

SIVASHANMUGAM, A., MURRAY, V., CUI, C., ZHANG, Y., WANG, J. and LI, Q., 2009. Practical protocols for production of very high yields of recombinant proteins using Escherichia coli. *Protein Science,* **18**(5), pp. 936-948.

SOKOLENKO, S., GEORGE, S., WAGNER, A., TULADHAR, A., ANDRICH, J.M. and AUCOIN, M.G., 2012. Co-expression vs. co-infection using baculovirus expression vectors in insect cell culture: Benefits and drawbacks. *Biotechnology Advances,* **30**(3), pp. 766-781.

STILLEBROER, A.B., MULDERS, P.F., BOERMAN, O.C., OYEN, W.J. and OOSTERWIJK, E., 2010. Carbonic anhydrase IX in renal cell carcinoma: implications for prognosis, diagnosis, and therapy. *European urology,* **58**(1), pp. 75-83.

STRUCTURAL GENOMICS, C., CHINA STRUCTURAL GENOMICS, C., NORTHEAST STRUCTURAL GENOMICS, C., GRASLUND, S., NORDLUND, P., WEIGELT, J., HALLBERG, B.M., BRAY, J., GILEADI, O., KNAPP, S., OPPERMANN, U., ARROWSMITH, C., HUI, R., MING, J., DHE-PAGANON, S., PARK, H.W., SAVCHENKO, A., YEE, A., EDWARDS, A., VINCENTELLI, R., CAMBILLAU, C., KIM, R., KIM, S.H., RAO, Z., SHI, Y., TERWILLIGER, T.C., KIM, C.Y., HUNG, L.W., WALDO, G.S., PELEG,

Y., ALBECK, S., UNGER, T., DYM, O., PRILUSKY, J., SUSSMAN, J.L., STEVENS, R.C., LESLEY, S.A., WILSON, I.A., JOACHIMIAK, A., COLLART, F., DEMENTIEVA, I., DONNELLY, M.I., ESCHENFELDT, W.H., KIM, Y., STOLS, L., WU, R., ZHOU, M., BURLEY, S.K., EMTAGE, J.S., SAUDER, J.M., THOMPSON, D., BAIN, K., LUZ, J., GHEYI, T., ZHANG, F., ATWELL, S., ALMO, S.C., BONANNO, J.B., FISER, A., SWAMINATHAN, S., STUDIER, F.W., CHANCE, M.R., SALI, A., ACTON, T.B., XIAO, R., ZHAO, L., MA, L.C., HUNT, J.F., TONG, L., CUNNINGHAM, K., INOUYE, M., ANDERSON, S., JANJUA, H., SHASTRY, R., HO, C.K., WANG, D., WANG, H., JIANG, M., MONTELIONE, G.T., STUART, D.I., OWENS, R.J., DAENKE, S., SCHUTZ, A., HEINEMANN, U., YOKOYAMA, S., BUSSOW, K. and GUNSALUS, K.C., 2008. Protein production and purification. *Nature Methods,* **5**(2), pp. 135-146.

SUMMERS, M.D. and SMITH, G.E., 1987. *A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures*. Texas Agricultural Experiment Station.

SUPURAN, C.T. and SCOZZAFAVA, A., 2007. Carbonic anhydrases as targets for medicinal chemistry. *Bioorganic & medicinal chemistry,* **15**(13), pp. 4336-4350.

SUPURAN, C.T., 2008. Development of small molecule carbonic anhydrase IX inhibitors. *BJU international,* **101**(Supplement 4), pp. 39-40.

SVASTOVA, E., HULIKOVA, A., RAFAJOVA, M., ZAT'OVICOVA, M., GIBADULINOVA, A., CASINI, A., CECCHI, A., SCOZZAFAVA, A., SUPURAN, C.T., PASTOREK, J. and PASTOREKOVA, S., 2004. Hypoxia activates the capacity of tumor-associated carbonic anhydrase IX to acidify extracellular pH. *FEBS letters,* **577**(3), pp. 439-445.

SVASTOVA, E., WITARSKI, W., CSADEROVA, L., KOSIK, I., SKVARKOVA, L., HULIKOVA, A., ZATOVICOVA, M., BARATHOVA, M., KOPACEK, J., PASTOREK, J. and PASTOREKOVA, S., 2012. Carbonic anhydrase IX interacts with bicarbonate transporters in lamellipodia and increases cell migration via its catalytic domain. *Journal of Biological Chemistry,* **287**(5), pp. 3392-3402.

SVASTOVA, E., ZILKA, N., ZAT'OVICOVA, M., GIBADULINOVA, A., CIAMPOR, F., PASTOREK, J. and PASTOREKOVA, S., 2003. Carbonic anhydrase IX reduces E-cadherin-mediated adhesion of MDCK cells via interaction with beta-catenin. *Experimental cell research,* **290**(2), pp. 332-345.

SWIETACH, P., HULIKOVA, A., VAUGHAN-JONES, R.D. and HARRIS, A.L., 2010. New insights into the physiological role of carbonic anhydrase IX in tumour pH regulation. *Oncogene,* **29**(50), pp. 6509-6521.

SYRJANEN, L., LUUKKAALA, T., LEPPILAMPI, M., KALLIOINEN, M., PASTOREKOVA, S., PASTOREK, J., WAHEED, A., SLY, W.S., PARKKILA, S. and KARTTUNEN, T., 2014. Expression of cancer-related carbonic anhydrases IX and XII in normal skin and skin neoplasms. *APMIS : Acta Pathologica, Microbiologica, et Immunologica Scandinavica,* **122**(9), pp. 880-890.

TURAN, S., ZEHE, C., KUEHLE, J., QIAO, J. and BODE, J., 2013. Recombinase-mediated cassette exchange (RMCE) - a rapidly-expanding toolbox for targeted genomic modifications. *Gene,* **515**(1), pp. 1-27.

VAUGHN, J.L., GOODWIN, R.H., TOMPKINS, G.J. and MCCAWLEY, P., 1977. The establishment of two cell lines from the insect Spodoptera frugiperda (Lepidoptera; Noctuidae). *In vitro,* **13**(4), pp. 213-217.

VENTINI, D.C., ASTRAY, R.M., LEMOS, M.A., JORGE, S.A., RIQUELME, C.C., SUAZO, C.A., TONSO, A. and PEREIRA, C.A., 2010. Recombinant rabies virus glycoprotein synthesis in bioreactor by transfected Drosophila melanogaster S2 cells carrying a constitutive or an inducible promoter. *Journal of Biotechnology,* **146**(4), pp. 169-172.

WESTERS, L., WESTERS, H. and QUAX, W.J., 2004. Bacillus subtilis as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Biochimica et biophysica acta,* **1694**(1-3), pp. 299-310.

WICKHAM, T.J., DAVIS, T., GRANADOS, R.R., SHULER, M.L. and WOOD, H.A., 1992. Screening of insect cell lines for the production of recombinant proteins and infectious virus in the baculovirus expression system. *Biotechnology progress,* **8**(5), pp. 391-396.

WINGO, T., TU, C., LAIPIS, P.J. and SILVERMAN, D.N., 2001. The catalytic properties of human carbonic anhydrase IX. *Biochemical and biophysical research communications,* **288**(3), pp. 666-669.

WYKOFF, C.C., BEASLEY, N.J., WATSON, P.H., TURNER, K.J., PASTOREK, J., SIBTAIN, A., WILSON, G.D., TURLEY, H., TALKS, K.L., MAXWELL, P.H., PUGH, C.W., RATCLIFFE, P.J. and HARRIS, A.L., 2000. Hypoxia-inducible expression of tumor-associated carbonic anhydrases. *Cancer research,* **60**(24), pp. 7075-7083.

YANG, L., SONG, Y., LI, X., HUANG, X., LIU, J., DING, H., ZHU, P. and ZHOU, P., 2012. HIV-1 virus-like particles produced by stably transfected Drosophila S2 cells: a desirable vaccine component. *Journal of virology,* **86**(14), pp. 7662-7676.

YE, J., ALVIN, K., LATIF, H., HSU, A., PARIKH, V., WHITMER, T., TELLERS, M., DE LA CRUZ EDMONDS,CELINA M., LY, J., SALMON, P. and MARKUSEN, J.F., 2010. Rapid protein production using CHO stable transfection pools. *Biotechnology progress,* **26**(5), pp. 1431-1437.

ZAVADA, J., ZAVADOVA, Z., PASTOREK, J., BIESOVA, Z., JEZEK, J. and VELEK, J., 2000. Human tumour-associated cell adhesion protein MN/CA IX: identification of M75 epitope and of the region mediating cell adhesion. *British journal of cancer,* **82**(11), pp. 1808-1813.

ZHU, J., 2012. Mammalian cell protein expression for biopharmaceutical production. *Biotechnology Advances,* **30**(5), pp. 1158-1170.

9 Appendix 1.

CAIX-C-His

Extras, BglII, optimized signal sequence, CA-domain, thrombin, 8xHIS, STOP, HindIII, extras

Nucleotide sequence

ggccagatctatggctcccctgtgcccctcccctggctgcccctgctgatccccgctcc
cgctcccggtctgaccgtgcagctgctgctgtccctgctgctgctgatgcccgtgcaccc
ccagcagagtcattggcgctatggaggcgacccgccctggccccgggtgtccccagcctg
cgcgggccgcttccagtccccggtggatatccgcccccagctcgccgccttctgcccggc
cctgcgcccccctggaactcctgggcttccagctcccgccgctcccagaactgcgcctgcg
caacaatggccacagtgtgcaactgaccctgcctcctgggctagagatggctctgggtcc
cgggcgggagtaccgggctctgcagctgcatctgcactgggggggctgcaggtcgtccggg
ctcggagcacactgtggaaggccaccgtttccctgccgagatccacgtggttcacctcag
caccgcctttgccagagttgacgaggccttggggcgcccgggaggcctggccgtgttggc
cgcctttctggaggagggcccggaagaaaacagtgcctatgagcagttgctgtctcgctt
ggaagaaatcgctgaggaaggctcagagactcaggtcccaggactggacatatctgcact
cctgccctctgacttcagccgctacttccaatatgaggggtctctgactacaccgccctg
tgcccagggtgtcatctggactgtgtttaaccagacagtgatgctgagtgctaagcagct
ccacaccctctctgacaccctgtggggacctggtgactctcggctacagctgaacttccg
agcgacgcagcctttgaatgggcgagtgattgaggcctccttccctctggtgccccgtgg
ttcccaccaccatcaccaccatcaccactaaaagcttggcgaa

Amino acid sequence

ARSMAPLCPSPWLPLLIPAPAPGLTVQLLLSLLLLMPVHPQSHWRYGGDPPWPRVSPACA
GRFQSPVDIRPQLAAFCPALRPLELLGFQLPPLPELRLRNNGHSVQLTLPPGLEMALGPG
REYRALQLHLHWGAAGRPGSEHTVEGHRFPAEIHVVHLSTAFARVDEALGRPGGLAVLAA
FLEEGPEENSAYEQLLSRLEEIAEEGSETQVPGLDISALLPSDFSRYFQYEGSLTTPPCA
QGVIWTVFNQTVMLSAKQLHTLSDTLWGPGDSRLQLNFRATQPLNGRVIEASFPLVPRGS
HHHHHHHHKLGE

CAIX-N-His

Extras, BglII, optimized signal sequence, 8xHIS, thrombin, CA-domain, STOP, HindIII, extras

Nucleotide sequence

ggccAGATCTatggctcccctgtgcccctcccctggctgcccctgctgatccccgctcc
cgctcccggtctgaccgtgcagctgctgctgtccctgctgctgctgatgcccgtgcaccc
ccagcgtctgcaccaccatcaccaccatcaccactggtgccccgtggttcccacaggga
caaagaaggggatgaccagagtcattggcgctatggaggcgacccgccctggccccgggt
gtccccagcctgcgcgggccgcttccagtccccggtggatatccgcccccagctcgccgc
cttctgcccggccctgcgcccccctggaactcctgggcttccagctcccgccgctcccaga
actgcgcctgcgcaacaatggccacagtgtgcaactgaccctgcctcctgggctagagat
ggctctgggtcccgggcgggagtaccgggctctgcagctgcatctgcactgggggctgc
aggtcgtccgggctcggagcacactgtggaaggccaccgtttccctgccgagatccacgt
ggttcacctcagcaccgcctttgccagagttgacgaggccttggggcgcccgggaggcct
ggccgtgttggccgcctttctggaggagggcccggaagaaaacagtgcctatgagcagtt
gctgtctcgcttggaagaaatcgctgaggaaggctcagagactcaggtcccaggactgga
catatctgcactcctgccctctgacttcagccgctacttccaatatgaggggtctctgac
tacaccgccctgtgcccagggtgtcatctggactgtgtttaaccagacagtgatgctgag
tgctaagcagctccacaccctctctgacaccctgtggggacctggtgactctcggctaca
gctgaacttccgagcgacgcagcctttgaatgggcgagtgattgaggcctccttctaaaa
gcttggcgaa

Amino acid sequence

ARSMAPLCPSPWLPLLIPAPAPGLTVQLLLSLLLLMPVHPQRLHHHHHHHLVPRGSHRD
KEGDDQSHWRYGGDPPWPRVSPACAGRFQSPVDIRPQLAAFCPALRPLELLGFQLPPLPE
LRLRNNGHSVQLTLPPGLEMALGPGREYRALQLHLHWGAAGRPGSEHTVEGHRFPAEIHV
VHLSTAFARVDEALGRPGGLAVLAAFLEEGPEENSAYEQLLSRLEEIAEEGSETQVPGLD
ISALLPSDFSRYFQYEGSLTTPPCAQGVIWTVFNQTVMLSAKQLHTLSDTLWGPGDSRLQ
LNFRATQPLNGRVIEASFKLGE

CAIX-Avd

Extras, BglII, Avidin signal, Avidin, GGSGGS, thrombin, CA-domain, STOP, HindIII, extras

Nucleotide sequence

ggccagatctatggtgcacgcaacctccccgctgctgctgctgctgctcagcctggc
tctggtggctcccggcctctctgccagaaagtgctcgctgactgggaaatggaccaacga
tctgggctccaacatgaccatcggggctgtgaacagcagaggtgaattcacaggcaccta
catcacagccgtaacagccacatcaaatgagatcaaagagtcaccactgcatgggacaca
aaacaccatcaacaagaggacccagcccacctttggcttcaccgtcaattggaagttttc
agagtccaccactgtcttcacgggccagtgcttcatagacaggaatgggaaggaggtcct
gaagaccatgtggctgctgcggtcaagtgttaatgacattggtgatgactggaaagctac
cagggtcggcatcaacatcttcactcgcctgcgcacacagaaggaggcggttccggtgg
ctccctggtgccccgtggttccacagggacaaagaaggggatgaccagagtcattggcg
ctatggaggcgacccgccctggccccgggtgtccccagcctgcgcgggccgcttccagtc
cccggtggatatccgcccccagctcgccgccttctgcccggccctgcgcccccctggaact
cctgggcttccagctcccgccgctcccagaactgcgcctgcgcaacaatggccacagtgt
gcaactgaccctgcctcctgggctagagatggctctgggtcccgggcgggagtaccgggc
tctgcagctgcatctgcactgggggctgcaggtcgtccgggctcggagcacactgtgga
aggccaccgtttccctgccgagatccacgtggttcacctcagcaccgcctttgccagagt
tgacgaggccttggggcgcccgggaggcctggccgtgttggccgcctttctggaggaggg
cccggaagaaaacagtgcctatgagcagttgctgtctcgcttggaagaaatcgctgagga
aggctcagagactcaggtcccaggactggacatatctgcactcctgccctctgacttcag
ccgctacttccaatatgaggggtctctgactacaccgccctgtgcccagggtgtcatctg
gactgtgtttaaccagacagtgatgctgagtgctaagcagctccacaccctctctgacac
cctgtggggacctggtgactctcggctacagctgaacttccgagcgacgcagcctttgaa
tgggcgagtgattgaggcctccttctaaaagcttggcgaa

Amino acid sequence

ARSMVHATSPLLLLLLLSLALVAPGLSARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTY
ITAVTATSNEIKESPLHGTQNTINKRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVL
KTMWLLRSSVNDIGDDWKATRVGINIFTRLRTQKEGGSGGSLVPRGSHRDKEGDDQSHWR
YGGDPPWPRVSPACAGRFQSPVDIRPQLAAFCPALRPLELLGFQLPPLPELRLRNNGHSV
QLTLPPGLEMALGPGREYRALQLHLHWGAAGRPGSEHTVEGHRFPAEIHVVHLSTAFARV
DEALGRPGGLAVLAAFLEEGPEENSAYEQLLSRLEEIAEEGSETQVPGLDISALLPSDFS
RYFQYEGSLTTPPCAQGVIWTVFNQTVMLSAKQLHTLSDTLWGPGDSRLQLNFRATQPLN
GRVIEASFKLGE

CAIX-PG

Extras, BglII, optimized signal sequence, 8xHIS, thrombin, PG-domain, STOP, HindIII, extras

Nucleotide sequence

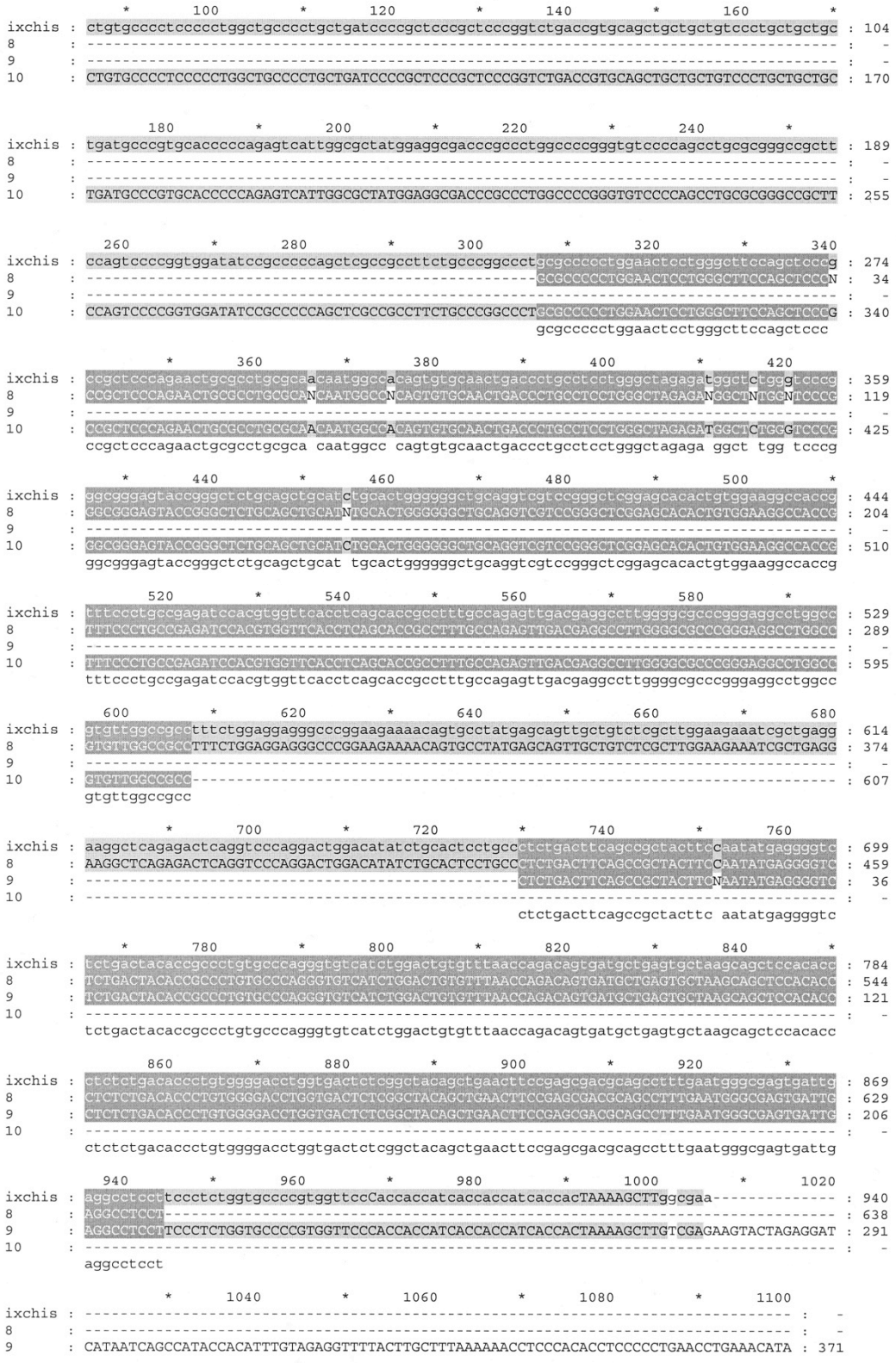
ggccAGATCTatggctcccctgtgcccctcccctggctgcccctgctgatccccgctcc
cgctcccggtctgaccgtgcagctgctgctgtccctgctgctgctgatgcccgtgcaccc
ccagcgtctgcaccaccatcaccaccatcaccactggtgccccgtggttcccagaggtt
gccccggatgcaggaggattccccttgggaggaggctcttctggggaagatgacccact
gggcgaggaggatctgcccagtgaagaggattcacccagagaggaggatccacccggaga
ggaggatctacctggagaggaggatctacctggagaggaggatctacctgaagttaagcc
taaatcagaagaagagggctccctgaagttagaggatctacctTAAAAGCTT ggcgaa

Amino acid sequence

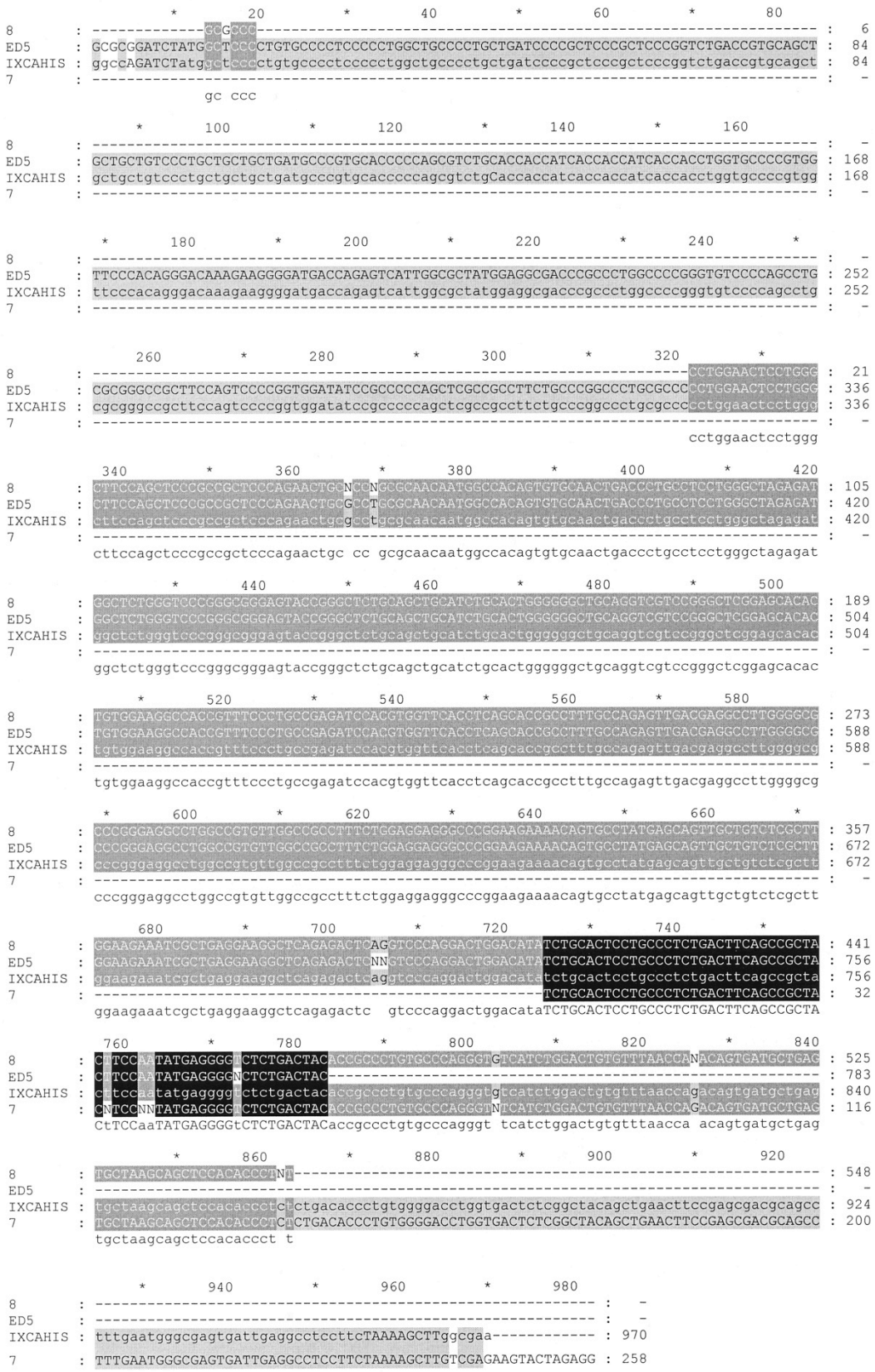
ARSMAPLCPSPWLPLLIPAPAPGLTVQLLLSLLLLMPVHPQRLHHHHHHHHLVPRGSQRL
PRMQEDSPLGGGSSGEDDPLGEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPEVKP
KSEEEGSLKLEDLPKLGE

10 Appendix 2.

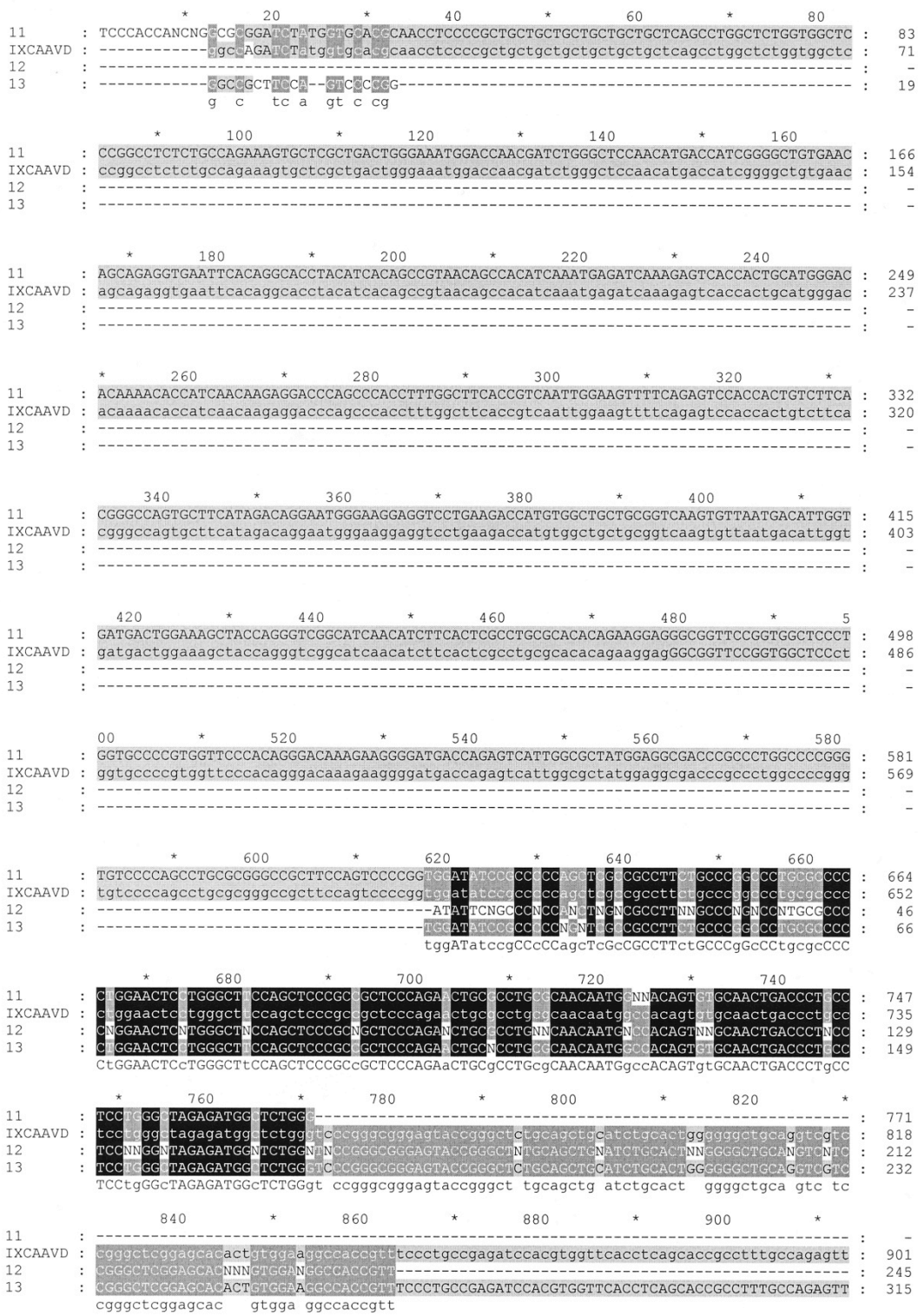
CAIX-C-His

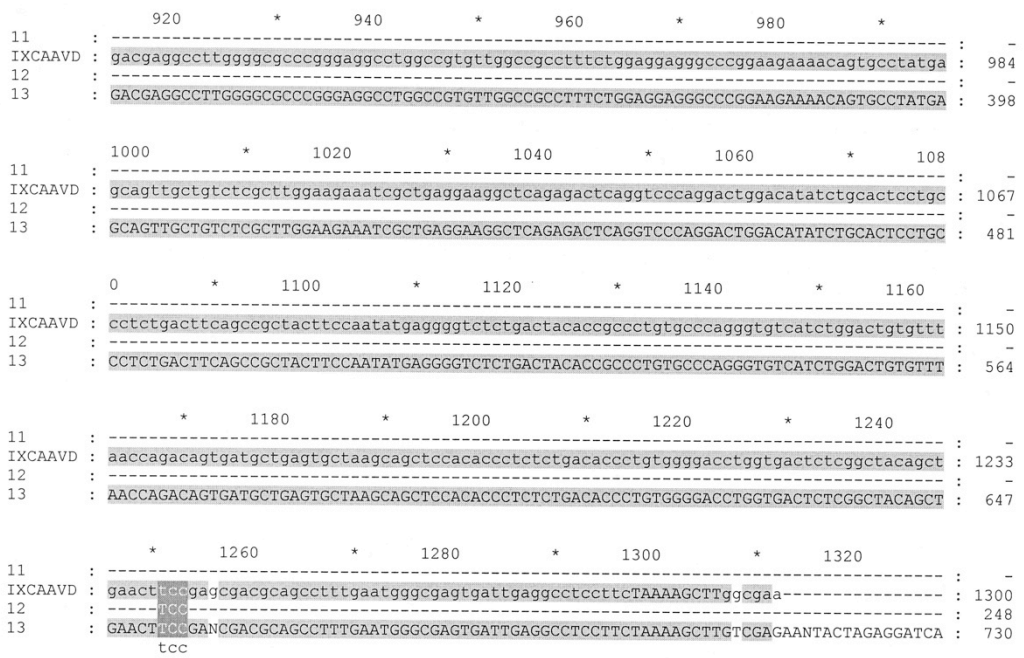


CAIX-N-His



CAIX-Avd




920 * 940 * 960 * 980 *
11 : - :
IXCAAVD : gacgaggccttggggcgcccgggaggcctggccgtgttggccgccttttctggaggagggcccggaagaaaacagtgcctatga : 984
12 : - :
13 : GACGAGGCCTTGGGGCGCCCGGGAGGCCTGGCCGTGTTGGCCGCCTTTCTGGAGGAGGGCCCGGAAGAAAACAGTGCCTATGA : 398
1000 * 1020 * 1040 * 1060 * 108
11 : - :
IXCAAVD : gcagttgctgtctcgcttggaagaaatcgctgaggaaggctcagagactcaggtcccaggactggacatatctgcactcctgc : 1067
12 : - :
13 : GCAGTTGCTGTCTCGCTTGGAAGAAATCGCTGAGGAAGGCTCAGAGACTCAGGTCCCAGGACTGGACATATCTGCACTCCTGC : 481
0 * 1100 * 1120 * 1140 * 1160
11 : - :
IXCAAVD : cctctgacttcagccgctacttccaatatgaggggtctctgactacaccgccctgtgcccagggtgtcatctggactgtgttt : 1150
12 : - :
13 : CCTCTGACTTCAGCCGCTACTTCCAATATGAGGGGTCTCTGACTACACCGCCCTGTGCCCAGGGTGTCATCTGGACTGTGTTT : 564
* 1180 * 1200 * 1220 * 1240
11 : - :
IXCAAVD : aaccagacagtgatgctgagtgctaagcagctccacaccctctctgacaccctgtggggacctggtgactctcggctacagct : 1233
12 : - :
13 : AACCAGACAGTGATGCTGAGTGCTAAGCAGCTCCACACCCTCTCTGACACCCTGTGGGGACCTGGTGACTCTCGGCTACAGCT : 647
* 1260 * 1280 * 1300 * 1320
11 : - :
IXCAAVD : gaacttccgagcgacgcagcctttgaatgggcgagtgattgaggcctccttctTAAAAGCTTggcgaa---------- : 1300
12 : -----TCC---------- : 248
13 : GAACTTCCGANCGACGCAGCCTTTGAATGGGCGAGTGATTGAGGCCTCCTTCTAAAAGCTTGTCGAGAANTACTAGAGGATCA : 730
tcc

CAIX-PG