Tampere University

MAARIT PATRIKAINEN

Carbonic Anhydrase VI

Functional characterization in mouse and zebrafish model organisms

Tampere University Dissertations 134

Tampere University Dissertations 134

MAARIT PATRIKAINEN

Carbonic Anhydrase VI

Functional characterization in mouse and zebrafish model organisms

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty Council of the Faculty of Medicine and Health Technology
of the Tampere University,
for public discussion in the auditorium F115
of the Arvo building, Arvo Ylpön katu 34, Tampere,
on 11th October 2019, at 12 o'clock.

ACADEMIC DISSERTATION
Tampere University, Faculty of Health Technology
Finland

Responsible supervisor or/and Custos
Professor Seppo Parkkila
Tampere University
Finland

Pre-examiner(s)
Assistant Professor
Mikko Metsä-Ketelä
University of Turku
Finland

Associate Professor
Mohammed Al-Haroni
UiT the Arctic University of Norway
Norway

Opponent(s)
Professor Juha Tuukkanen
University of Oulu
Finland

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

Cover design: Roihu, Inc.

ISBN 978-952-03-1257-2 (print)
ISBN 978-952-03-1258-9 (pdf)
ISSN 2489-9860 (print)
ISSN 2490-0028 (pdf)
http://urn.fi/
URN:ISBN:978-952-03-1258-9

PunaMusta Oy – Yliopistopaino
Tampere 2019

TO MY FAMILY

"Everybody is genius.
But if you judge a fish by its ability to climb a tree,
It will live its whole life believing that it is stupid."
-Albert Einstein

ACKNOWLEDGEMENTS

This PhD thesis was carried out at the Faculty of Medicine and Health Technology, Tampere University (Tampere, Finland) during 2013-2019 under the supervision of Professor Seppo Parkkila.

I wish to express my gratitude to my supervisor Professor Seppo Parkkila (MD, PhD) for providing me the opportunity to do scientific research in the field of anatomy and tissue biology and for giving me a place in your CA group.

I would like to thank my external reviewers of my thesis, Associate Professor Mikko Metsä-Ketelä (PhD) and Associate Professor Mohammed Al-Haroni (DDS, PhD), for their valuable comments and kind words. I would also thank Professor Juha Tuukkanen (DDS, PhD) for giving me the honor to be my opponent.

I am extremely grateful for my awesome teacher Martti Tolvanen (PhD) who has been there for me since I started my bioinformatics studies and have introduced me to the world of CAs, and been my mentor during my PhD journey. I want to thank you for introducing me Turku and all those scientific and non-scientific discussions we have had over these years. Next, I like to thank former and present CA group members who have been working with me during these years: Peiwen Pan (PhD), Ashok Aspatwar (PhD), Harlan Barker (MSc), Linda Urbanski (MSc(Tech)), Sami Purmonen (PhD), Heimo Syvälä (PhD), Susanna Haapanen (BM), Reza Zolfaghari Emameh (PhD), Alma Yrjänäinen (MSc), Leo Syrjänen (MD, PhD), and Heini Kallio (PhD). I am grateful for being able to work with two excellent laboratory technicians Aulikki Lehmus and Marianne Kuuslahti. Your expertise, insight and help have been crucial for carrying out the laboratory work for this thesis.

I want to thank all the other collaborators and co-authors for their valuable contribution: Chaba Ortutay (PhD), Vootele Voikar (PhD), Natalia Kulesskaya (PhD), Janne Jänis (PhD), Mikko Laitaoja (PhD), Vesa Hytönen (PhD), Latifeh Azizi (MSc), Prajwol Manandhar (MSc), Edit Jáger (PhD), Daniela Vullo (PhD), Sampo Kukkurainen (MSc), Mika Hilvo (PhD), Claudiu Supuran (PhD), people in the Zebrafish Facility (Tampere University), and people in Laboratory animal Centre in Oulu. I want to give special thanks to all of you I spend my coffee breaks with, bringing lots of joy, you bright up my days, and those that I have forgotten to mention here.

I want to give special thanks to my family and friends: “Vanhemmilleni, Antille ja Railille suuret kiitokset kaikesta avusta ja tuesta, jota ilman en olisi tässä nyt. Olette antaneet minulle evät elämään ja vanhemmuuteen, sekä tukeneet minua hyvinä ja haastavina aikoina. Siskoani Minnaa haluan kiittää kaikista unohtumattomista seikkailuista ja tuestasi johon voin aina luottaa. Olet minulle rakas ja erittäin tärkeä esikuva. Rakkaalle tyttärelleni Miralle erityiskiitos kaikista halauksista ja hymyistä, sekä tuesta jota olet minulle antanut vuosien varrella. Sanat eivät riitä kertomaan kuinka tärkeä olet minulle. Lopuksi haluan kiittää ystäviäni joista on tullut minun toinen perheeni: Hanna, Heidi, Elina, Päivi ja Tuija, tuhannet kiitokset antamastanne tuesta. On ollut ilo ja kunnia jakaa nämä vuodet teidän kanssanne.”

This thesis work was financially supported by the Marjatta Melkas-Rusanen and Anneli Melkas Memorial Foundation of the Finnish Cultural Foundation Pirkanmaa Regional Fund, Tampere Tuberculosis Foundation, Sigrid Juselius Foundation, Jane & Aatos Erkko Foundation, Academy of Finland, VTR-Funding of Tampere University Hospital, and Doctoral Programme on the Faculty of Medicine and Health Technology, Tampere University. All of the financial supporters are greatly acknowledged.

Tampere, September 2019

Maarit Patrikainen

Contents

LIST OF ORIGINAL COMMUNICATIONS

This thesis is based on the following original communications, which are referred to in the text by their Roman numerals (I-III).

I **Patrikainen M**, Pan P, Kulesskaya N, Voikar V, Parkkila S (2014): The role of carbonic anhydrase VI in bitter taste perception: Evidence from the *Car6*$^{-/-}$ mouse model. J Biomed Sci 21(1):82.

II **Patrikainen MS**, Pan P, Barker HR, Parkkila S (2016): Altered gene expression in the lower respiratory tract of *Car6* $^{-/-}$ mice. Transgenic Res 25(5):649-64.

III **Patrikainen MS***, Tolvanen MEE*, Aspatwar A*, Barker H, Ortutay C, Jänis J, Laitaoja M, Hytönen V, Azizi L, Manandhar P, Jáger E, Vullo D, Kukkurainen S, Hilvo M, Supuran CT, Parkkila S (2017) Identification and characterization of a novel zebrafish (*Danio rerio*) pentraxin–carbonic anhydrase. PeerJ 5:e4128.

*equal contribution

ABBREVIATIONS

AZA	Acetazolamide (CA inhibitor)
APH	Amphipathic helix
ATP	Adenosine triphosphate
BALT	Bronchus-associated lymphoid tissue
Bluo-gal	5-Bromo-3-indolyl β-D-galactopyranoside
BSA	Bovine serum albumin
BSA-PBS	Bovine serum albumin in phosphate-buffered saline
CA	Carbonic anhydrase
CA VI	Carbonic anhydrase VI protein
ca6	*Carbonic anhydrase 6* gene (*Danio rerio*, Zebrafish)
CA6	*Carbonic anhydrase 6* gene (*Homo sapiens*, Human)
Calca	*Calcitonin/calcitonin-related polypeptide, alpha*
Car6	*Carbonic anhydrase 6* gene (*Mus musculus*, Mouse)
CA-RP	Carbonic anhydrase-related protein
CALT	Conjunctiva-associated lymphoid tissue
cAMP	Cyclic adenosine monophosphate
cDNA	Complementary deoxyribonucleic acid
cGMP	Cyclic guanosine monophosphate
CHOP	CCAAT/Enhancer-binding protein homologous protein
CID-MS/MS	Collision Induced Dissociation with Tandem Mass Spectrometry
CRP	C-reactive protein
DAB	3,3'-diaminobenzidine
DALT	Duct-associated lymphoid tissue
DAVID	Database for Annotation, Visualization and Integrated Discovery
Dbp	*D site albumin promoter binding protein* gene
Dcpp1	*Demilune cell and parotid protein 1* gene
DMFT	Decayed, missing, and filled teeth
DNA	Deoxyribonucleic acid
DOG1	Discovered on GIST1 (Calcium-dependent chloride channel protein)
DPF	Day(s) post fertilization
EDTA	2,2',2",2'"-(Ethane-1,2-diyldinitrilo)tetraacetic acid

ER Endoplasmic reticulum
Faim *Fas apoptotic inhibitory molecule 3* gene
FC Fold change
FELASA Federation for Laboratory Animal Science Associations
FFPE Formalin-fixed paraffin-embedded
FPKM Fragments per kilobase of transcript per million mapped reads
G3PDH Glyceraldehyde-3-phosphate dehydrogenase
GABA gamma-Aminobutyric acid
GalNAc N-acetylgalactosamine
GALT Gut-associated lymphoid tissue
gapdh *glyceraldehyde 3-phosphate dehydrogenase* gene (*Danio rerio*, Zebrafish)
Gdpd3 *Glycerophosphodiester phosphodiesterase domain containing 3* gene
GI Gastrointestinal
GIALT Gill-associated lymphoid tissue
GIST Gastrointestinal stromal tumor
GlcNAc2(Fuc)Man3 N-Diacetylglucosamine core-fucosylated trimannose
GlcNAc2Man3 N-Diacetylglucosamine trimannose
GPCR G-protein-coupled receptors
GPR157 *G Protein-Coupled Receptor 157* gene
H3R8me2s Histone H3 dimethyl R8
HEPES 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HPF Hour(s) post fertilization
IgA Immunoglobulin A
IL-6 Interleukin 6
IL-12 Interleukin 12
IP_3 Inositol 1, 4, 5-trisphosphate
IPTG Isopropyl-β-D-thiogalactopyranoside
K_{cat} Turnover number, the maximum number of enzymatic reactions catalyzed per second
KD Knockdown
kDa Kilodalton, unit of molecular mass
K_m Michaelis constant, substrate concentration at which the rate of the enzyme reaction is half of the maximum
KO Knockout
LDALT Lacrimal duct-associated lymphoid tissue
LALT Larynx-associated lymphoid tissue
LB Luria broth

Lyz1	*Lysozyme 1* gene
mActb	*Mouse beta-actin* gene
MALT	Mucosa-associated lymphoid tissue
MCTD	Mixed connective tissue disease
MO	Morpholino-oligonucleotide
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
NALT	Nasopharynx-associated lymphoid tissue
NGS	Normal goat serum
Nppa	*Natriuretic peptide precursor type A* gene
NRS	Normal rabbit serum
OD	Optical density
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
pI	Isoelectric point, pH at which overall net charge of a protein is zero
pIgR	Polymeric immunoglobulin receptor
PKD2L1	Polycystic kidney disease 2-like 1 protein
PM	Polymyositis
PRM	Pattern-recognition molecule
PRMT5	N-methyltransferase 5
PTX	Pentraxin
RA	Rheumatoid arthritis
qRT-PCR	Quantitative Reverse-Transcription Polymerase Chain Reaction
RNA	Ribonucleic acid
RT	Room temperature
SALT	Skin-associated lymphoid tissue
SAP	Serum amyloid P
SC	Secretory component
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
Sftpc	*Surfactant associated protein C* gene
Sf9	*Spodoptera frugiperda* cell line
SIgA	Secretory Immunoglobulin A
SLC2A5	*Solute carrier family 2, facilitated glucose transporter member* 5 gene
SLC2A7	*Solute carrier family 2, facilitated glucose transporter member* 7 gene
SS	Sjögren's syndrome
SSc	Accompanying systemic sclerosis

sSS	Secondary Sjögren's syndrome
TAE	Tris-Acetate-Ethylenediaminetetra-acetic acid
T1R1	Taste receptor type 1 member 1
TAS2R38	Bitter receptor gene
TBS	Tris-buffered saline
TM	Transmembrane
TMH	Transmembrane helix
TRC	Taste receptor cell
Tricaine	Ethyl 3-aminobenzoate methane sulfonate salt
Tris	2-Amino-2-hydroxymethyl-propane-1,3-diol
Ucp1	*Uncoupling protein 1(mitochondrial, proton carrier)* gene
VLAD	Gene List Analysis and Visualization tool
WT	Wild Type

ABSTRACT

Carbonic anhydrases (CAs) efficiently catalyze the reversible hydration reaction of carbon dioxide ($CO_2 + H_2O \leftrightarrow HCO_3^- + H^+$). Major CA families belong to the α-, β-, or γ-classes and have different features in terms of both their overall structure and active site. Carbonic anhydrase VI (CA VI) is the only secreted isozyme in the α-CA enzyme family. In the 1970s, a novel taste-associated secreted protein was identified in human saliva and named Gustin. The same protein was observed in a sheep parotid gland and saliva in 1979 and was classified as a CA based on its protein sequence. Two decades later, in 1998, Gustin was proven to be identical to the CA VI enzyme. CA VI is one of the major protein constituents of human saliva and milk. The general aim of this study was to reveal new information about the expression and function of secretory CA VI.

We used two independent animal models in our studies focused on CA VI. First, a *Car6* knockout mouse (KO) model was utilized to evaluate the differences in taste modalities between CA VI-deficient and wild-type (WT) mice. Then, we further investigated by cDNA microarray whether the gene expression profiles differed between KO and WT mice with a focus on the trachea and lung. Second, we used zebrafish to investigate the role and distribution of CA VI in another vertebrate animal model. The recombinant zebrafish CA VI protein was produced in insect cells. The protein was purified using chromatography methods and characterized by static and dynamic light scattering (SLS and DLS) analyses and by mass spectrometry (MS). Verification of glycosylation sites and glycan structures was carried out using Collision Induced Dissociation with Tandem Mass Spectrometry (CID-MS/MS). Sequence analysis was performed by various bioinformatics tools. The functional significance of the CA VI enzyme to zebrafish was studied by silencing the gene encoding the enzyme, and its effect on the phenotype was examined.

The comparison of taste modalities showed that *Car6*$^{-/-}$ mice significantly preferred a0.003 mM quinine solution to water, whereas WT mice preferred water. The microarray data analysis results showed a number of differentially expressed genes in the trachea and lung when comparing the WT and *Car6* $^{-/-}$ mice groups. A Gene List Analysis and Visualization tool (VLAD) analysis resulted in changes in a

metabolic process, biological regulation, single-organism process, and an immune response in mucosal-associated lymphoid tissue.

Sequence analysis of zebrafish CA VI showed that there was an additional pentraxin (PTX) protein attached to the C-terminus of CA VI. Light scattering combined with gel filtration analysis indicated oligomeric assembly for the protein, with a pentameric configuration being the most plausible form. Both the CA VI and PTX domains contained either intra- or interpeptide disulfide bonds. Localization analysis of CA VI-PTX in various zebrafish tissues showed the strongest positive staining signal on cell surfaces in the skin, heart, gills, and swim bladder. Underdeveloped or deflated swim bladders of CA VI-PTX knockdown (KD) larvae were observed at 4 days post fertilization (dpf), whereas fish gained the ability to inflate their swim bladder and started to swim normally at 5 dpf.

Our KO mouse model confirms that CA VI is involved in bitter taste perception. As one of the major protein constituents of saliva and milk, CA VI may be one of the factors that contributes to the avoidance of bitter, potentially harmful, substances. Several findings on zebrafish CA VI give novel insights into the evolution, structure, and function of this unique CA form.

TIIVISTELMÄ

Hiilihappoanhydraasit (Carbonic anhydrase, CA) katalysoivat tehokkaasti käänteistä hiilidioksidin ja veden reaktiota bikarbonaatiksi ja protoniksi ($CO_2 + H_2O \leftrightarrow HCO_3^- + H^+$). Useimmat CA-entsyymit kuuluvat α-, β-, tai γ- luokkaan, jotka eroavat toisistaan sekä aktiivisen keskuksen että ulkoisen rakenteen osalta. CA VI on ainoa α-luokan erittyvä entsyymi. Se havaittiin ihmisen syljestä ja yhdistettiin makuaistimukseen ensimmäistä kertaa 1970-luvulla ja nimettiin 'gustiini' -entsyymiksi. Vuonna 1979 CA-entsyymi löydettiin lampaan sylkirauhasesta ja syljestä. Ihmisen syljestä eristetty entsyymi nimettiin CA VI:ksi vuonna 1987. Vuonna 1998 osoitettiin, että 'gustiini' ja CA VI ovat keskenään identtisiä. Ihmisellä CA VI-entsyymiä erittyy syljen lisäksi myös maitoon. Oman tutkimukseni tarkoituksena on ollut hankkia uutta tietoa erittyvän CA-entsyymin ilmentymisestä ja toiminnasta.

Tutkimuksessa käytettiin kahta koe-eläinmallia: *Car6* poistogeenisiä hiiriä sekä seeprakaloja, joiden vastaava geeni oli hiljennetty morfoliinotekniikalla.

Car6 poistogeenistä hiirimallia käytettiin sekä makuaistimuksen että geenien ilmentymisen tutkimiseen. Tuloksia verrattiin normaalien villin tyypin (WT) hiirten tuloksiin. Tutkimuksessa tarkasteltiin cDNA-mikrosirutekniikan avulla, miten CA VI-entsyymin poisto vaikuttaa geenien ilmentymiseen henkitorvessa ja keuhkoissa. Kahden hiiriryhmän välisiä eroja selvitettiin tilastotieteellisillä menetelmillä.

Hyönteissoluissa geenitekniikan avulla tuotetun seeprakalan CA VI-proteiinin rakenteen kartoitukseen käytettiin staattista (SLS) ja dynaamista (DLS) valonsirontamenetelmää, joita ennen proteiini puhdistettiin kromatografiamenetelmillä. Massaspektrometriaa varten proteiinirakenne pilkottiin trypsiinin avulla. Proteiinirakenne, siinä esiintyvät sokerointikohdat ja sokeriosien rakenteet analysoitiin CID-MS/MS-massaspektrometrialla.

Tutkimuksen eri vaiheissa tuloksina saadut nukleotidi- ja aminohapposekvenssit analysoitiin käyttämällä lukuisia bioinformatiikan työkaluja. CA VI- entsyymin merkitystä seeprakalan kehityksessä tutkittiin hiljentämällä *ca6*-geeni alkioista, minkä jälkeen tarkasteltiin kalojen kehitystä ja uintikäyttäytymistä.

Makuaistimuskokeen tulokset osoittivat *Car6*$^{-/-}$ -hiirten suosivan veden juomisen sijaan 0.003 mM kiniiniliuosta, kun taas WT hiiret suosivat vettä.

Mikrosirumenetelmällä saadun datan analyysi osoitti geenien ilmentymiseroja sekä henkitorvessa että keuhkoissa *Car6*$^{-/-}$ poistogeenisten ja WT hiirten välillä. VLAD-analyysi osoitti muutoksia aineenvaihdunnan prosesseissa, biologisessa säätelyssä, yhden organismin prosesseissa sekä limakalvoon liittyvän kudoksen immuunivasteessa.

Sekvenssianalyysi osoitti, että seeprakalan CA VI:n C-terminaalipäähän on liittynyt pentraksiiniproteiini. Valonsironta- ja geelisuodatusanalyysien tuloksena proteiini ilmenee oligomeerina, ja molekyylin koon perusteella sen voidaan arvioida esiintyvän pentameerina. Sekä CA VI- että PTX-domeeneissa on rikkisiltoja. Seeprakalan CA VI-PTX -kudosvärjäys osoitti positiivista signaalia solun pinnalta ihossa, sydämessä, kiduksissa sekä uimarakossa. *ca6* geenin hiljentämisen jälkeen neljän päivän ikäiset poikaset eivät pystyneet uimaan normaalisti lähellä veden pintaa joko uimarakon hidastuneen kehittymisen tai sen täyttymiseen liittyvien ongelmien vuoksi, kun taas WT poikaset uivat normaalisti. Viiden päivän iässä *ca6*-hiljennettyjen poikasten uimarakko täyttyi ja uintikyky tuli normaaliksi.

Poistogeenisen hiirimallin avulla osoitimme, että CA VI osallistuu karvaan maun aistimukseen. CA VI, jota erittyy runsaasti sylkeen ja maitoon voi olla yksi tekijä, joka auttaa välttämään karvaan makuisia ja mahdollisesti myrkyllisiä aineita. Seeprakalan CA VI-entsyymiin liittyvien tutkimusten avulla saatiin runsaasti uutta tietoa CA-entsyymien evoluutiosta sekä CA VI-PTX-proteiinin ilmentymisestä, rakenteesta ja toiminnasta.

1 INTRODUCTION

Carbonic anhydrases (CAs) efficiently catalyze the reversible hydration reaction of carbon dioxide $CO_2 + H_2O \leftrightarrow HCO_3^- + H^+$ (Chegwidden & Carter, 2000, Gilmour, 2010). CAs may have been among the earliest enzymes to appear, and their long convergent evolution has resulted in seven genetically distinct unrelated enzyme families: α-, β-, γ-, δ-, ζ-, η-, and θ-CAs (Supuran, C. T., 2010, Del Prete et al, 2014, Supuran, Claudiu T. & De Simone, 2015, Kikutani et al, 2016). The major CA families belong to the α-, β-, or γ-classes and have different structures including their active sites. Out of the 16 α-CA isoforms, CA-related proteins (CA-RPs) VIII, X and XI lack catalytic activity (Tashian et al, 2000, Aspatwar et al, 2010, Gilmour, 2010).

The carbonic anhydrase VI (CA VI) is the only secreted isozyme in the α-CA enzyme family (Fernley et al, 1979, Murakami, H. & Sly, 1987). In the 1970s, a novel taste-associated secreted protein was identified in human saliva and named Gustin (Henkin et al, 1975). The same protein was observed in the sheep parotid gland and saliva in 1979 but was classified as a CA (Fernley et al, 1979). Two decades later, in 1998, Gustin was proven to be identical to the CA VI enzyme (Thatcher et al, 1998). CA VI is one of the major protein constituents of human saliva and milk (Kadoya et al, 1987, Parkkila et al, 1990, Parkkila et al, 1993).

Henkin´s group linked Gustin/CA VI to the regulation of taste function in 1981 (Shatzman & Henkin, 1981). Later, Barbarossa and coworkers found a link between bitter taste modality and *CA6*. They first reported a polymorphism in the *CA6* gene (rs2274333 A/G) that contributed to 6-n-propylthiouracil taster status (Padiglia et al, 2010). They also showed that the alterations in bitter taste function are due to polymorphic changes in the bitter receptor gene (*TAS2R38*) and *CA6* gene and require contributions from other still unknown factors (Calo et al, 2011). CA VI is located in a thin oral biofilm, called the enamel pellicle, where it is strategically located to perform its functions on the tooth surface (Leinonen, J. et al, 1999). Kivelä and coworkers suggested that salivary CA VI plays a role in the natural defense systems against dental caries because the concentration of CA VI exhibited a negative correlation with the number of decayed, missing, and filled teeth (DMFT index) in individuals with poor oral hygiene (Kivela et al, 1999b).

CA VI may play a role in immune defense systems. A recent article showed that Immunoglobulin A (IgA) anti-CA VI autoantibodies are frequently seen in patients with long-term Sjögren's Syndrome (SS) and suggested further studies to determine whether these antibodies could be used in the diagnosis of secondary SS (sSS) accompanying systemic sclerosis (SSc), sSS in rheumatoid arthritis (RA), and polymyositis (PM) and mixed connective tissue disease (MCTD) (Shen et al, 2012, De Langhe et al, 2017). In 1999, Sok and coworkers discovered a novel form of murine CA VI, CA VI-b, which is a highly responsive enzyme to certain forms of cell stress (Sok et al, 1999). Recently, another group found that CA VI-b is directly connected to the innate immune response by selectively inducing cytokine IL-12 production through protein arginine N-methyltransferase 5 (PRMT5) and regulating symmetric dimethylation at Arg-9 histone H3 dimethyl R8 (H3R8me2s) modification, independent of its CA activity (Xu, J. et al, 2017)

This thesis aimed to increase our knowledge about the distribution, function, and structure of the secretory CA VI enzyme. We utilized various methods, including the tools of functional genomics and protein chemistry, to investigate the features of this unique enzyme in mouse (*Mus musculus*) and zebrafish (*Danio rerio*) animal models.

2 REVIEW OF THE LITERATURE

2.1 General aspects of carbonic anhydrases

In 1932, an enzyme, carbonic anhydrase (CA) (EC 4.2.1.1), was discovered in red blood cells as a result of the realization that the non-catalyzed HCO_3^- dehydration rate was too low to support CO_2 excretion during the time blood spent on the gas exchange surface (Forster, 2000). Gaseous stage CO_2 is generated in most metabolic oxidative processes and must be converted to a soluble form to prevent accumulation and damage to cells and other organelles (Supuran, C. T., 2018). CAs efficiently catalyze the reversible hydration reaction of carbon dioxide ($CO_2 + H_2O \leftrightarrow HCO_3^- + H^+$). CAs are among the fastest enzymes known (the turnover number or K_{cat} of some CA isoforms exceeds 1×10^6 s^{-1}) and take part in a remarkable range of physiological processes, such as calcification, photosynthesis, respiration, ionic, acid-base and fluid balance, metabolism, and cell growth (Chegwidden & Carter, 2000, Gilmour, 2010).

CAs may have been among the earliest enzymes to appear, and during their long evolutionary history, CA genes have undergone many rounds of duplication, resulting in the convergent evolution of seven genetically distinct unrelated enzyme families: α-, β-, γ-, δ-, ζ-, η-, and θ-CAs (Supuran, C. T., 2010, Del Prete et al, 2014, Supuran, Claudiu T. & De Simone, 2015, Kikutani et al, 2016). All CAs are metalloenzymes and involved in a two-step catalytic mechanism. A substitution reaction with CO_2 and the metal ion-bound OH^- ion that yields a coordinated HCO_3^- ion is shown in E.1. A water molecule is subsequently displaced from the metal ion (Figure 1 A-C). An addition reaction, in which OH^- regeneration involves the transfer of an H^+ from the metal ion-bound water molecule to the solvent, is shown in E.2 (Figure 1 C-A) (Lindskog & Silverman, 2000, Supuran, Claudiu T. & De Simone, 2015).

$$EM^{2+} - OH^- + CO_2 \rightleftarrows EM^{2+} - CHO_3^- \overset{H_2O}{\rightleftarrows} EM^{2+} - H_2O + CHO_3^- \quad \text{(E. 1)}$$

$$EM^{2+} - H_2O + B \rightleftarrows EM^{2+} - OH^- + BH^+ \quad \text{(E. 2)}$$

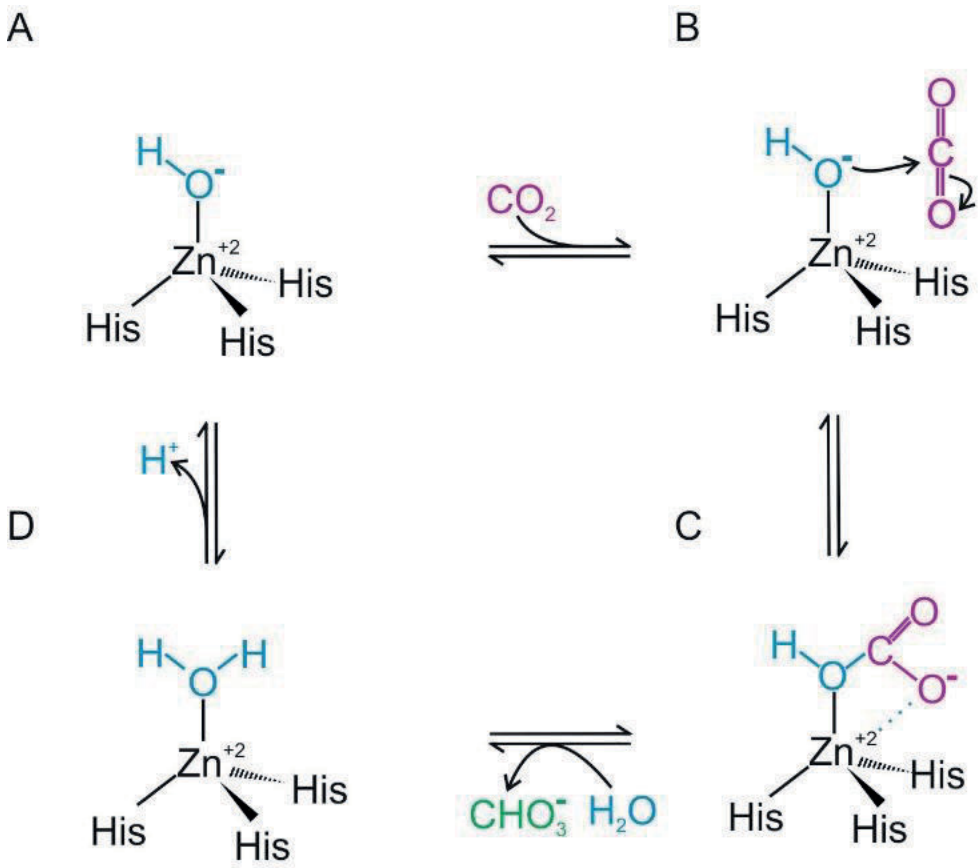

Figure 1. Schematic illustration of the mechanism of the alpha-CA active site (Supuran, C. T., 2008a). This figure has been modified by Maarit Patrikainen.

Major CA families belong to the α-, β-, or γ-classes and have different features in their overall structures, and importantly, their active sites have different configurations. Three histidine residues and an oxygen from a water molecule, which coordinate the zinc atom binding, form the active site of α-CAs. The active site of β-CAs has two cysteine residues and one histidine residue coordinating the zinc atom. γ-CAs differ from the other CA families in that they can use either zinc, cobalt or iron in their active site (Ferry, 2010). The metal ion is coordinated by two histidine residues and a glutamine residue together with the water molecule. Figure 2 shows examples of the overall three-dimensional structures and active sites of major CA families.

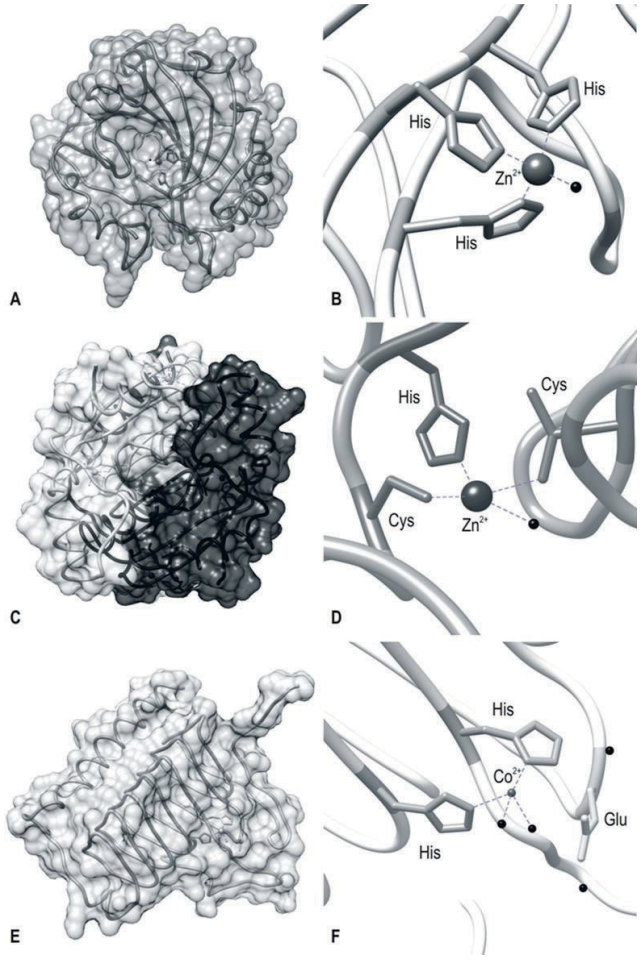

Figure 2. Structures and metal ion coordination in the major CA enzyme families. The above structures of 3D molecules were retrieved from the PDB (Berman et al, 2000), and the molecular graphics were generated with the UCSF Chimera package (Pettersen et al, 2004), developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311); the graphics were modified by Maarit Patrikainen. (A) The structure of α-class CA II (PDB ID: 2ILI, (Fisher et al, 2007)) and its active site (B), where the zinc ion is coordinated by three histidine residues and a water molecule/hydroxide ion (black sphere). (C) The dimeric mycobacterial enzyme Rv1284 (PDB ID: 4YF4, (Nienaber et al, 2015)) represents the structure of a β-class CA. Two cysteine residues and one histidine together with the water molecule/hydroxide ion coordinate the zinc ion in the active site of β-CAs (D). (E) The γ-class CA structure of a methanogenic archaeon, *Methanosarcina thermophila* (PDB ID: 1QRE, (Iverson et al, 2000)), and its active site (F), where two histidine residues and glutamine, together with a water molecule/hydroxide ion, coordinate a cobalt ion.

The β-CAs are found in most species belonging to the Archaea and Bacteria domains and probably all species of plants and fungi among Eukarya (Hewett-Emmett, 2000). While α-CAs are typically found as monomers and γ-CAs as trimers, β-CAs are found in many oligomerization states. The crystal structures of dimeric, tetrameric and even octameric β-CAs have been reported in the literature (Kimber & Pai, 2000; Smith, Cosper, Stalhandske, Scott, & Ferry, 2000; Strop, Smith, Iverson, Ferry, & Rees, 2001). The first tetrameric γ-CA was chemically characterized from *Methanosarcina thermophila* (Alber & Ferry, 1994).

Minor CA families include the δ-class CA from the marine diatom *Thalassiosira weissflogii*, TweCA, which was discovered in 1997 and characterized in 2013 (Roberts et al, 1997, Lee, R. B. et al, 2013). The ζ-class cadmium-containing CA (CDCA) from the same species was reported in 2005 (Lane et al, 2005, Xu, Y. et al, 2008). In 2014, Del Prete and coworkers introduced a novel CA from the protozoan *Plasmodium falciparum*, which belonged to a new η-CA class (Del Prete et al, 2014), and most recently, Kikutani´s group presented a θ-CA from the marine diatom *Phaeodactylum tricornutum* (Kikutani et al, 2016). The distribution of different CA classes over seven kingdoms is shown in Table 1.

Table 1. CA classes identified in different organisms using the classification of seven kingdoms (Zolfaghari Emameh et al, 2014, Ruggiero et al, 2015, Vullo et al, 2017)

Bacteria	α-, β-, and γ-CAs
Archaea	β- and γ-CAs
Protozoa	α-, β-, and η-CAs
Chromista	α-, β-, δ- (marine phytoplankton, haptophytes, dinoflagellates, chlorophytic prasinophytes, diatoms), ζ- (marine diatoms), η- (*P. falciparum*), and θ-CAs (*P. tricornutum*)
Plantae	α- (chloroplasts of green plants), β- (chloroplasts of mono- and dicotyledons), and γ-CAs
Fungi	α- and β-CAs
Animalia	α- (vertebrates, corals) and β-CAs (insects)

2.1.1 α-Carbonic anhydrase isozymes

Sixteen α-CA isoforms have been identified in mammals, and the evolutionary relationships among these isoforms have been investigated (Supuran, C. T., 2008b).

Out of the 16 α-CA isoforms, three isoforms, CA-RP VIII, X and XI, lack catalytic activity (Tashian et al, 2000, Aspatwar et al, 2010, Gilmour, 2010). Classification of α-CA isozymes, examples of their main locations of expression and putative functions in addition to the maintenance of pH homeostasis in mammals are described in Table 2.

Table 2. Classification of mammalian alpha-CA isozymes, their main locations of expression, and putative functions in addition to the maintenance of pH homeostasis.

Localization		CAs	Distribution	Function	(Selected) References
Intracellular	Cytosolic	I	Gastrointestinal (GI) tract, red blood cells	Gas exchange and ion transport	(Headings & Tashian, 1971)
		II	Almost all cells	Bone resorption, osteoclast differentiation, secretion of cerebrospinal fluid, production of aqueous humor, urine acidification, gas exchange, production of gastric acid, production of alkaline pancreatic juice and duodenal bicarbonate, acidification of bile, sperm motility, GABAergic neuronal excitation	(Ruusuvuori et al, 2013, Sly et al, 1983, Krishnan et al, 2018, Parkkila & Parkkila, 1996, Lehenkari et al, 1998, Vaananen & Parvinen, 1983, Leppilampi, Parkkila et al, 2005, Juvonen et al, 1994, Parkkila et al, 1991)
		III	Skeletal muscle, adipose tissue, osteocytes, liver	Antioxidative agent, mitochondrial ATP synthesis	(Carter, N. et al, 1978, Raisanen et al, 1999, Kim, G. et al, 2004, Liu et al, 2007, Shi et al, 2018)
		VII	Liver, brain, colon, skeletal muscle	Antioxidative agent, GABAergic neuronal excitation	(Montgomery et al, 1991, Monti et al, 2017, Ruusuvuori et al, 2013)
		XIII	Colon, brain, kidney, oligodendrocytes, sperm cells, cervical and endometrial mucosa, odontoblasts	Unknown	(Reibring et al, 2014, Lehtonen et al, 2004, Hilvo et al, 2008)
	Mitochondrial	VA	Liver	Ammonia detoxification, ureagenesis and gluconeogenesis	(Fujikawa-Adachi et al, 1999b, Shah et al, 2013)
		VB	Heart, skeletal muscle, pancreas, kidney, salivary glands, spinal cord	Ammonia detoxification, ureagenesis and gluconeogenesis	(Fujikawa-Adachi et al, 1999b, Shah et al, 2013)

Table 2 continues

Localization		CAs	Distribution	Function	(Selected) References
Extracellular	Secreted	VI	Saliva, salivary glands, nasal glands, lacrimal glands, von Ebner's glands, mammary glands, milk, odontoblasts	Protection against dental caries, taste function	(Kivela et al, 1999b, Fernley et al, 1979, Murakami, H. & Sly, 1987, Aidar et al, 2013, Reibring et al, 2014)
	Membrane associated	IV	Kidney, lung, gallbladder, pancreas, eye, colon, blood capillary endothelium, salivary glands, heart, skeletal muscle, odontoblasts	Bicarbonate reabsorption, potential tumor suppressor gene	(Hageman et al, 1991, Carter, N. D. et al, 1990, Fleming et al, 1993, Chen et al, 2017, Reibring et al, 2014, Fleming et al, 1995)
		IX	Small intestine, colon, gallbladder, odontoblasts, fetal intervertebral disc, fetal joint cartilage, various tumors	Tumor-associated CA isoenzyme	(Opavsky et al, 1996, Pastorekova et al, 1997, Kim, J. H. et al, 2013, Reibring et al, 2014, Mboge et al, 2018)
		XII	Aorta, bladder, choroid plexus, colon, esophagus, kidney, liver, lung, lymph nodes, mammary glands, ovaries, prostate, pancreas, peripheral blood lymphocytes, rectum, stomach, skeletal muscle, skin, spleen, testis, trachea, uterus, various tumors	Tumor-associated CA isoenzyme	(Tureci et al, 1998, Ivanov et al, 1998, Waheed & Sly, 2017)
		XIV	Brain, kidney, colon, small intestine, urinary bladder, liver, spinal cord, odontoblasts	Implicated in epileptogenesis and some retinopathies, neuronal signal transduction	(Fujikawa-Adachi et al, 1999a, Makani et al, 2012, Alterio et al, 2014, Reibring et al, 2014)
		XV	Kidney, brain, testis		(Hilvo et al, 2005, Saari et al, 2010, Tolvanen et al, 2013)
CA-RPs	No enzymatic activity	VIII	Brain, lung, liver, salivary gland, stomach, human osteosarcoma cells	Neuronal functions, controls ER-dependent cytosolic Ca^{2+} homeostasis by binding to IP_3 receptor, promotes glucose uptake in human osteosarcoma cells	(Kelly et al, 1994, Aspatwar et al, 2010, Lo et al, 2018)
		X	Brain	Unknown	(Aspatwar, Tolvanen and Parkkila, 2013)
		XI	Brain, associated with several cancers	Unknown	(Aspatwar, Tolvanen and Parkkila, 2013)

2.2 General aspects of secretory carbonic anhydrase VI

2.2.1 Expression of CA VI

CA VI is the only secreted isozyme in the α-CA enzyme family (Fernley et al, 1979, Murakami, H. & Sly, 1987). In the 1970s, Henkin's group discovered a novel taste-associated protein in human saliva. This secreted protein was named Gustin (Henkin et al, 1975). Two decades later, in 1998, Henkin's group further showed that Gustin was identical to CA VI (Thatcher et al, 1998), which had been independently identified as a CA enzyme present in the sheep parotid gland and saliva in 1979 (Fernley et al, 1979). To date, many research groups have studied the CA VI enzyme, but its exact functions still remain a matter of debate. The CA VI enzyme has been isolated from rat (Feldstein & Silverman, 1984) and human (Murakami, H. & Sly, 1987). This enzyme is highly expressed in the serous acinar cells of the parotid and submandibular glands and is one of the major protein constituents of human saliva (Kadoya et al, 1987, Parkkila et al, 1990, Parkkila et al, 1993).

The expression of CA VI is not restricted only to the salivary glands. In 2002, Ogawa and coworkers (Ogawa et al, 2002) demonstrated CA VI expression in the lacrimal gland. Milk contains high levels of secretory CA VI, which shares 100 % amino acid sequence homology with salivary CA VI (Karhumaa et al, 2001). Colostrum contains an eight times higher concentration of CA VI than mature milk; the latter contains concentrations comparable to the mean levels in saliva.

In recent years, many studies have identified CA VI in several different species. Examples of these species and the main localizations of the CA VI enzyme are listed in Table 3.

Table 3. Expression of the *carbonic anhydrase 6* gene in different studies found in the literature.

Species	mRNA	CA VI Protein	Reference
Human (*Homo sapiens*)	Salivary gland, breast, skin, epididymis, ovary, spleen, lymph node, placenta, bone marrow, pancreas, appendix, heart muscle, tonsil, cerebellum	Saliva, von Ebner's glands, serum, serous acinar and ductal cells of the parotid and submandibular glands, secretory granules and cytosol of serous acinar cells, colostrum, milk	(Parkkila et al, 1994, Fernley et al, 1995, Kivela et al, 1999a, Jiang & Gupta, 1999, Karhumaa et al, 2001, Leinonen, Jukka et al, 2001, Ogawa et al, 2002, Lizio, Marina et al, 2015, Lizio, M. et al, 2017)
Horse (*Equus caballus*)	Duodenum, jejunum, ileum, cecum, colon, salivary glands, testis, thyroid gland, liver		(Ochiai et al, 2009)
Pig (*Sus scrofa*)	Parotid gland, kidney	Saliva, serum, bile, seminal plasma, epithelial cells of distal straight tubule of kidney, parotid gland, submaxillary gland, sublingual gland, gallbladder, colostrum	(Nishita et al, 2011, Nishita et al, 2014)
Dog (*Canis lupus familiaris*)	Parotid gland, submandibular gland, sublingual gland, zygomatic gland, vestibule region, respiratory region, olfactory region, lateral nasal gland, esophagus	Parotid gland, submandibular gland, sublingual gland, zygomatic gland, epithelium of the nasal mucosa, serous acinar and ductal epithelial cells of the nasal mucosa.	(Murakami, M. et al, 2003, Ichihara et al, 2007, Kasuya et al, 2007, Sugiura et al, 2008)
Cow (*Bos taurus*)	Liver, mammary gland	Serum, saliva, major salivary glands, forestomach, large intestine, liver, mammary gland, milk, colostrum	(Hooper et al, 1995, Kitade et al, 2003, Kaseda et al, 2006, Nishita et al, 2007)
Sheep (*Ovis aries*)	Lacrimal acinar cells	Lacrimal gland, parotid gland, submandibular gland, milk	(Fernley et al, 1991, Fernley et al, 1979, Penschow et al, 1997, Ogawa et al, 2002)
Rat (*Rattus norvegicus*)		Milk, Saliva, serous acinar and duct cells of the tracheobronchial glands, secretory cells at the base of the ciliated cells of the tracheobronchial surface epithelium, Club cells of the bronchiolar surface epithelium	(Karhumaa et al, 2001, Feldstein & Silverman, 1984, Smith et al, 1986, Etzel et al, 1997, Leinonen, Jukka et al, 2001)
Mouse (*Mus musculus*)	nasal mucosa	Duct contents of the anterior gland of the nasal septum, lateral nasal gland, the mucus covering the respiratory and olfactory mucosa, the lumen of the nasolacrimal duct, CA VI-b in submandibular gland	(Sok et al, 1999, Kimoto et al, 2004)

2.2.2 Structure of CA VI

The human *CA6* gene is located at chromosome 1 p36.22-p36.33 (Jiang & Gupta, 1999, Sutherland et al, 1989). The CA VI transcript (Ensembl protein entry ENST00000377443.6) has eight exons encoded by the *CA6* gene (Ensembl gene entry ENSG00000131686) (Flicek et al, 2012), including a signal peptide positioned at N-terminal amino acids 1-17 plus a main polypeptide chain with amino acid residues at positions 18-308 (The UniProt Consortium, 2010).

The secondary structure of hCA VI adopts the canonical α-CA fold, having a central 10-stranded β-sheet surrounded by several short α-helices and β-strands. There is a disulfide linkage between Cys42 and Cys224, which stabilizes the active site of the enzyme for catalysis (Pilka et al, 2012). The secondary structure and active site of CA VI are shown in Figure 3.

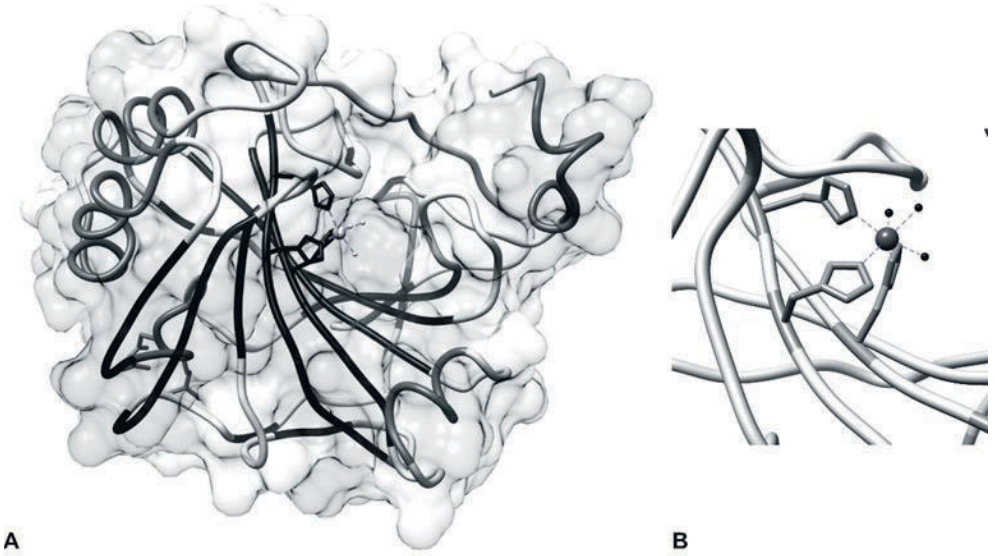

Figure 3. Structure of human CA VI. (A) The secondary structure of human CA VI showing alpha helices (gray) and beta-sheet (black) structures. (B) The active site of human CA VI. The 3D structure (PDB ID: 3FE4, (Pilka et al, 2012)) was retrieved from PDB (Berman et al., 2000), and graphics were created by Maarit Patrikainen with the UCSF Chimera package (Pettersen et al., 2004).

Both human salivary and milk CA VI have a similar molecular mass of 42 kilodalton (kDA), and the polypeptide contains three potential N-glycosylation sites (Asn-X-Thr/Ser) (Aldred et al, 1991, Karhumaa et al, 2001). Thatcher and coworkers detected that N-linked carbohydrate chains contained N-acetyl glucosamine, galactose, mannose, and fucose interior to di-, tri- and tetra-sialyated termini in hCA VI. They also determined five increasingly acidic pI values by isoelectric focusing that were consistent with an addition of sialic acid as the terminal carbohydrate residue on the N-linked glycoforms of the protein. (Thatcher et al, 1998)

Hooper and coworkers studied bovine submaxillary and parotid CA VI and demonstrated that the majority of Asn-linked oligosaccharides present on the secreted CA VI synthesized by the bovine submaxillary gland terminate with N-acetylgalactosamine 4- sulfate (GalNAc-4-SO_4). A similar proportion of Asn-linked oligosaccharides on CA VI synthesized in the bovine parotid gland terminate with β1,4-linked N-acetylgalactosamine (GalNAc), which is not sulfated. The researchers suggested that submaxillary and parotid CA VI may have distinct biological roles based on the differences in the structures of their Asn-linked oligosaccharides and that the 45-kDa protein might be a major endogenous acceptor for the GalNAc-4-sulfotransferase in the submaxillary gland. (Hooper et al, 1995)

Miller and coworkers reported that in HEK 293T cells, when CA VI epitope-tagged with the V5His vector at its carboxyl terminus is expressed, at least 51 % of the secreted CA VI is bound by immobilized biotinylated Wisteria Floribundia Agglutinin (WFA), a lectin that binds oligosaccharides bearing terminal β1,4-linked GalNAc, as in the luteinizing hormone α subunit. CA VI has a recognition determinant that results in its selective modification with GalNAc when expressed by HEK 293T cells. (Torres et al, 1988, Miller et al, 2008)

Several inhibitors have been tested for their ability to inhibit CA VI enzyme activity: iodide, coumarins, sulfonamide derivatives, such as acetazolamide (AZA) and methazolamide, saccharin, and Foscarnet (phosphonoformate trisodium salt) (Murakami, H. & Sly, 1987, Kohler et al, 2007, Temperini et al, 2007, Maresca et al, 2009, Crocetti et al, 2009). A series of sulfonamides incorporating sugar moieties, such as glucose, ribose, arabinose, xylose, and fucose derivatives, showed excellent CA VI inhibition properties (K_i 0.56-5.1 nm), whereas galactose, mannose, and rhamnose scaffolds were weaker inhibitors (K_i 10.1-34.1 nm) (Winum et al, 2009).

2.2.3 Function of CA VI

After four decades of research, the actual function of CA VI has remained unknown. The following chapters describe in more detail the potential roles of CA VI in saliva for taste function and dental health.

2.2.3.1 CA VI in saliva

The serous acinar cells of the parotid and submandibular glands secrete hCA VI into saliva at the rate of 10.2 ± 7.9 μg/min, and this secretion follows circadian periodicity (Parkkila et al, 1993, Parkkila et al, 1995). It has been shown that smokers have unaltered CA activity, salivary secretion rates, and amylase activity levels, whereas both salivary pH and buffer capacity pH values are lower compared to those of non-smokers (Kivela et al, 1997). It has been suggested that, as an enzymatically active CA, salivary CA VI can maintain optimal pH homeostasis within the oral cavity and upper alimentary tract, but it is not directly involved in the regulation of the pH of secreted saliva (Parkkila & Parkkila, 1996, Kivela et al, 1999a, Kivela et al, 1999b). CA VI secretion is not significantly affected by the hormonal alterations associated with pregnancy, nor is it involved in the regulation of actual salivary buffer capacity (Kivela et al, 2003).

The second messenger molecules cyclic adenosine monophosphate (cAMP) and Ca^{+2} regulate a large number of cellular events in eukaryotes (Sharma & Kalra, 1994). cAMP and its associated kinases function in several biochemical processes, including the regulation of glycogen, sugar, and lipid metabolism, and Ca^{+2} works through calmodulin, a ubiquitous calcium-binding protein (Sharma & Kalra, 1994, Ali et al, 2016). Human salivary CA VI stimulates brain calmodulin-dependent cyclic nucleotide phosphodiesterase (cAMP PDEase) activity in the absence of calmodulin, 5- to 6-fold over physiological levels (Law et al, 1987). To compare the levels of salivary cAMP and cyclic guanosine monophosphate cGMP in patients with taste and smell dysfunction with those in normal subjects, parotid saliva was collected from 61 normal volunteers and 253 patients with taste and smell dysfunction. The results showed lower mean concentrations of both cAMP and cGMP in patients than in normal subjects. (Henkin et al, 2007)

2.2.3.2 Role of CA VI for taste function

Henkin´s group became the first to link Gustin (CA VI) to the regulation of taste function (Shatzman & Henkin, 1981). Their observations of the biochemical characteristics of CA VI were similar whether the protein was isolated from subjects with normal taste acuity or from patients with hypogeusia (decreased taste acuity), and hypogeusic subjects had salivary CA VI concentrations as low as 20 % that of normal subjects. They also reported that zinc treatment can affect both taste and CA VI concentrations in hypogeusia, but they did not link CA VI to any specific taste modality (Shatzman & Henkin, 1981). A study on patients with an acute influenza-type illness described a clinical disorder formulated as a syndrome of hyposmia (decreased smell acuity), hypogeusia, dysosmia (distorted smell function), dysgeusia (distorted taste function) and the decreased secretion of parotid saliva CA VI with associated pathological changes in taste bud anatomy. (Henkin et al, 1999a). They further performed a study in which exogenous zinc was administered to the patients to stimulate the synthesis and/or secretion of CA VI to correct the symptoms of this disorder. The results led them to propose a novel hypothesis that CA VI promotes growth and the development of taste buds through its action on taste bud stem cells (Henkin et al, 1999b). Additionally, the high concentrations of CA VI in milk and colostrum suggested that it can participate in the developmental processes of the GI canal during the postnatal period (Karhumaa et al, 2001).

Polymorphism of the *CA6* gene has been the research topic of Barbarossa's group in several publications, and indeed, they have found a link between bitter taste modality and polymorphic changes in the *CA6* gene. They first found a polymorphism (rs2274333 A/G) that contributed to 6-n-propylthiouracil taster status (Padiglia et al, 2010). Later, they elegantly showed that alterations of bitter taste function are due to polymorphic changes in the bitter receptor gene (*TAS2R38*) and *CA6* gene and require contributions from other still unknown factors (Calo et al, 2011). In 2013, Barbarossa's group suggested that polymorphisms of the *CA6* gene are associated with the concentrations of secreted CA VI and that the rs2274333 polymorphic change in the *CA6* gene affects 6-n-propylthiouracil sensitivity by acting on fungiform papilla development and maintenance (Aidar et al, 2013, Melis et al, 2013).

2.2.3.3 Salivary CA VI and dental caries

Saliva contains inorganic compounds and multiple proteins that affect conditions in the oral cavity and initiate various systemic defense mechanisms, including bicarbonate ions, leukocytes, secretory IgA, agglutinating proteins, and a number of enzymes, to the actual site of microbial growth on the tooth surface (Lamkin & Oppenheim, 1993, Lagerlof & Oliveby, 1994, Edgar et al, 1994, Kivela et al, 1999b). Saliva is involved in food debris clearance and also provides inorganic ions for both the neutralization of acidic microbial metabolic products and enamel remineralization after the consumption of fermentable carbohydrates (Edgar et al, 1994). The salivary bicarbonate diffuses into dental plaque, combining with H^+ to form carbonic acid. CA VI is believed to contribute to the neutralization of plaque acid by accelerating the chemical reaction from carbonic acid to CO_2 and H_2O (Kivela et al, 1999a). Bicarbonate itself is an important factor for the pH balance on dental surfaces because it is the main contributor to the salivary buffering capacity (Bardow et al, 2000).

Dental plaque is formed on the tooth surface as a biofilm of microorganisms embedded in a matrix of polymers of host and bacterial origin (Marsh, 2006). Dental caries is a consequence of the dental hard tissues dissolving under the acid conditions prevailing beneath dental plaque (Kivela et al, 1999b). In 1999, Kivelä and coworkers suggested that salivary CA VI played a specific role in the natural defense systems against dental caries because the concentration of CA VI exhibits a negative correlation with the DMFT index in individuals with poor hygiene (Kivela et al, 1999b). Perez and coworkers showed a positive association between salivary buffer capacity and the rs2274327 (C/T) polymorphism when analyzing the allele and genotype distribution of three polymorphisms in the coding sequences of the *CA6* gene in children aged 7-9 years (Peres et al, 2010). Studies performed within a Chinese population showed no association between the rs2274327 polymorphism and dental caries susceptibility, whereas the rs17032907 genetic variant and the haplotype (rs2274328, rs17032907 and rs11576766) of *CA6* may be associated with dental caries susceptibility (Li, Z. Q. et al, 2015). Sengul and coworkers also observed no correlation of rs2274327 polymorphism frequencies between study groups of carious and non-carious children in Turkey (Sengul et al, 2016). A recent study showed higher activity of CA VI in saliva but a lower salivary flow rate associated with the development of dental caries among school children. In the caries-free group, a higher concentration of CA VI, higher salivary flow rate, higher pH and higher buffering capacity were observed, suggesting that CA VI is able to neutralize

the acid in the oral environment and thus provides greater protection against tooth decay. (Picco et al, 2017)

The same researchers also reported a contradictory result: that salivary CA VI activity was, in fact, higher in children with caries. Similarly, Borghi and coworkers showed that CA VI activity was 53.8 % higher in the saliva of children with early childhood caries than in the saliva of the caries-free children. (Borghi et al, 2017)

A previous study using a mouse cariogenesis model revealed that after infection with *S. mutans* strain UA159, total smooth surface caries and sulcal caries were more than 6-fold and 2-fold lower, respectively, in *Car6*$^{-/-}$ mice than in WT mice (Culp et al, 2011). The recovery of *S. mutans* and total microbiota from molars was lower for the oral microbiota of *Car6*$^{-/-}$ mice. Additionally, *Lactobacillus murinus* and an unidentified *Streptococcus* species were cultivated at higher levels in *Car6*$^{-/-}$ mice, suggesting that salivary CA VI may promote caries via the enzymatic production of acid within plaque and/or modulation of the oral microbiota to favor *S. mutans* colonization (Culp et al, 2011). A recent study hypothesized that CA VI and *CA6* gene variations contribute to the microbial environment in both the nasopharynx and gastro-intestinal tracts. The results, indeed, supported the hypothesis by showing that *CA6* gene polymorphisms rs10864376 (T), rs3737665 (T), rs12138897 (G) and haploblock TTG of *CA6* are associated with *S. mutans* colonization, overall microbiota composition, and dental caries. It was also observed that secreted salivary CA VI tended to be linked to *CA6* gene variation. (Esberg et al, 2019)

2.2.3.4 Other suggested functions of CA VI

In general, CA VI has played a minor role in previous studies focusing on cancer. Hsieh and coworkers investigated the potential diagnostic utility of CA VI in differentiating acinic cell carcinoma (AciCC) of the salivary gland from its morphological mimic, mammary analogue secretory carcinoma. They reported that CA VI shows staining sensitivity and specificity as high as that previously detected for a calcium-dependent chloride channel protein (DOG1). Therefore, they concluded that a combination of CA VI and DOG1 could serve as an ideal immunohistochemical panel for the diagnosis of AciCC. (Hsieh et al, 2016)

Sjögren's syndrome (SS) is a complex autoimmune disease in which damaged salivary and lacrimal glands lead to dry eyes and dry mouth and also secondary problems from the lack of protective secretions. Some patients with primary SS (pSS) develop systemic manifestations including lung and kidney disease, peripheral neuropathy, vasculitis, and lymphoma (Giusti et al, 2007, Baldini et al, 2011). SS may

also occur secondary to other autoimmune diseases including SSc, RA, systemic lupus erythematosus, PM, and MCTD (Manoussakis & Moutsopoulos, 2000, Peri et al, 2012). In 2011, Baldini and coworkers identified 15 differentially expressed proteins in pSS. Among them, α-amylase precursor, carbonic anhydrase VI, β-2 microglobulin, glyceraldehyde-3-phosphate dehydrogenase, epidermal fatty acid binding protein, and immunoglobulin κ light chain apparently showed the most significant differences in pSS when compared with control groups (Baldini et al, 2011). A recent article, from 2017, showed that IgA type anti-CA VI autoantibodies were frequently seen in patients with long-standing SS, and further studies were suggested to determine the usefulness of anti-salivary gland protein 1, anti-CA VI, and anti-parotid secretory protein antibodies in the diagnosis of sSS in SSc, sSS in RA, and MCTD (Shen et al, 2012, De Langhe et al, 2017).

2.2.3.5 Nuclear carbonic anhydrase 6B (CA VI-b)

In 1999, Sok and coworkers discovered a novel form of murine CA VI, CA VI-b, while studying the transcription factor CCAAT/Enhancer-Binding Protein Homologous Protein (CHOP), which is a highly responsive enzyme to certain forms of stress (Sok et al, 1999). Their study focused on stressed (tunicamycin-treated) *chop*$^{+/+}$ and *chop*$^{-/-}$ mouse embryo fibroblasts, and they identified several cDNA fragments from differentially expressed genes. Sequence analysis of CA VI-b revealed the presence of divergent N-terminal sequences in proteins, thus indicating that both the secreted form and the stress-induced form of CA VI are likely the products of the same gene. Untreated and tunicamycin-treated fibroblasts revealed that type CA VI-b was expressed in response to stress in these cells, whereas the secreted form of CA VI was restricted to the salivary glands (Sok et al, 1999).

In 2014, it was confirmed using the CN1.4 cell line that the CA VI-b isoform is retained in the cell, where it is distributed between the nucleus and the cytoplasm. CA VI-b induced by endoplasmic reticulum stress in neurons was dependent upon CHOP. (Matthews et al, 2014)

Recently, it was found that CA VI-b is directly connected to the innate immune response by selectively inducing cytokine IL-12 production through PRMT5 and regulating histone H3R8me2s arginine modification, independent of its CA activity (Xu, J. et al, 2017).

2.3 Taste perception

Taste perception is one of the most important features for animals to prevent the ingestion of toxic substances, evaluate the nutritious content of food, guide essential appetitive behaviors, and help maintain a healthy diet. It is generally accepted that our taste evokes from a few distinctive sensations: sweet, bitter, sour, salty, and umami (savory). (Chandrashekar et al, 2006)

Taste receptor cells (TRCs) assemble into taste buds, which are distributed into different papillae within the tongue epithelium. Circumvallate papillae are found at the very back of the tongue and contain hundreds (mice) to thousands (human) of taste buds. Foliate papillae localize to the posterior lateral edge of the tongue and contain dozens to hundreds of taste buds, and fungiform papillae contain a single or a few taste buds and are located at the frontal part of the tongue (Hoon et al, 1999). Filiform papillae are the smallest and most numerous in humans and do not contain any taste buds. The filiform papillae are located over the entire anterior-dorsal surface of the tongue, with their tips pointing backward, and serve only a mechanical role (Ross & Pawlina, 2011). Figure 4 shows the localization of these different papillae.

Taste modalities are mediated by different mechanisms. Salty tastants modulate taste-cell function by the direct entry of Na^+ and H^+ through specialized membrane channels on the apical surface of the cell. The entry of Na^+ through amiloride-sensitive Na^+ channels is believed, at least partly, to stimulate TRC activation (Heck et al, 1984, Avenet & Lindemann, 1988). Sour taste detection functions as a warning mechanism and provides important sensory input to discourage the ingestion of foods spoiled by acid-producing microorganisms (DeSimone et al, 2001). Huang and coworkers identified a candidate mammalian sour taste sensor, polycystic kidney disease 2-like 1 protein (PKD2L1), an ion channel that is expressed in a subset of TRCs distinct from those responsible for sweet, bitter, and umami taste modalities (Huang et al, 2006). This sour taste receptor responds to acids, such as citric acid, tartaric acid, acetic acid, and hydrochloric acid (Chandrashekar et al, 2006).

G-protein-coupled receptors (GPCRs) taste receptor type 1 member 1 (T1R1), T1R2, and T1R3 are expressed in subsets of co-expressing TRCs and mediate sweet and umami taste perception (Nelson et al, 2001, Zhao et al, 2003). In mice, TRCs in both fungiform and circumvallate papillae express each T1R receptor alone or in all possible combinations (T1R1 + T1R2, T1R1 + T1R3, T1R2 + T1R3, and T1R1 + T1R2 + T1R3) (Kim, M. R. et al, 2003). The heterodimer of T1R3 and T1R2 forms a sweet taste receptor that responds to all classes of sweet tastants, including natural

sugars, D-amino acids, intensely sweet proteins, and artificial sweeteners (Nelson et al, 2001, Li, X. et al, 2002). T1R1 and T1R3 form a heterodimer, which is an amino acid (umami) taste receptor (Li, X. et al, 2002, Nelson et al, 2002, Zhao et al, 2003). This receptor is strongly stimulated by purine nucleotides in both mouse and human (Li, X. et al, 2002, Nelson et al, 2002).

Bitter taste is mediated by a family of highly divergent GPCRs, the T2Rs, which are selectively expressed in subsets of TRCs except those containing sweet and umami receptors (Adler et al, 2000, Matsunami et al, 2000).

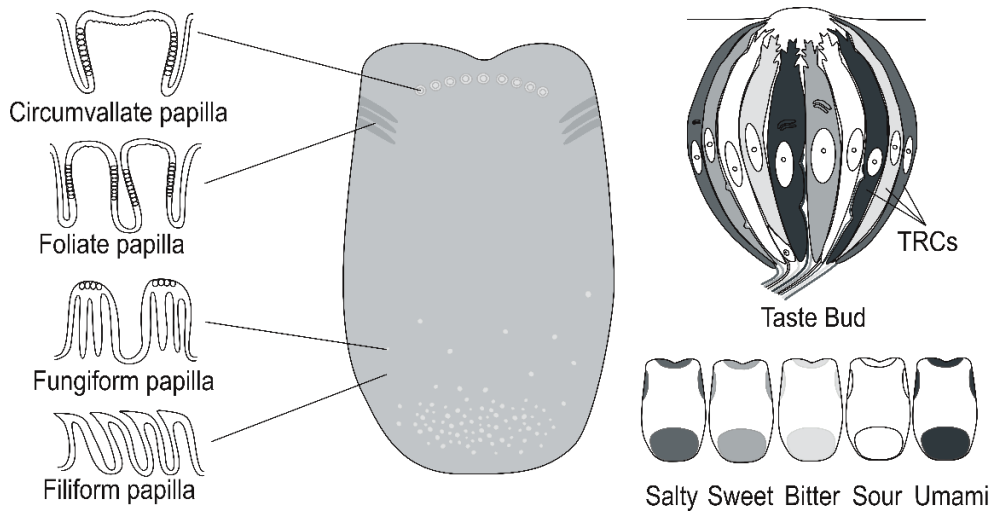

Figure 4. The left side of the image shows the specialized mucosa of the oral cavity on a dorsal surface of a human tongue. Filiform papillae distribute over an anterior dorsal surface of the tongue, while fungiform papillae project above the filiform papillae, are scattered and are more numerous near the tip of the tongue. The circumvallate papilla position on the posterior side, forming a V-shaped configuration. The foliate papillae occur on the lateral side of the tongue. A diagram of the taste bud shows sensory cells (gray), supporting, and basal cells. Nerve fibers have synapses with neuroepithelial (sensory) cells (Chandrashekar et al, 2006), modified by Maarit Patrikainen.

2.4 Pentraxin

Innate immune systems, based on pattern recognition, exist in some form in all metazoan organisms (Medzhitov, 2007). Pattern recognition molecules (PRMs) recognize conserved structures on the surface of pathogens and activate the innate immune response. Pentraxins (PTXs), belonging to a humoral arm of innate immunity, are a superfamily of fluid phase pattern recognition molecules

characterized by a cyclic multimeric structure with a regulatory role in inflammation (Bottazzi et al, 2016). PTXs share a 200-amino-acid-long domain at the protein C-terminus and are divided into groups based on their primary subunit structures with a distinctive primary motif, His-X-Cys-X-Ser/Thr-Trp-X-Ser, where X is any amino acid. Here, I concentrate only on short PTXs, a C-reactive protein (CRP), and serum amyloid P (SAP) (Garlanda et al, 2005, Bottazzi et al, 2010), which are 25-kDa proteins characterized by a common structural organization in five identical subunits arranged in a pentameric radial symmetry. Human CRP and SAP share approximately 51 % amino acid sequence identity and are thought to have originated from a single gene duplication (Rubio et al, 1993). The promoter region of human CRP comprises two acute phase response elements, each containing a binding site for the liver-specific transcription factor 1 and two C/EBP (CCAAT/enhancer binding protein β) binding sites, both necessary and sufficient for interleukin 6 (IL-6)-induced transcription (Toniatti et al, 1990, Ramji et al, 1993, Li, S. P. & Goldman, 1996). SAP is a highly conserved plasma glycoprotein and a normal component of basement membranes (Zahedi, 1997).

SAP plays a dual role in bacterial infections, exhibiting a host defense function against pathogens to which it does not bind. When SAP binds to bacteria, a strong anti-opsonic effect is observed, resulting in enhanced virulence of the infectious agent. The physiological functions attributed to CRP and SAP involve Ca^{2+}-dependent ligand binding. (Noursadeghi et al, 2000, Garlanda et al, 2005)

2.5 Functional genomics

The work described within this thesis could not be performed without using several different tools based on a concept called functional genomics. Functional genomics covers high-throughput technologies for DNA and protein analysis, in which instead of studying one protein or gene at the time, entire genomes and proteomes are studied. There are several different subdivisions and tools of functional genomics, such as gene disruption, proteomics, and structural genomics. (Gasperskaja & Kucinskas, 2017)

As an example, KO or KD animal models have proven very useful for studying the roles of various genes or proteins. In many cases, proteomics and genome-wide studies have been efficiently deployed to reveal multiple phenotypic effects that are related to complete or partial silencing of target genes. There are several model animals that have been widely used in research: mouse (*Mus musculus*), rat (*Rattus*

norvegicus), zebrafish (*Danio rerio*), fruit fly (*Drosophila melanogaster*), and nematode worm (*Caenorhabditis elegans*) (Kawakami et al, 2017, Mitani, 2017, Yamaguchi & Yoshida, 2018).

2.5.1 Knockdown or knockout animal models for CAs

In the last two decades, several KD or KO animal models have been developed for certain CA isoforms. These models have then been utilized to study the corresponding phenotypes. Table 3 lists the developed KO models and their phenotypes in mouse and zebrafish strains.

Table 4. Target genes/proteins of KO animal mouse models used in CA studies and their observed phenotypes.

KO target	Strain	Phenotype	Reference
Car2	C57BL/6J x DBA/2J	*Car2* null mice were smaller and had renal tubular acidosis and vascular calcification. Lack of duodenal bicarbonate secretory response to prostaglandin E2.	(Leppilampi, Parkkila et al, 2005, Lewis et al, 1988, Spicer et al, 1989)
Car3	C57BL/6J	Growth, fertility, and life span of the mutant mice were similar to those of WT controls. Half-relaxation times of single twitches and of tetani are reduced in soleus muscle.	(Kim, G. et al, 2004)
Car4	C57BL/6J	Viable, healthy, fertile and grew normally when crossed with WT partners. Matings between male and female homozygous *Car4* null mice produced small litters, and many of the pups did not survive. Extracellular buffering was disturbed in the central nervous system.	(Shah et al, 2005)
Car5A	C57BL/6J	*Car5A* null mice were smaller than WT littermates and bred poorly. Their fasting blood sugars were normal but blood ammonia concentrations were markedly elevated.	(Shah et al, 2013)
Car5B	C57BL/6J	*Car5B* null mice showed normal growth, normal blood ammonia, and normal fasting blood sugar levels.	(Shah et al, 2013)
Car5A/B	C57BL/6J	Double KO mice were even smaller and had greater hyperammonemia than *Car5A* null mice. They were fertile, but pups were produced in less than predicted numbers. Survival after weaning was reduced, especially for males.	(Shah et al, 2013)
Car6	C57BL/6J	Growth, fertility, and life span of the mutant mice were similar to those of WT controls. Lower cariogenesis rate in *Car6* null mice compared to that in WT mice. A greater number of lymphoid follicles was observed in the small intestinal Peyer's patches. No other morphological differences were found.	(Culp et al, 2011, Pan et al, 2011)
Car8	Was found in C57BL/KS	Wobbly side to side ataxic movements that were seen when the mice reached two weeks of age and lasted throughout their life span. Frequent tail elevation and intermittent Straub tail reaction were seen.	(Jiao et al, 2005)
Car9	C57BL/6J	Gastric hyperplasia of the glandular epithelium with numerous cysts. Gastric submucosal inflammation. Disruption of brain histology, abnormal locomotor activity, and poor performance in a memory test.	(Gut et al, 2002, Leppilampi, Karttunen et al, 2005, Pan et al, 2012)
Car12	Tg(tetL)1Bjd/J	*Car12* null mice resulted in reduced fitness.	(Geurts et al, 2006)
Car14	C57BL/6J	*Car14* null mice were viable, fertile, and grew normally. Extracellular buffering was disturbed in the central nervous system.	(Shah et al, 2005)
Car4/Car14	C57BL/6J	Double KO mice appeared healthy, grew and were fertile, but both males and females were significantly smaller than WT mice. Female double KO mice died early, and the few that survived were infertile. *Car4/Car14* double mutants showed a greater deficit in a light response than the *Car14*-null retina.	(Shah et al, 2005, Ogilvie et al, 2007)
ca5	TL/4608 ENU	Fish had the collapse of the medial fins followed by necrosis. Rapid embryo degeneration.	(Wienholds et al, 2003, Postel & Sonnenberg, 2012)
ca8	AB, TL	Fish had shortened tail, a curved body axis, the absence of a swim bladder and otolith vesicles.	(Aspatwar, Tolvanen, Jokitalo et al, 2013)
ca10a and *ca10b*	AB	Fish had curved body, a pericardial edema, abnormalities in the head and eye and increased apoptotic cell death in the brain region. Developmental delay with a high rate of mortality in larvae.	(Aspatwar et al, 2015)

3 AIMS OF THE PRESENT STUDY

The general aim of the study was to reveal new information on the expression and function of CA VI. The specific aims were as follows:

To evaluate the differences of taste modalities in CA VI-deficient mice compared with WT mice (I).

To investigate whether gene expression profiles differ in the trachea and lung of CA VI-deficient mice compared with WT mice (II).

To trace evolutionary development of CA VI using phylogenetics (III).

To study the structure, location and function of the novel Pentraxin - CA VI protein using protein modeling, immunohistochemistry, enzyme kinetics, and a *ca6* KD zebrafish model (III).

4 MATERIALS AND METHODS

4.1 Animal models

Car6$^{-/-}$ mice were designed, produced and originally characterized by Dr. Peiwen Pan at the University of Tampere (Pan et al, 2011). To obtain mice with a pure C57BL/6 strain background, the mice with the targeted allele were backcrossed for more than ten generations. The mice were housed in pathogen-free conditions at the University of Oulu. The health status of the animals was monitored on a regular basis according to the Federation for Laboratory Animal Science Associations (FELASA) recommendations. The Animal Experimentation Committee approved the production and characterization of the KO mice. All of the *Car6*$^{-/-}$ and WT mice used in this study were bred and raised at the Animal Centre, University of Oulu, and transferred before the experiments to either the Neuroscience Center/Laboratory Animal Center, University of Helsinki, or the animal facility of the University of Tampere.

WT, AB strain, zebrafish (*Danio rerio*) were maintained under standard conditions at +28.5 °C (Westerfield, 2007). The experiments involving zebrafish and their embryos/larvae were carried out according to the rules of the Provincial Government of Western Finland, protocol #LSLH-2007-7254/Ym-23. Using maximum 5-days-old zebrafish larvae as a model organism requires no ethical permission.

4.2 *ca6* knockdown zebrafish (III)

Zebrafish embryos/larvae were collected from the breeder tanks with a sieve and moved to Petri dishes (Sarstedt, Nümbrecht, Germany) in embryonic medium supplemented with 1-phenyl-2-thiourea (Sigma-Aldrich, Inc., St. Louis, Missouri, USA) at +28.5 ºC until they were used in experiments. KD of *ca6* was carried out using the translation-blocking antisense morpholino oligonucleotides (MOs) (GeneTools LLC, Philomath, Oregon, USA) listed in Table 5. MOs targeted the sequence 5' upstream of the translation initiation site of the *ca6* transcript. Random

control MOs were used as a control. Injection of one nanoliter of MO dilution was performed into the yolks of one- to two-cell stage larvae. Screenings of injected larvae for successful injections were carried out using a Lumar V1.1 fluorescence stereomicroscope (Carl Zeiss MicroImaging GmbH; Göttingen, Germany) and AxioVision software version 4.9 (Carl Zeiss MicroImaging GmbH) at 24 hours post fertilization (hpf). Phenotypic appearance was observed from the images of the developing embryos/larvae from one dpf to five dpf of WT and MO-injected zebrafish. Larvae were first euthanized using 0.05 % ethyl 3-aminobenzoate methane sulfonate salt (Tricaine) (Sigma-Aldrich, Inc.) in embryo medium and embedded in 17 % high molecular weight methylcellulose in 15 x 30 mm transparent polypropylene Petri dishes. Images were taken with a Zeiss Stereo Microscope (Carl Zeiss MicroImaging GmbH) with a NeoLumar S 1.5x Objective (Carl Zeiss MicroImaging GmbH). Further details are described in article III.

Table 5. Translation-blocking antisense MO targets and sequences.

Target	MO sequences (5'-3')
ca6 mRNA	CTG CCT GTG CTC TGA ACT GTT TCT C
Random control	CCT CTT ACC TCA GTT ACA ATT TAT A

4.3 DNA and total RNA extraction and synthesis of cDNA (I, II, III)

Genomic DNA extractions from mice ear samples were carried out using a NucleoSpin Tissue Kit (Macherey & Nagel GmbH & Co. KG, Düren, Germany) according to the manufacturer's instructions. Further details are described in article II.

For microarray, eight WT and six *Car6*$^{-/-}$ KO male mice at the age of eight weeks were sacrificed using CO_2, and trachea and lung samples were collected. Total RNA from tissue samples was extracted using an RNeasy® Mini Kit (Qiagen, Hilden, Germany), including DNase digestion according to the manufacturer's instructions. The RNA concentrations and purity were determined by measuring optical density at 260 and 280 nm, using an ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, USA). Triplicates of each sample, including eight WT and six KO for the trachea, eight WT and six KO for the lung, and four sample duplicates as internal controls, were pipetted into a 96-well plate. One hundred fifty nanograms of total RNA from each sample was converted to double-stranded cDNA and analyzed individually using Illumina's Sentrix® Mouse Ref-8 v2 Bead Chips (Illumina, Inc.,

San Diego, California, USA) by Gen-Probe Life Sciences Ltd. (Oak Business Park, Manchester, UK).

For the *Car6* expression analysis and microarray validation, the RNA samples extracted from the trachea and lung of WT and *Car6*$^{-/-}$ mice were converted to cDNA using a High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Inc., Foster City, California, USA) with random hexamer primers according to the manufacturer's instructions. Further details are described in article II.

For recombinant protein production, total RNA was isolated from 30-mg samples of adult *D. rerio* homogenate as described above. For *ca6* expression analysis, total RNA from the 0-168 hpf embryos/larvae and from different organs of the adult zebrafish was isolated from 30-µg samples as described above. A reverse transcriptase step was performed using 0.1-5 µg of total RNA for synthetizing the first-strand cDNA using a First Strand cDNA Synthesis Kit (High-Capacity cDNA Reverse Transcription Kits, Applied Biosystems, Inc.) with random primers and M-MuLV reverse transcriptase enzyme, according to the manufacturer's instructions. Further details are described in article III.

4.4 Polymerase chain reaction, PCR (I, II, III)

Genotyping of *Car6*$^{-/-}$ mice was performed by multiplex PCR using RedHot DNA Polymerase (Thermo Fisher Scientific, Waltham, Massachusetts, USA). The PCR was run using primers F1 + R1, producing a 434-bp product for WT (P1), and F1 + R2, producing a 673-bp product for *Car6*$^{-/-}$KO (P2) as described previously (Pan et al, 2011). Primer sequences are shown in Table 6, and PCR conditions are described in Table 7.

For microarray validation, several genes of interest were chosen as external standards. The external standards of the trachea tissue were prepared using primers designed for the genes *Fas apoptotic inhibitory molecule 3* (*Faim*), *Calcitonin/calcitonin-related polypeptide, alpha* (*Calca*), *Lysozyme 1* (*Lyz1*), *Natriuretic peptide precursor type A* (*Nppa*), *Demilune cell and parotid protein 1* (*Dcpp1*), *Surfactant associated protein C* (*Sftpc*), and *Uncoupling protein 1 (mitochondrial, proton carrier)* (*Ucp1*). *D site albumin promoter binding protein* (*Dbp*) and *Glycerophosphodiester phosphodiesterase domain containing 3* (*Gdpd3*) were chosen as external standard genes for microarray of the lung tissue. The DNA samples extracted from WT mouse tissues were used as templates, and murine *β-Actin* (*mActb*) was used as a housekeeping gene for qRT-PCR normalization. The

primer sequences are listed in Table 6, and PCR conditions for P3-P9 for the trachea samples, P10-P11 for the lung samples, and P12 for *mActb* are shown in Table 7.

To build a construct for zebrafish CA VI recombinant protein production, primers F12 and R13 were used for coding sequence identification, and the amplification was carried out using conditions for P13. A P14 reaction was used to add a BamHI restriction site into the N-terminus and a SalI restriction site, a tag of six histidines for protein purification, and a thrombin cleavage site for histidine tag removal into the C-terminus of the insert (F13 and F14). To confirm the coding sequence for the recombinant CA VI protein, the insert was sequenced (P15) using a ABI 209 PRISM BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Inc.). The primer sequences and PCR conditions are described in Table 6 and Table 7, and the reactions were performed using the Applied Biosystems 3130*xl* Platform (Applied Biosystems, Inc.). Sample precipitation and sequence analyses are described in detail in article III.

4.5 Quantitative reverse-transcription PCR (qRT-PCR) (II, III)

qRT-PCR of duplicate cDNA samples of trachea and lung tissue, from both *car6*$^{-/-}$ and WT mice, was performed on a 96-Well Optical Reaction Plate (ABI PRISM™, Applied Biosystems, Inc.) using Power SYBR® Green PCR Master Mix (Applied Biosystems, Inc.). A total of 15 μl of each qRT-PCR reaction contained 0.5 μl of first-strand cDNA, each primer at 0.5 μM (Oligomer Oy, Helsinki, Finland), and 1× SYBR Green Master Mix (Applied Biosystems, Inc.). Amplification and detection were carried out using the ABI PRISM®7000 Detection System (Applied Biosystems, Inc.) using the qRT-PCR protocol described in Table 7, with the annealing temperatures (P3-P12) listed in Table 6. The relative expression levels of genes in the microarray validation were assessed by the Sequence Detection System v 1.2.3 (Applied Biosystems, Inc.) using mouse housekeeping gene *beta-actin* (*mActb*) values as an internal control to normalize the qRT-PCR C_t-values. Comparison of the relative expression values between the two groups is described in data analysis section 4.14. Detailed protocol information is described in article II.

The qRT-PCR reactions of samples isolated at different stages of development, from 0-168 hpf whole embryos/larvae, and from different organs of the adult zebrafish were carried out using primers for *ca6* (P16) and the housekeeping gene *gapdh* (P17). The reaction mixture was the same as that described above. Information on the primer sequences of chosen genes and qPCR conditions is shown in Table 6

and Table 7, and data analysis is described in Section 4.14. Detailed protocol information is described in article III.

Table 6. Primer sequences and PCR steps used in this study.

PCR	T_m	Target	Primer sequences (5'- 3')
P1 (F1+R1)	62	*Car6* WT	CCT GGA GTT CAC TAT GAC TAA C (F1)
P2 (F1+R2)	62	*Car6* KO	GGA TCC AGC TTG TTA GGC TT (R1)
			GGC CTA CCC GCT TCC ATT GC (R2)
P3 (F2+R3)	62	*Faim3*	CAG GTG ATC GGT GTG TGT CA (F2)
			CCA GAC ACT GGC AGG AAG TA (R3)
P4 (F3+R4)	60	*Calca*	CCA CTG CCA GGA TCA AGA GT (F3)
			AAG ATG CTG ACA ACC AGG AA (R4)
P5 (F4+R5)	62	*Lyz1*	CCA GGC CAA GGT CTA CAA TC (F4)
			ATA GTC GGT GCT TCG GTC TC (R5)
P6 (F5+R6)	60	*Nppa*	TGG GCT TCT TCC TCG TCT TG (F5)
			CTA GCA GGT TCT TGA AAT CC (R6)
P7 (F6+R7)	60	*Dcpp1*	GCT GGC CTT ACT GAT TCT TG (F6)
			CAC TTG ACC GTC CTT GTT GT (R7)
P8 (F7+R8)	60	*Sftpc*	AGA GGT CCT GAT GGA GAG TC (F7)
			CAC CAC AAC CAC GAT GAG AA (R8)
P9 (F8+R9)	58	*Ucp1*	AGA GTT ATA GCC ACC ACA GA (F8)
			GCA ACA AGA GCT GAC AGT AA (R9)
P10 (F9+R10)	62	*Dbp*	ACC GCG CAG GCT TGA CAT CT (F9)
			GTC TCA TGG CCT GGA ATG CT (R10)
P11 (F10+R11)	60	*Gdpd3*	TTC TCA CCA GGC CAT TTT GC (F10)
			GCT ACC TTG TGA ATG AGT TC (R11)
P12 (F11+F12)	57	*mActh*	AGA GGG AAA TCG TGC GTG AC (F11)
			CAA TAG TGA TGA CCT GGC CGT (R12)
P13 (F12+R13)	55	*ca6*	ATG GAG CAG CTG ACT CTA GTC (F12)
			TTT CTC TGT TTC TCT ATT ATT ATT AT (R13)
P14 (F13+R14)	62	P13	GGC CGG ATC CAT GGA GCA GCT GAC TCT AGT (F13)
			GCC GTC GAC TTA ATG GTG GTG ATG GTG GTG GGA ACC ACG
			GGG CAC CAG TTT CTC TGT TTC TCT ATT ATT ATT ATT ATC C (R14)
P15 (F14+R15)	50	P14	AAT GAT AAC CAT CTG GCA (F14)
			GGT ATG GCT GAT TAT GAT (R15)
P16		*ca6*	CAA ACA TTT ATT TGC CAG CAC TCC (F15)
			TAT GTC CAA TAA TCT CCA TCT ACT CC (R16)
P17		*gapdh*	AGT GTC AGG ACG AAC AGA GGC T (F16)
			GCC AAT GCG ACC GAA TCC GTT A (R17)

F forward primer, *R* reverse primer, T_m annealing temperature

Table 7. PCR conditions used in this study.

	PCR 1-2		PCR 3-12		PCR 13-14		PCR 15		PCR 1-2		qRT-PCR	
	T/ °C	Time	T/ °C	Time	T/ °C	Time	T/ °C	Time	T/ °C	Time	T/ °C	Time
1) Denaturation	94	3 min	95	2 min	94	3 min	-	-	94	3 min	95	10 min
2) Denaturation	94	30 s	95	25 sec	94	30 s	96	10 s	94	30 s	95	15 sec
3) Annealing*	62	30 s	57-62	35 sec	62	30 s	50	5 s	62	30 s	60	1 min
4) Elongation	72	1 min	72	45 sec	72	1 min	55	4 min	72	1 min	-	-
5) Final extension	72	5 min	72	5 min	72	5 min	-	-	72	5 min	-	-
Cycles	40		30		35		35		40		40	
Steps	2-4		2-4		2-4		2-4		2-4		2-3	

*) Annealing temperatures of target genes are listed in Table 4 as T_m.

4.6 Recombinant zebrafish (*Danio rerio*) CA VI-PTX protein production (III)

The baculoviral genome construct encoding the *ca6* recombinant protein was generated according to the bac-to-bac© Baculovirus Expression System instructions (Invitrogen, Camarillo, California, USA) and was optimized for *Spodoptera frugiperda*. A restriction enzyme digestion of purified insert DNA and pFastBac™ (Invitrogen) target vector was carried out using BamHI and SalI restriction enzymes (New England Biolabs, Ipswich, Massachusetts, USA) in 1 % BSA (Sigma-Aldrich, Inc.) and 10 % NEBuffer 3 (New England Biolabs) incubated at +37 °C for two hours. Purification of DNA was carried out with the Illustra™ GFX PCR DNA and GEL Band Purification Kit (GE Healthcare Life Sciences, Chicago, Illinois, USA). The purified insert DNA and vector were incubated at room temperature (RT) for 15 min in 10 × T4 ligase buffer and T4 ligase (New England Biolabs), followed by transformation into One Shot® TOP10 chemically competent *Escherichia coli* cells (Invitrogen).

The ligation mixture and TOP10 cells melted on ice were incubated for 30 min on ice followed by heat shock at +42 °C for two min. S.O.C. medium (Invitrogen) was added, and the solution was incubated on a shaker at +37 °C for one hour. The cells were plated in culture plates containing pre-warmed Luria Broth (LB) (Sigma-Aldrich, Inc.), Bacto™ Agar (Becton, Dickinson and Company, Franklin Lakes, NJ, USA), and 50 µg/ml ampicillin (Roche Diagnostics, Basel, Switzerland) and incubated at +37 °C overnight. Colonies containing a correct size insert were cultured overnight (5 ml), and plasmids were purified using a QIAprep Spin Miniprep Kit™ (Qiagen).

Transposition from the pFastBac™ donor vector with the *ca6* coding sequence to MAX Efficiency® DH10Bac™ Competent Cells (Invitrogen) containing the baculovirus shuttle vector genome was performed according to the protocol provided by the manufacturer. A streak of DH10Bac bacteria cultured overnight on LB plates containing 40 µg/ml kanamycin sulfate (Roche Diagnostics) was suspended in 100 µl of 100 mM $CaCl_2$ and incubated on ice for 15 min. After this, approximately 100 ng of plasmid was added to the suspension, and the cells were incubated on ice for 30 min. Heat shock was performed at 37 ºC for 2 min and placed on ice for two min, followed by the addition of 450 µl of S.O.C. medium (Invitrogen). The cells were grown at 37 ºC for four hours in an orbital shaker.

A solution containing 60 µl of 20 mg/ml 5-bromo-3-indolyl β-D-galactopyranoside (Bluo-gal , Sigma-Aldrich, Inc.) and 10 µl of 1 M isopropyl-β-D-thiogalactopyranoside (IPTG, Thermo Fisher Scientific) was spread evenly on LB agar plates containing 7 µg/ml gentamicin sulfate (Sigma-Aldrich, Inc.), 40 µg/ml kanamycin sulfate (Roche Diagnostics), and 10 µg/ml tetracycline hydrochloride (Sigma-Aldrich, Inc.) and incubated at +37 ºC in the dark. After incubation for 4 hours, the cells were centrifuged at 6000 x g 1 min at +20 ºC, resuspended to 100 µl of S.O.C. medium and 70 µl of cells were spread on the Bluo-gal plates; the cells were grown at 37 ºC for 24 hours. Recombinant bacmid selection was carried out according to the manufacturer´s instructions using blue/white colony selection. Selected colonies were cultured in LB, 7 µg/ml gentamicin sulfate, 40 µg/ml kanamycin sulfate, and 10 µg/ml tetracycline hydrochloride medium at +37 °C for three days, followed by bacmid purification using a PureLink® HiPure Plasmid Miniprep Kit (Invitrogen) according to the manufacturer's instructions.

For transfection of the insect cells, a primary stock was prepared. First, *Spodoptera frugiperda* Sf9 cells (Thermo Fisher Scientific) were calculated, 2 ml of cells (0.5 million cells/ml) was pipetted into a 6-well plate (Sarstedt), and the plate was kept in a +27 ºC incubator (Memmert GmbH + Co.KG, Schwabach, Germany) for at least 30 min for cells to attach to the bottom of the wells. Solution A containing 10 µl of purified bacmid and 90 µl of Insect-XPRESS™ Protein-free Insect Cell Medium with L-glutamine (Lonza Group Ltd, Basel, Switzerland) was mixed with solution B containing 9 µl of Cellfectin® II Reagent (Thermo Fisher Scientific) and 91 µl of medium. The A+B mixture was incubated for 15-45 min at RT. After incubation, 800 µl of medium was added to the A+B mixture, and 1 ml of this solution was added on top of the cells in the 6-well plate. The cells were incubated at +27 ºC for 5 hours. The medium was removed, 2 ml of fresh medium was added, and incubation continued at +27 ºC for 3 days, followed by isolation of the primary stock. The medium was collected and centrifuged at 1000 x g for five minutes. Supernatant was isolated and kept in the dark at +4 ºC. A secondary stock was prepared by infecting 20 ml of cells (2 million cells/ml) with 300 µl of the primary stock and incubated at +27 ºC in a shaker for 3 days, followed by a centrifugation step at 2000 x g for five minutes. The supernatant was collected, and 1 ml of the secondary stock was used to infect 100 ml of cells (2 million cells/ml). Four hundred milliliters of culture was used for protein production.

The cultured Sf9 cells were centrifuged at 5000 rpm for 10 min, and the supernatant was collected and diluted 1:1 with a binding buffer, pH 8.0, containing 50 mM Na_2HPO_4 (Sigma-Aldrich, Inc.) and 0.5 M NaCl (Sigma-Aldrich, Inc.).

Prewashed Protino Ni-NTA Agarose (Macherey & Nagel GmbH & Co. KG) was added and mixed with a magnetic stirrer for two hours at RT. The solution was filtered, and the remaining resin was washed three times with a binding buffer containing 20 mM imidazole (Sigma-Aldrich, Inc.) and packed in a column. Protein was eluted from resin with an elution buffer containing 0.25 M imidazole. Molecular mass was estimated using SDS-PAGE with a 10 % gel, and the Precision Plus Protein™ Standards Dual Color (Bio-Rad Laboratories, Inc.) marker was used. The gel was run with MiniPROTEAN® TetraCell PowerPac® Basic (Bio-Rad Laboratories, Inc.) at 200 V for 30 min.

The produced protein was concentrated by centrifugation at 4000 rpm for 20 min using an Amicon Ultracel 30 k (Millipore, Merck KGaA, Darmstadt, Germany) tube. Protein concentration was measured with the Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific) according to the manufacturer's instructions.

4.7 Enzyme activity measurement (III)

The enzyme activity of zebrafish CA VI and the inhibitor study for AZA were carried out using an Applied Photophysics stopped-flow instrument under the supervision of Professor Claudiu T. Supuran at the University of Florence, Italy, according to the protocol described in article III.

4.8 Light scattering analysis (III)

Analysis of zebrafish CA VI-PTX protein was performed using a liquid chromatography instrument (CBM-20A, Shimadzu Corporation, Kyoto, Japan) equipped with an autosampler (SIL-20A), UV-VIS (SPD-20A) and fluorescence detector (RF-20Axs). The size of the protein was determined by light scattering analysis using a Malvern Zetasizer μV instrument (Malvern Instruments Ltd, Worcestershire, UK) running static light scattering (SLS) and dynamic light scattering (DLS) methods under the supervision of Associate Professor Vesa Hytönen at the University of Tampere as described in article III.

4.9 Mass spectrometry analysis (III)

Mass spectrometry experiments were performed at the University of Eastern Finland using a 12-T Bruker Solarix-XR FT-ICR mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany), equipped with an Apollo-II electrospray ionization (ESI) source and a dynamically harmonized ParaCell ICR-cell. The studies were carried out under the supervision of Professor Janne Jänis according to the protocol described in article III.

4.10 Behavioral monitoring (I, III)

Behavioral monitoring of mice was carried out using an IntelliCage system under the supervision of Dr. Vootele Voikar at the Neuroscience Center/Laboratory Animal Center of the University of Helsinki according to protocol described in detail in the article I.

Zebrafish larvae represented another animal model in which the role of CA VI was investigated. We measured the swimming pattern of *ca6* KD, uninjected WT, and random control MO-injected larvae at 4 dpf and 5 dpf. The larvae (approximately 10/flask) were placed in a 23 x 43 x 45 mm TC Flask T25 (Sarstedt) containing 40 ml of embryo medium at 3 dpf and allowed to acclimate for 24 h at +28.5 °C standard conditions. At 4 dpf and 5 dpf, their swimming patterns were observed by recording a one-minute video, with a printed 1 cm x 1 cm grid behind the flask. In total, the movement patterns of 284 zebrafish were recorded, measured, and analyzed using the MtrackJ plugin (Meijering et al, 2012) within the ImageJ program (Schneider et al, 2012). Movement analysis was performed by Harlan Barker, MSc, according to the protocol described in article III.

4.11 Histochemical staining and immunohistochemistry (I, II, III)

To investigate the effect of CA VI deficiency on tongue morphology, whole tongue samples were taken from WT (n=9) and *Car6*$^{-/-}$ (n=9) adult mice and prepared for staining. The samples were dissected and fixed in 4 % formaldehyde (Sigma-Aldrich, Inc.) overnight, followed by paraffin (Thermo Fisher Scientific) embedding. Tissue samples were sectioned to a thickness of five μm, placed on Menzel Gläzer microscope slides (Gerhard Menzel GmbH, Braunschweig, Germany) and attached

at 60 ˚C in an incubator (Memmert GmbH + Co. KG) for two hours. For morphological analysis, the paraffin was removed with xylene (J. T. Baker, A Division of Mallinckrodt Baker, Inc., Deventer, The Netherlands), and the sections were rehydrated in a descending ethanol series (Altia Oyj, Helsinki, Finland) and stained with Mayer's hematoxylin (Oy FF-Chemicals Ab, Haukipudas, Finland) and eosin (Reagena, Toivala, Finland). Dehydration steps were carried out using an ascending ethanol series, and the tissue sections were mounted using Pertex (Histolab Products AB, Göteborg, Sweden) covered with Deckgläser cover slips (Heinz Herenz Medizinalbedarf GmbH, Hamburg, Germany). Morphologies of different papillae (circumvallate, fungiform, and filiform) and von Ebner´s glands were evaluated by light microscopy using a Zeiss Axioskop 40 microscope (Carl Zeiss MicroImaging GmbH).

To study the apoptosis rate of tongue epithelial cells, we used a FragEL DNA Fragmentation Detection Kit, Fluorescent-TdT Enzyme (Calbiochem, EMD Chemicals, Inc., San Diego, California, USA) according to the manufacturer's instructions. Prior to staining, the sections were incubated with 20 µg/ml proteinase K (Thermo Fisher Scientific) at RT for 20 min. Prior to the equilibration step, positive control sections were treated with 1 µg/µl DNase I (Roche Diagnostics) in 1 mM $MgSO_4$ (Merck KGaA) and 1 × Tris-buffered saline (TBS) and incubated for 20 min at RT. After incubation in 1 x TdT Equilibration buffer for 15 min, the sections were covered with TdT Labeling Reaction Mixture. A negative control was prepared at the labeling step using dH_2O instead of an enzyme in the TdT Labeling Reaction mixture. Samples were incubated at +37 ˚C for 1.5 hours. Sections were mounted with Fluorescein-FragEL mounting medium and photographed using a Nikon Mikrophot-FXA microscope with a 450-490 nm filter (Nikon Instruments Europe B.V.) with original magnification × 100.

Immunohistochemistry for the proliferation marker Ki67 was performed using a Vectastain Elite ABC kit (Vector Laboratories, Inc., Burlingame, CA, USA) following the manufacturer's instructions. Prior to staining, the sections were boiled in a 0.01 M sodium citrate buffer, pH 6.0, for 20 min. Endogenous peroxidase activity was blocked with a 3 % H_2O_2 solution (J.T. Baker, Inc., Phillipsburg, NJ, USA) for 10 min. Nonspecific binding of the primary antibody was prevented by incubating samples in 10 % normal rabbit serum (NRS) (Thermo Fisher Scientific) as a blocking agent for 30 min. The sections were incubated overnight at +4 ˚C in a primary antibody solution containing monoclonal rat anti-mouse Ki67 (#M7249, Dako Denmark A/S, Glostrup, Denmark) diluted 1:100 in 1 % NRS, 0.1 % Tween-20 (Sigma-Aldrich, Inc.), and 1 × phosphate buffered saline (PBS). Biotinylated

rabbit anti-rat serum (Vector Laboratories, Inc.) diluted 1:1000 in 1 % NRS, 0.1 % Tween-20, and 1 × PBS was used as a secondary antibody, and the sections were incubated for 30 min at RT. Samples were then incubated in ABC reagent for 30 min. For precipitation, substrate 3, 3'-diaminobenzidine (DAB) (Invitrogen) was used for 3 min, and sections were counterstained using Mayer's hematoxylin (Oy-FF-Chemicals Ab). Washing steps were carried out using PBS with 0.1 % Tween-20. Dehydration and mounting steps were carried out as described above.

Immunohistochemistry of tissue samples collected from the trachea, lung and submandibular gland of adult WT and *Car6*$^{-/-}$ mice was carried out using a Vectastain Elite ABC kit (Vector Laboratories, Inc.) following the manufacturer's instructions. Prior to staining, endogenous peroxidase was removed with 3 % H_2O_2 (J. T. Baker) in methanol (Sigma-Aldrich, Inc.) for five min, and samples were blocked with 10 % normal goat serum (NGS) (Thermo Fisher Scientific). Sections were incubated with primary antibody solution containing rabbit anti-rat CA VI serum (Leinonen, Jukka et al, 2001) diluted 1:1000 in 2 % NGS, 0.1 % Tween-20, and 1 × PBS overnight at +4 ˚C. A biotinylated goat anti-rabbit immunoglobulin G (IgG) (Vector Laboratories, Inc.) diluted 1:500 in 2 % NGS, 0.1 % Tween-20, and 1 × PBS solution was used as the secondary antibody, and samples were incubated for 30 min at RT. Samples were then incubated in ABC reagent for 30 min. DAB treatment, counterstaining, washing steps, dehydration, and mounting steps were carried out as described above. Photography of the sections was performed using a Zeiss Axioskop 40 microscope (Carl Zeiss MicroImaging GmbH) with original magnification × 400.

An antibody against zebrafish CA VI-PTX was manufactured by Innovagen AB (Innovagen AB, Lund, Sweden) according to their standard immunization schedule, with boosters at 14, 28, 49, and 70 days. Pre-immune serum and three samples of polyclonal antiserum at day 41 (bleed 1), day 62 (bleed 2), and day 83 (bleed 3) were tested using dot blotting. A Bio-Dot®Microfiltration Apparatus (BioRad Laboratories, Inc., Hercules, California, USA) was used to attach 500 ng of produced and purified native zebrafish CA VI-PTX protein to a PROTRAN® nitrocellulose (NC) transfer membrane (Schleicher & Schuell GmbH, Dassel, Germany) according to the manufacturer's instructions. Prior to staining, nonspecific binding of the primary antibody was prevented by using diluted colostrum 1:10 in TBS with Tween 20 (TBST) as a blocking agent, for 30 min. Four serum samples were diluted 1:100 in TBST and were added to NC strips that were incubated at RT for one hour. A donkey anti-rabbit IgG horseradish peroxidase-linked antibody (Amersham Biosciences, GE Healthcare Life Sciences, Little Chalfont, UK) was diluted 1:25,000 in TBST and used as the secondary antibody. All washing steps were carried out

using TBST. Staining was carried out using the ImmPACTTM DAB Peroxidase Substrate Kit (Vector Laboratories, Inc.). Antiserum of bleed 2 was used for immunocytochemical analysis of adult zebrafish tissue samples.

The adult zebrafish were anaesthetized in a final concentration of 0.05 % Tricaine (Sigma-Aldrich, Inc.) and sacrificed by keeping them on ice for more than 10 min. Different organs were harvested under the microscope and immediately transferred to 4 % paraformaldehyde (Sigma-Aldrich, Inc.) in PBS for fixation for 24 h at +4 ºC. After fixation, the tissues were transferred to 20 % sucrose (Sigma-Aldrich, Inc.) in PBS and stored at +4ºC before embedding them in Tissue-Tek® O.C.T.™ Compound (Sakura Finetek Europe B.V., Alphen aan den Rijn, The Netherlands). The embedded samples were cut into 10-µm slices using a cryotome and were stored at -20 ºC until staining was performed. Prior to staining, the sections were fixed at +37 ºC overnight. Nonspecific binding of the primary antibody was prevented by using 1 % BSA (Sigma-Aldrich, Inc.) in PBS as a blocking agent. The sections were incubated for an hour in a primary antibody solution containing polyclonal rabbit anti-zebrafish CA VI-PTX (Innovagen AB, Lund, Sweden) diluted 1:100 in [1 % BSA in PBS]. Washing steps were carried out using [1 % BSA in PBS] three times for five min. Alexa Fluor® goat anti-rabbit IgG (Life Technologies, Carlsbad, USA) was used as a secondary antibody at the dilution of 1:1000 in PBS containing 1 % BSA and incubated for one hour. Sections were mounted with Vectashield Hard Set Mounting Medium with DAPI (Vector Laboratories, Inc.). All of the above procedures were carried out at RT and in the dark after the addition of the secondary antibody. The sections were photographed using a Zeiss LSM780 Laser Scanning Confocal Microscope with a Zeiss Cell Observer.Z1 microscope and Plan-Apochromat 40 x/1.4 (oil) objective, with a pulsed diode laser at 405 nm, a multiline Argon laser at 488 nm, and a Quasar spectral GaAsP PMT array detector (Carl Zeiss Microscopy GmbH, Goettingen, Germany). Images were analyzed with Zeiss ZEN2Lite.

4.12 Bioinformatics (III)

Dr. Martti Tolvanen, University of Turku, Finland, performed bioinformatic studies of zebrafish CA VI. In 2007, during preliminary studies, the presence of an additional PTX domain in the CA VI of four non-mammalian species was observed, but since no PTX domains were found in the proximity of any human or mouse CA, this observation was not followed up at that time. Later, we realized that the PTX

domain is, indeed, present in non-mammalian CA VI too consistently to be an annotation artifact. In 2015, non-mammalian CA VI sequences were retrieved from the NCBI (NCBI Resource Coordinators, 2016) nr protein database using BLASTP (Altschul et al, 1990), with human CA VI (ENSP00000366662 from Ensembl (Flicek et al, 2012) as the query sequence. Short sequences and sequences containing runs of X characters or long gaps in the alignment, which were interpreted as missing data for internal exons, were discarded. The final sequence set contained 78 sequences from 75 species. All sequences had complete PTX domains and were aligned to observe the conservation of two domains. To see if the orphan fragment Contig22468 was from the platypus genome (which contains the exon coding for a "CA VI-type" PTX domain), we performed a BlastN search in Ensembl (http://www.ensembl.org/Homo_sapiens/Tools/Blast?db=core). BlastN was run against the platypus genome with the full 11,311-nt sequence of supercontig:OANA5:Contig22468 as the query sequence.

Exon comparisons were made for *hCA6* (ENST00000377443), *hCA9* (ENST00000378357), *hCA12* (ENST00000178638 and ENST00000344366), *hCA14* (ENST00000369111) and zebrafish *ca6* (ENSDART00000132733) retrieved from Ensembl. The exon lengths preceding that coding for the PTX domain were noted from Ensembl transcripts for *hCRP* (ENST00000255030) and *hAPCS* (SAP, ENST00000255040). To visualize the conserved features in the region between the CA and PTX domains, a small alignment of CA VI sequences was constructed, followed by the creation of helical wheel diagrams by the PepWheel program of the EMBOSS suite (Rice et al, 2000).

Phylogenetic trees were built using the program MrBayes v 3.2 (Ronquist et al, 2012), and the corresponding analyses were performed by Dr. Csaba Ortutay. The first tree was constructed to show the evolutionary distances between CA 6, 9, 12 and 14 cDNA sequences, and the corresponding protein translations were used. The second tree was built using the catalytic domains of CA VI sequences only, and the third tree used the CA VI-associated PTX domains from selected species and human PTXs.

Prajwol Manandhar, MSc, built the 3D model of zebrafish CA VI under the supervision of Dr. Martti Tolvanen. The details of the tools used for building the model are listed in Table 6. Detailed experimental procedures can be found in article III. All bioinformatics tools used in this study are listed and described in Table 8.

4.13 Data analysis (I, II, III)

4.13.1 qRT-PCR (II, III)

To compare the expression of genes of interest of the WT and *Car6*$^{-/-}$ mice trachea samples, result normalization were carried out with a mathematical model (E.3 and E.4) for relative quantification by real-time qRT-PCR (Pfaffl, 2001), using *mActb* values as internal controls. The median value of WT data for each gene was used as a control for the normalization of the KO and WT data for that particular gene. Detailed experimental procedures can be found in article II.

$$E = 10^{-1/Slope} \qquad \text{(E. 3)}$$

$$R = \frac{E_{\text{Target}}^{\Delta CP(Control-Sample)_{Target}}}{E_{Actb}^{\Delta CP(Control-Sample)_{Actb}}} \qquad \text{(E. 4)}$$

The quantitative expression analysis of *ca6* mRNA was calculated between the WT and *ca6* morphant zebrafish at different stages of development and adult zebrafish tissue samples using formulas E.3 and E.4. For normalization, zebrafish *gapdh* was used as a housekeeping gene. Detailed experimental procedures can be found in III.

4.13.2 Statistical analyses (II, III)

The mouse cDNA samples for microarray were analyzed using GenomeStudio (2010.3) Gene Expression Module (1.8.0) software (Illumina, Inc.). The quality control of the data analysis included non-metric multidimensional scaling, dendrograms, and hierarchical clustering. The data were filtered according to the standard deviation (SD) of the probes. The data percentage that did not pass through the filter was adjusted to 99 %, implicating an SD value of almost 3. For the comparison of two groups (WT and *Car6*$^{-/-}$), statistical analysis was performed using the empirical Bayes t-test. At this point, data were filtered according to p-values. The remaining 750 probes in the trachea and 51 probes in the lung were further filtered according to up- and downregulated expressions of fold change (FC) with a cut-off value of ±1.4, which is considered to show a high correlation between microarray and qRT-PCR data (Morey et al, 2006). The functional annotation tool DAVID

(Database for Annotation, Visualization and Integrated Discovery) (http://david.abcc.ncifcrf.gov/) was used for the identification of enriched biological categories among the regulated genes, compared with all the genes present in Illumina's Sentrix® Mouse Ref-8 v2 Bead Chips (Illumina, Inc.).

The final qRT-PCR results for the gene of interest were presented as the expression relative to the housekeeping gene m*Actb*. Statistical analyses were performed by statistician Tiina Luukkaala, using IBM SPSS Statistics version 23.0.0 (IBM SPSS Statistics, IBM Corporation, Chicago, IL). The evaluation of relative expression values from *Car6*$^{-/-}$ and WT mice was carried out using the Mann–Whitney test because of the small sample size.

To measure the swimming activity of morphant vs. WT zebrafish, total distances traveled were calculated for individual larvae using the Matplotlib Python library (Hunter, 2007). Statistical analyses between relevant group pairs were performed to determine whether they could have been drawn from the same distribution using a two-sample Kolmogorov-Smirnov test using the Stats module of the SciPy Python library (S. van der Walt et al, 2011).

4.13.3 Histochemical staining and immunohistochemistry (I, II, III)

To study the morphology of mouse tongues, the number of fungiform papillae was evaluated from 21 sections/group (three sections/mouse). The anterior half of the tongue was examined for the presence of fungiform papillae. The number of fungiform papillae from each section was counted. Because of the small sample size and the potential non-normal distribution of section cell counts, the Mann-Whitney test (GraphPad Prism 5.01 for Windows, GraphPad Software, La Jolla, California, USA) was used to compare two groups.

The labeled nuclei of apoptotic cells were analyzed from seven *Car6*$^{-/-}$ and seven WT mouse tongues (two fields/tongue). Statistical comparison between the two groups was performed using the Mann-Whitney test (GraphPad Prism 5.01 for Windows). Cells that were positively stained for Ki67 were counted from one field/tongue and photographed with original magnification × 100 using Nikon Microphot-FXA microscope (Nikon Instruments Europe B.V., Amstelveen, Netherlands). Positively stained cells were calculated from one field/tongue. To determine the difference between the two groups, 21 fields from WT and 16 from KO mice were analyzed. Statistical analysis was performed using the Mann-Whitney test (GraphPad Prism 5.01 for Windows).

4.13.4 RNA-Seq and GO enrichment analysis

The expression of *CA6* mRNA in the lung was analyzed using two different data sets. *Car6* expression in mouse lung RNA-Seq data was retrieved from the ArrayExpress database (Kolesnikov et al, 2015), and data needed for the gene ontology (GO) enrichment analysis for human *CA6* was retrieved from the Medisapiens database (Kilpinen et al, 2008). Harlan Barker, MSc, performed the data analysis (tools are listed in Table 8), and detailed information about the analyses is described in article II.

Table 8. Bioinformatics and data analysis tools used in the thesis.

Program	Description	Reference	Used in
NCBI (*National Center for Biotechnology Information)*	Sequence database.	(NCBI Resource Coordinators, 2016)	II, III
Blastp (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins)	Nr protein database.	(Altschul et al, 1990)	III
Ensembl	Genome browser.	(Zerbino et al, 2018)	I, II, III
Clustal Omega	Sequence alignment.	(Sievers et al, 2011)	III
PAL2NAL web server v. 14 (http://www.bork.embl.de/pal2nal/)	Multiple protein sequences alignment.	(Suyama et al, 2006)	III
MrBayes v 3.2	Phylogeny analyses	(Ronquist et al, 2012)	III
APE (Analyses of Phylogenetics and Evolution) R package	Visualization tool.	(Paradis et al, 2004)	III
EMBOSS, PepWheel (http://www.bioinformatics.nl/cgi-bin/emboss/help/pepwheel)	Helical wheel diagram.	(Rice et al, 2000)	III
I-TASSER (Iterative Threading ASSEmbly Refinement) v. 4.0	Protein structure and function prediction.	(Roy et al, 2010)	III
HADDOCK (High Ambiguity Driven protein-protein DOCKing) 2.1 (http://haddock.science.uu.nl/)	Modeling of biomolecular complexes.	(de Vries et al, 2007)	III
CPORT (http://milou.science.uu.nl/services/CPORT/)	Prediction of protein-protein interface residues.	(de Vries & Bonvin, 2011)	III
MODELLER server	Modeling 3D structures of proteins and their assemblies.	(Webb & Sali, 2002)	III
UCSF Chimera (Extensible molecular modelling system) v. 1.10	Interactive visualization and analysis of molecular structures and related data.	(Pettersen et al, 2004)	III
MatchMaker tool	Superimposing protein or nucleic acid structures.	(Meng et al, 2006)	III
GraphPad Prism 5.01 for Windows (www.graphpad.com)	Combines scientific graphing, comprehensive curve fitting, understandable statistics, and data organization.	GraphPad Software, La Jolla California, USA	I, II, III
Chipster™ v2.9.0 (http://chipster.csc.fi)	Analysis software for high-throughput data containing over 350 analysis tools for next-generation sequencing (NGS), microarray, proteomics and sequence data	(Kallio et al, 2011)	II

Table 8 continues

Program	Description	Reference	Used in
VLAD Gene List Analysis and Visualization tool (http://proto.informatics.jax.org/prototypes/vlad/)	Test for the enrichment of ontology terms in a set of genes based on their annotations to bio-ontologies	(Richardson & Bult, 2015)	II
DAVID (Database for Annotation, Visualization and Integrated Discovery) (http://david.abcc.ncifcrf.gov/)	High-throughput and integrative gene functional annotation environment to systematically extract biological themes behind large gene lists	(Dennis et al, 2003, Huang da et al, 2009a, Huang da et al, 2009b)	II
The ArrayExpress Archive of Functional Genomics Data (http://www.ebi.ac.uk/arrayexpress)	International functional genomics database at the European Bioinformatics Institute (EMBL-EBI)	(Kolesnikov et al, 2015)	II
Bowtie of the Tophat package	An alignment program for aligning short DNA sequence reads to large genomes.	(Langmead et al, 2009)	II
Tophat of the Tophat package	Read-mapping algorithm designed to align reads from an RNA-Seq experiment to a reference genome without relying on known splice sites.	(Trapnell et al, 2009)	II
Cuffdiff module of the Cufflinks package	Includes algorithms that are not restricted by prior gene annotations and that account for alternative transcription and splicing.	(Trapnell et al, 2010)	II
Medisapiens database	Database of gene expression patterns	(Kilpinen et al, 2008)	II
Gene Ontology (http://geneontology.org)	Bioinformatics resource that supplies information about gene product function using ontologies to represent biological knowledge	(Ashburner et al, 2000, Gene Ontology Consortium, 2015)	II
ImageJ program	An open tool for scientific image analysis	(Schneider et al, 2012)	III
MtrackJ plugin	ImageJ plugin to facilitate tracking of moving objects in image sequences and the measurement of track statistics	(Meijering et al, 2012)	III
Matplotlib Python library	2D graphics package used for Python for application development, interactive scripting, and publication-quality image generation	(J. D. Hunter, 2007)	III
Chromas LITE v 2.1.1	Trace viewer for simple DNA sequencing projects.	(Technelysium Pty Ltd, South Brisbane, Australia)	III
GeneDoc v 2.6 (www.nrbsc.org)	Editing, analyzing and shading tool for multiple sequence alignments.	(Nicholas et al, 1997)	III

5 RESULTS

5.1 Gene expression studies

5.1.1 Microarray analysis

All of the extracted RNA samples of the trachea and lungs of eight WT and six *Car6*$^{-/-}$ mice that were sent for microarray analysis had an OD260/OD280 ratio 2.0 or higher, which is generally considered "pure" for RNA. Microarray data analysis results showed differentially expressed genes in the trachea and lung when compared between the WT and *Car6*$^{-/-}$ mice groups. All of the alterations of gene expression, either up- or downregulation, after filtering with an FC ±1.4, were statistically significant with p-values > 0.05. Five genes showed induced expression and 39 genes showed repressed expression in the trachea, whereas only single up- and downregulated genes were above or below the set FC threshold in the lung. A truncated list of differentially expressed genes in the trachea and lung, their symbols, gene descriptions, FC- and p-values are shown in Table 9.

5.1.2 qRT-PCR analysis

Results of the quantitative RT-PCR revealed statistical significance for genes *Faim3*, *Sftpc*, and *Dcpp1* in the trachea and *Dbp* in the lung. The relative expression values of *Faim3* and *Sftpc* showed induced expression, whereas *Dcpp1* and *Dbp* showed reduced expression in concordance with the microarray results. The same phenomena were observed with the other genes of interest: *Ucp1* and *Nppa* in the trachea and *Gdpd3* in the lung were upregulated, and *Lyz1* and *Calca* in the trachea were downregulated in both the qRT-PCR and microarray analyses. The relative expression ratios showing the median with the range for genes analyzed by qRT-PCR from *Car6*$^{-/-}$ and WT mice are shown as boxplots in Figure 5.

Table 9. Microarray results of differentially expressed genes in the trachea of *Car6*$^{-/-}$ mice compared with WT mice

TRACHEA

Gene	GenBank No	Description	FC	p-value
Sftpc	NM_011359	surfactant associated protein C	2.13	0.0303
Nppa	NM_008725	natriuretic peptide precursor type A	1.87	0.0451
Ucp1	NM_009463	uncoupling protein 1 (mitochondrial, proton carrier)	1.79	0.0072
Myl4	NM_010858	myosin, light polypeptide 4	1.46	0.0354
Faim3	NM_026976	Fas apoptotic inhibitory molecule 3	1.41	0.0101
Myh4	NM_010855	myosin, heavy polypeptide 4, skeletal muscle	-2.05	0.0036
Gp2	NM_025989	glycoprotein 2 (zymogen granule membrane)	-2.17	0.0191
Lipf	NM_026334	lipase, gastric	-2.25	0.0259
Expi	NM_007969	extracellular proteinase inhibitor	-2.34	0.0200
Dmbt1	NM_007769	deleted in malignant brain tumors 1	-2.52	0.0192
Tpo	NM_009417	thyroid peroxidase	-2.73	0.0198
Dcpp3	NM_001077633	demilune cell and parotid protein 3	-2.75	0.0131
Tg	NM_009375	thyroglobulin	-3.07	0.0148
Dcpp2	NM_001039238	demilune cell and parotid protein 2	-3.39	0.0055
Dcpp1	NM_019910	demilune cell and parotid protein 1	-3.48	0.0055
Calca	NM_007587	calcitonin/calcitonin-related polypeptide, alpha	-3.48	0.0056

LUNG

Gene	GenBank No	Description	FC	P-value
Gdpd3	NM_024228	glycerophosphodiester phosphodiesterase domain containing 3	1.45	0.0005
Dbp	NM_016974	D site albumin promoter binding protein	-1.45	0.0241

The quantitative expression analysis of *ca6* morphant zebrafish was carried out at different stages of development. The levels of *ca6* mRNA were consistently higher in the *ca6* KD embryos compared with those in the WT, possibly due to compensatory upregulation of the gene caused by the absence of the CA VI protein. The mRNA expression in wild-type embryos was highest at 24 hpf, while the peak in the expression of *ca6* was seen at 48 hpf in the *ca6* morphant embryos.

To gain further insights into *ca6* expression in zebrafish, the expression pattern in different tissues of adult zebrafish was measured. The relative expression was found to be prominent in the fins/tail (relative expression level 214.82) and brain (293.46). Lower levels of expression were observed in the gills (68.70), kidney (19.44), teeth (6.81), skin (4.32), and spleen (2.37), whereas very faint signals were seen in the swim bladder (0.61), intestine (0.41), pancreas (0.27), liver (0.02), eggs (0.01), and heart (0.01). The qPCR results in zebrafish are described in detail in article III.

5.1.3 Computational analysis

Computational analyses of differentially expressed genes in the trachea were carried out using the VLAD and DAVID tools. The VLAD results revealed changes in the following biological phenomena: a metabolic process, biological regulation, single-organism process (genes with induced expression), and immune response in mucosal-associated lymphoid tissue. Significant changes were also observed in the cellular component analysis with changes in the genes encoding cytoskeletal, cytoplasmic and extracellular proteins. The analyses of molecular functions indicated differentially expressed genes in the processes linked to transmembrane transporter, antioxidant and catalytic activity, protein binding, and structural constituents of cytoskeleton.

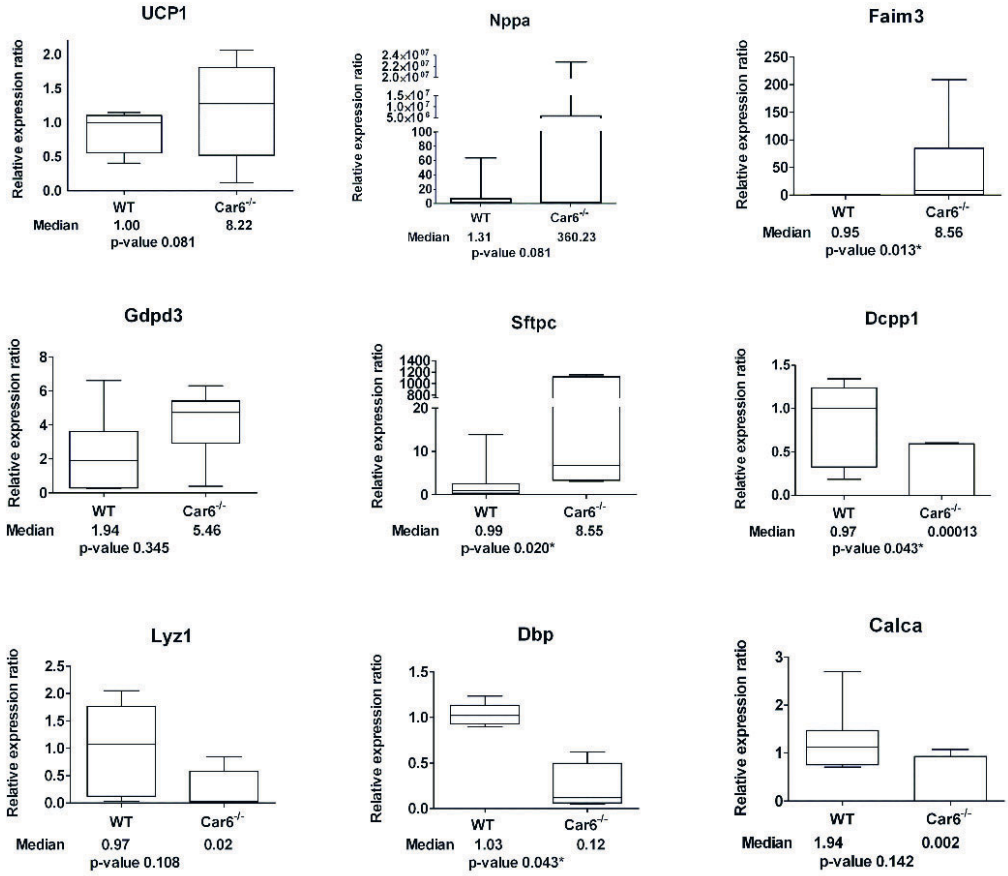

Figure 5. Relative expression ratios of the genes of interest in *Car6*$^{-/-}$ and WT mice based on qRT-PCR analysis. The median values of the relative expression values for both groups are shown in charts, and statistical comparison results with *p<0.05 were considered statistically significant.

The DAVID functional table of annotations reported 30 genes with enriched terms. The genes with induced expression in the trachea were functionally categorized by the associated terms: *Myl4* (muscle protein, myosin, and motor protein), *Nppa* (secreted, a signal, a hormone, and disulfide bond), *Sftpc* (secreted), *Faim3* (signal and disulfide bond). The genes with reduced expression were categorized mostly as signal (15 genes), secreted (10 genes), and disulfide bond (10 genes); thus, many of the repressed genes were annotated with several functional category terms.

As seen from the microarray analysis, the CA VI-deficient mice showed no significant alterations of gene expression in the lung samples. RNA-Seq data retrieved from databases also showed that *Car6* is not expressed in the normal lung tissues. However, positive *Car6* expression was registered in an occasional lung sample when the mice were infected with an influenza virus. Additionally, the influenza virus-infected mice also showed *Car6* expression in several bronchoalveolar lavage (BAL) samples and to a lesser extent in the interferon-treated lung tissue. The results are shown as fragments per kilobase of a transcript per million mapped reads (FPKM) in Figure 6.

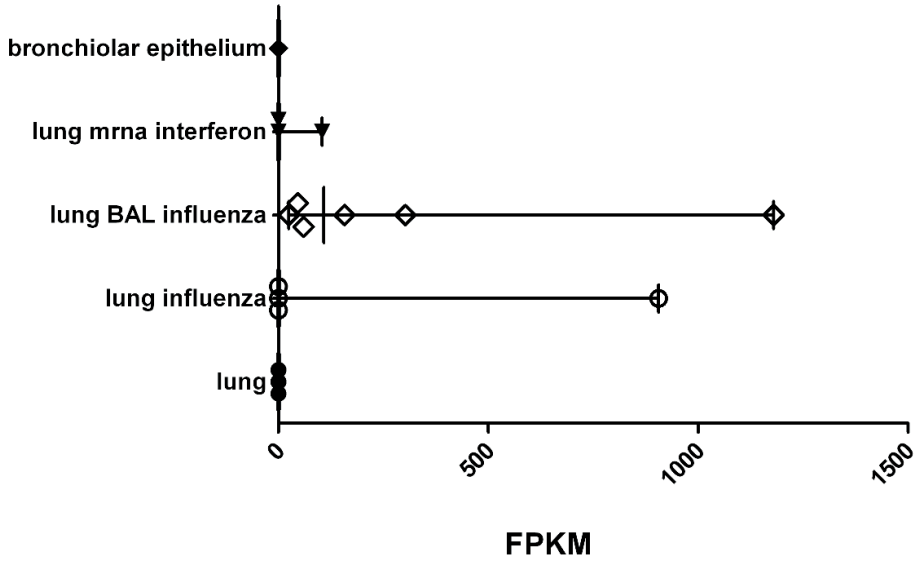

Figure 6. RNA-Seq results of *Car6* expression retrieved from the ArrayExpress database. RNA-Seq reads were mapped to CA genes in the mouse genome and compared with reference transcript structures and shown here as fragments per kilobase of a transcript per million mapped reads (FPKM).

GO enrichment analysis of the genes significantly co-expressed with human *CA6* in the lung tissue was performed using data from the Medisapiens Database. The analysis results identified three biological terms: axonemal dynein complex assembly (GO:0070286) over-represented more than 26-fold, cilium assembly (GO:0042384) over-represented more than 7-fold, and cilium organization (GO:0044782) over-represented more than 6-fold. All of these terms are related to the structure or function of cellular cilia.

5.2 Behavioral studies

5.2.1 *Car6*$^{-/-}$ mice

Comparison between the WT and *Car6*$^{-/-}$ mice showed no differences in body weight, activity (a number of corner visits in IntelliCage system), or total liquid consumption (a number of licks) (Figure 7). The mice showed similar preference for saccharin at the concentration of 5 g/l, whereas both groups identically avoided the 50 g/l concentration. Both groups showed similar preference for the water and avoidance of the 150 and 450 mM concentrations of NaCl. The *Car6*$^{-/-}$ mice avoided more 50 mM NaCl than the WT mice. Both groups showed equal preference for water and 0.1 mM citric acid, and they avoided the citric acid concentrations of 10 mM and 100 mM in a similar manner. Bitter taste testing resulted in differences between the groups. *Car6*$^{-/-}$ mice showed a preference for 0.003 mM quinine solution, whereas WT mice preferred water. Both WT and KO mice clearly avoided higher concentrations of quinine.

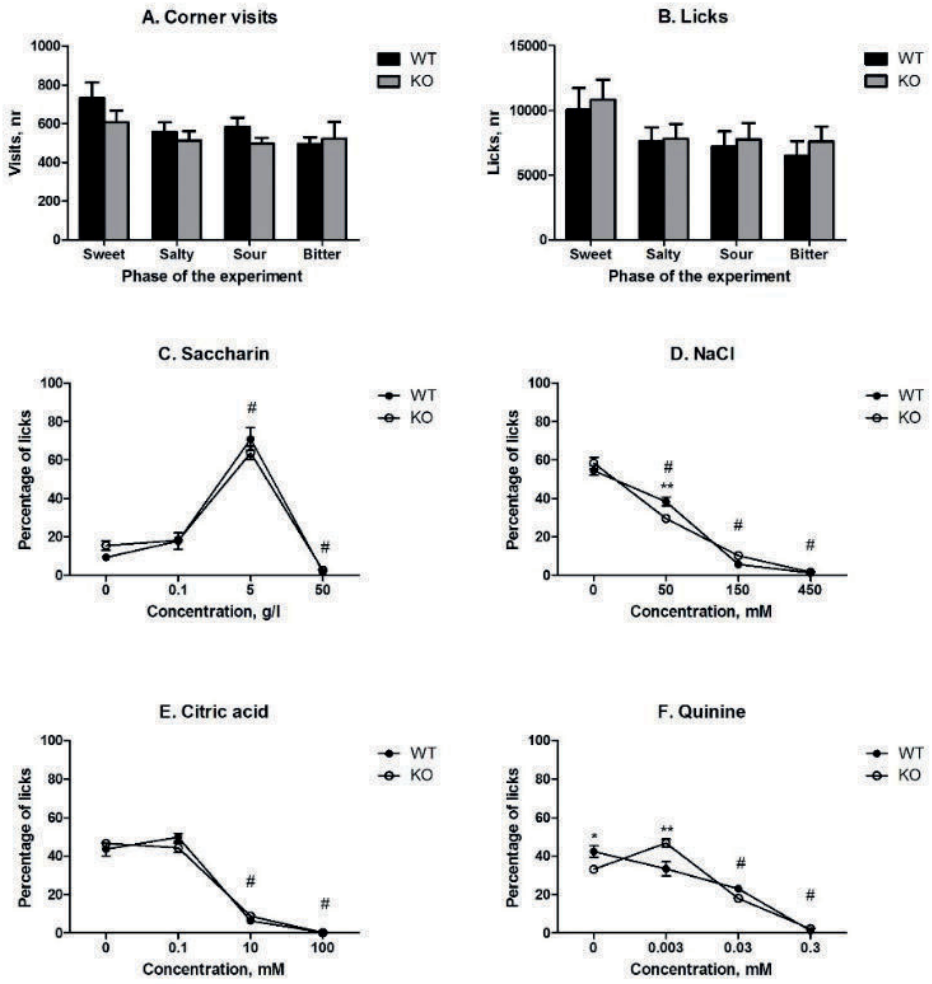

Figure 7. Taste preference assessment in the IntelliCage system. (A) The total number of corner visits in four-day experiment phases. (B) The total number of licks in four-day experiment phases. (C) The percentages of licks for different concentrations of the sweet (saccharin) solution, (D) salty (NaCl) solution, (E) sour (citric acid) solution and (F) bitter (quinine) solution. Statistically significant difference in taste preference shown between the genotypes: * $p < 0.05$ and ** $p < 0.01$. # $p < 0.01$ compared with water.

5.2.2 *ca6* knockdown zebrafish

Gene-specific antisense MOs were used to inhibit gene expression in zebrafish larvae. Two different MOs, one for translational blocking of *ca6* mRNA and the other for blocking intron splicing before exon 9, showed equally efficient KD in both kinds of *ca6* morphants. The *ca6* KD zebrafish embryos/larvae of one and three dpf were devoid of any notable morphological changes compared with the WT zebrafish (Figure 8A-J). Similarly, no morphological differences were observed between the uninjected and random control MO-injected embryos/larvae over the period of five days of development as shown in Figure 8A-L. An underdeveloped or deflated swim bladder was observed in *ca6* KD larvae at four dpf, whereas at five dpf the fish gained their ability to inflate the swim bladder and started to swim normally (Figure 8I-L).

For the swimming activity measurements of morphant vs. WT fish, total distances traveled for individual larvae were calculated and presented as boxplots in Figure 8Q and 8R. The day four KD larvae swam less (median 0.00 cm/min) than the day four WT larvae (median 13.80 cm/min, p-value 4.28×10^{-19}), and similarly the day five KD larvae swam less (median 4.75 cm/min) than the day five WT larvae (median 10.22 cm/min, p-value 1.16×10^{-7}). Both the swim bladder defect in four dpf larvae and the presence of CA VI in the adult zebrafish swim bladder suggest that CA VI is required either for swim bladder development or function. When CA VI expression was largely restored in five dpf larvae, the swimming pattern returned to almost normal (Figure 8M-P). Full details of the swimming data are shown in Table 4 of article III.

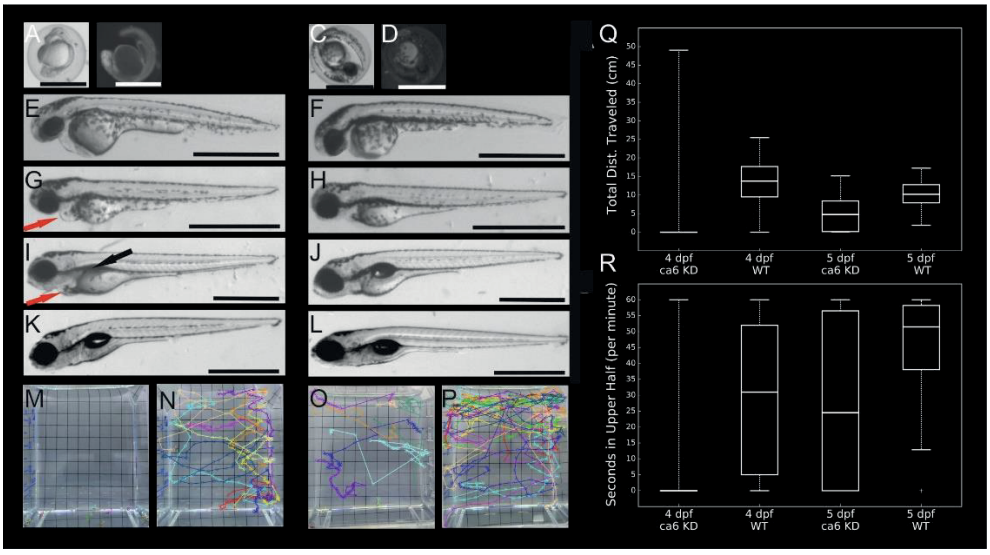

Figure 8. Swimming pattern measurements of zebrafish larvae over one minute of swimming. A-L show the representative images of developmental differences in *ca6* KD (A, B, E, G, I, K) and WT (C, D, F, H, J, and L) zebrafish larvae during 0-5 dpf. Morpholino-injected (A) and WT (C) embryos were photographed at 1 dpf using white light and screened under fluorescent light (B and D). On two dpf, the larvae (E, F) show no differences, but on three dpf, *ca6* KD larvae (G) show the signs of pericardial edema (red arrow) when compared with WT larvae (H). The swim bladder is deflated in four dpf *ca6* KD larvae (I, a black arrow), whereas WT larvae (J) have an inflated swim bladder. After morpholino detachment, the swim bladder structure/function normalizes on the fifth dpf, showing an inflated swim bladder for both the *ca6* KD (K) and WT larvae (L). The swimming distances of each *ca6* KD (M, N) and WT (O, P) fish were tracked and measured at day four (M, O) and five (N, P). Boxplots (Q) show the total distances traveled based on the video analyses of zebrafish larvae and the time the fish spent in the upper half of the tank in seconds/min (R).

5.3 Histological and immunocytochemical studies

Histological tongue samples were analyzed from *Car6*$^{-/-}$ and WT mice, including the morphology and number of taste buds and different types of papillae. Figure 9A-F shows representative images of different analysis results. Tongue samples showed no morphological changes in the von Ebner´s glands, nor did the circumvallate, fungiform and filiform papillae exhibit changes between the two groups of mice. As expected, immunohistochemistry for the proliferation marker Ki67 showed positive nuclei in the tongue epithelium of all specimens, and again no difference was observed between the *Car6*$^{-/-}$ and WT mice.

Trachea and lung tissue samples stained with anti-rat CA VI serum showed the absence of CA VI protein in the submandibular gland of *Car6*$^{-/-}$ mice, whereas the

corresponding samples of the WT mice presented clear positive CA VI staining in the acinar cells. Both trachea and lung tissue samples showed only faint background staining. No morphological differences were observed between the *Car6*$^{-/-}$ and WT mice in the submandibular gland, trachea, and lung tissues. Figure 9 G-L shows representative images of staining in different tissues.

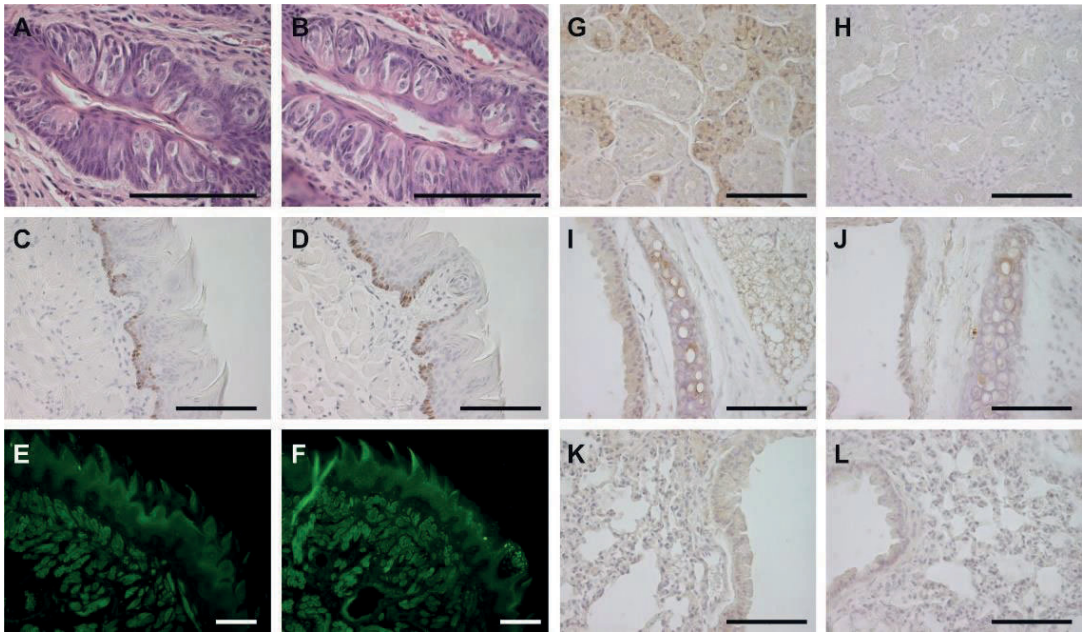

Figure 9. Histological and immunohistochemical staining of tongue sections from *Car6*$^{-/-}$ and WT mice. No significant differences in taste bud morphology are observed in the HE-stained samples of WT (A) and CA VI-deficient (B) mice. Immunohistochemical staining of the WT (C) and *Car6*$^{-/-}$ (D) mouse tongue samples using the proliferation marker Ki67 demonstrates a number of positive nuclei located in the basal layer of the stratified epithelium. The level of apoptosis in the tongue specimens analyzed by a DNA fragmentation detection kit indicated no differences between WT (E) and (F) CA VI-deficient mice. Positive CA VI staining is located in the acinar cells of the submandibular gland section in a WT mouse (G), whereas the enzyme is absent in the submandibular gland sample from a *Car6*$^{-/-}$ mouse (H). Weak background staining is seen in both WT (I) and *Car6*$^{-/-}$ (J) trachea sections. Lung sections show similar background staining for both WT (K) and CA VI-deficient mice (L). Scale bar 100 µm.

Recombinant zebrafish CA VI-PTX protein was used to raise a rabbit polyclonal antiserum to show the cell surface localization of CA VI-PTX in various tissues. The strongest positive staining signal was seen on the cell surfaces in the skin, heart, gills, and swim bladder, while intracellular staining was detectable but weaker, as shown in Figure 10.

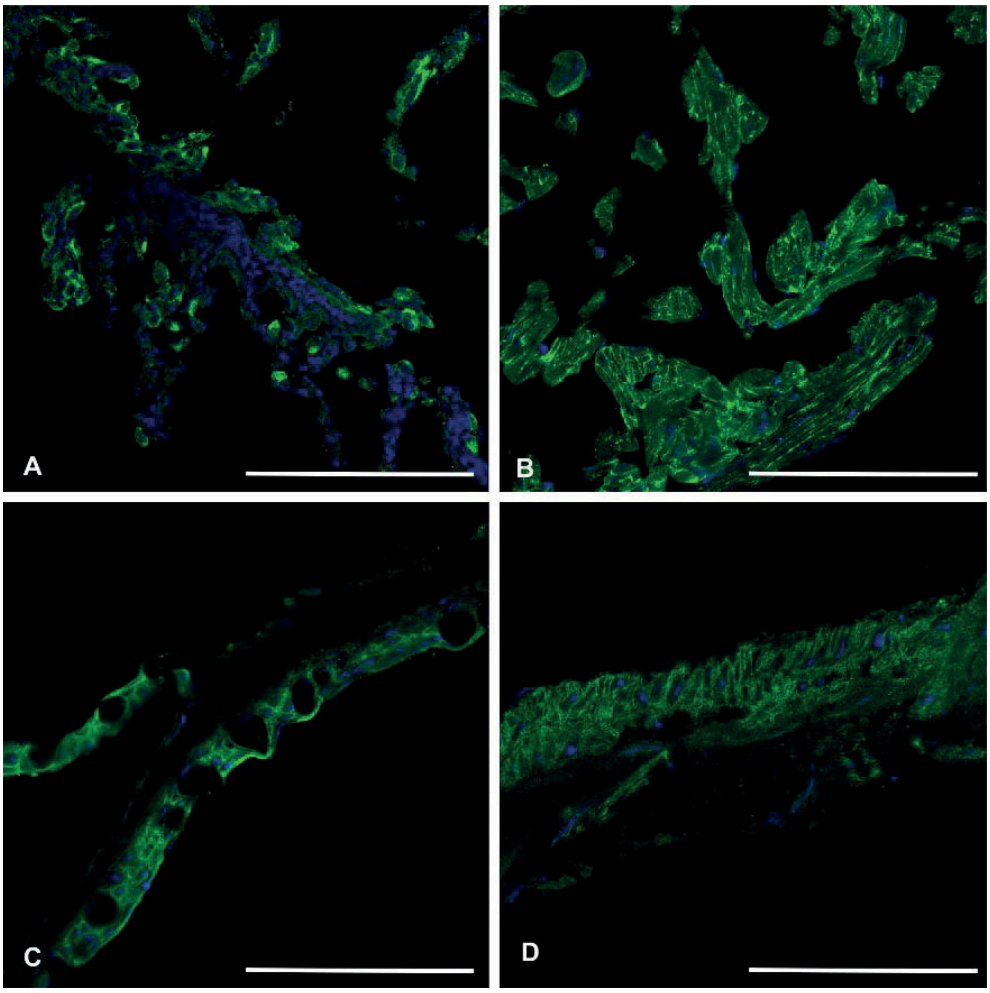

Figure 10. Immunostaining of adult zebrafish tissues using anti-zebrafish CA VI-PTX (green), and nuclear DAPI (blue) staining. The strongest signal was seen on the cell surface in gills (A), heart (B), skin (C), and swim bladder (D). Scale bar 100 µm.

5.4 Characterization of zebrafish CA VI-PTX

The presence of CA and PTX domains in non-mammalian CA VI sequences was previously documented in the Pfam database (Finn et al, 2016). To study the features of these non-mammalian CA VI sequences of different species, 78 CA VI protein sequences from 75 non-mammalian species were retrieved in NCBI GenPept, all of which had the C-terminal PTX domain. Sequence analysis showed that the PTX domain of CA VI is less conserved than the CA domain and that the longer, non-

mammalian isoforms (with a PTX domain) are orthologs of mammalian CA VI. Further analysis of the PTX domain sequences showed that the novel PTX domains are most closely related to short PTXs, CRP, and SAP (Garlanda et al, 2005, Bottazzi et al, 2010).

Mammalian CA VI proteins contain an additional C-terminal region of at least 25 residues. This region is unique compared to any other vertebrate CA isoforms and has an unknown structure. Non-mammalian CA VI proteins contain a spacer region between the CA and PTX domains, and this sequence is homologous to the mammalian C-terminal extension. The length of the exon coding for the spacer between the CA and PTX domains in zebrafish *ca6* is 84 bp, and the homologous exon coding the sequence of the human *CA6* is 83 bp. The penultimate exons coding for the regions containing the transmembrane (TM) helices in *CA9*, *CA12*, and *CA14* genes are 82, 85, and 85 bp in length, respectively.

Multiple sequence alignment indicated a pattern of hydrophobic residues repeating approximately every fourth residue in the region following the CA domain (the final part of mammalian CA VI or the segment between the CA and PTX domains of non-mammalian CA VI) (Figure 4 and S1 in the article III). The helical wheel visualizations indicated that when folded as an alpha helix, this region of human and zebrafish CA VI would form an APH with one side lined with mainly hydrophobic residues (in blue and lilac, Figure 4 and S1 in the article III). The APH region seemed to be a unique feature of CA VI and present in both non-mammalian and mammalian sequences. The findings concerning the APH region suggested that the exon coding for the cytoplasmic domain of ancestral CA VI was lost and replaced by the single exon coding for the PTX domain. The results also implied that the last exon in mammalian *CA6* and the penultimate exon of non-mammalian *CA6*, predicted to code for an APH, and the penultimate exons of *CA9*, *CA12*, and *CA14*, coding for the TM helix, are likely to share a common ancestry.

Based on phylogenetic analyses, we looked for an ornAna1 assembly in the Platypus Genome Browser Gateway (WUGSC 5.0.1/ornAna1 (Warren et al, 2008)) and found a genomic fragment not assigned to any chromosome. Contig22468 contains an exon, which codes for a PTX domain unlike any that we find in other mammalian species. This platypus PTX sequence is orthologous with the PTX sequences associated with CA VI in non-mammalian species. The phylogenetic trees are shown in original article III.

5.5 Characterization of recombinant CA VI-PTX protein

5.5.1 Sequencing analysis

To make recombinant CA VI-PTX protein, we first produced zebrafish *ca6* cDNA, with a sequence containing five synonymous substitutions compared with Ensembl ENSDART00000132733 (Supplemental information S5 in the original article III) and three unresolved bases leading to one unknown amino acid residue. Supplemental material S6 in original article III shows that the translation product is identical to the predicted 530-residue protein (Ensembl ENSDARP00000119189 or UniProt E9QB97). The cDNA sequence was submitted to the ENA database (http://www.ebi.ac.uk/ena) as LT724251 and its translation to UniProt as A0A1R4AHH7. The multiple sequence alignments of the recombinant zebrafish *ca6* cDNA and protein sequences are shown in article III.

5.5.2 Catalytic activity measurement of recombinant CA VI-PTX

The insect cell production of zebrafish CA VI-PTX resulted in high yields of the protein. After purification, the protein showed a single band close to the expected theoretical size of 58.107 kDa without glycans and excluding signal peptide. The molecular weight measured from the SDS-PAGE gel was 58.6 kDa. The activity measurements of the produced protein were carried out in the presence or absence of AZA. The kinetic parameters of CA VI-PTX CA activity (k_{cat} of 8.9 x10^5 s^{-1} and k_{cat}/K_m of 1.3 x 10^8 M^{-1} x s^{-1}) were then compared with those of thoroughly investigated Cas: the cytosolic and ubiquitous human isozymes hCA I and hCA II (Table 10).

Table 10. Kinetic parameters for CO_2 hydration reaction catalyzed by some human α-CA isozymes (hCA I, hCA II, and zebrafish CA VI-PTX, at 20 °C and pH 7.5), and their inhibition data with AZA, a clinically used drug.

Enzyme	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (M^{-1}x s^{-1})	K_I(AZA) (nM)
hCA I [a]	2.0x10^5	4.0	5.0 x 10^7	250
hCA II [a]	1.4x10^6	9.3	1.5 x 10^8	12
CA VI-PTX [b]	8.9x10^5	6.5	1.3 x 10^8	5

[a] Human recombinant isozymes, stopped flow CO_2 hydrase assay method (pH 7.5) (Innocenti et al, 2009).
[b] Recombinant enzyme, stopped flow CO_2 hydrase assay method (pH 7.5), this work.

5.5.3 Light scattering analysis

The recombinantly produced zebrafish CA VI-PTX was used for molecular size estimation by SLS and DLS after liquid chromatography. The results of the analysis are shown in Figure 7 in original article III, representing the main peak (black curve according to A280) eluting in a 1.52-ml retention volume by gel filtration analysis and associated with the SLS intensity peak with identical shape. An estimate of the protein MW by SLS was 280 ± 11 kDa for the peak, and the estimate was homogeneous throughout the elution peak (near-horizontal line across the peak in dark gray). The hydrodynamic radius (Rh) for the eluted peak indicated a 7.69 ± 0.29 nm particle size based on DLS data collection. The particle size was consistent with the determined molecular weight. The MW estimate from the gel filtration was slightly smaller (214 ± 10 kDa), possibly due to a non-globular shape of the molecule. A small peak (retention volume ~1.1 ml) before the main peak was observed, indicating the presence of aggregated protein (<5 % of the sample). The light scattering analyses combined with gel filtration results clearly indicate oligomeric assembly for the protein – most probably a pentameric form.

5.5.4 Mass spectrometry analysis

The structural characterization of trypsin-digested CA VI-PTX was performed by mass spectrometry under non-reducing conditions to preserve disulfide bonds in the structure. The digestion resulted in 64 specific tryptic peptides with 97 % sequence coverage (Figure 8 and 9 in the original article III). Three disulfide bonds were confirmed within the sequence. The confirmed disulfide bonds observed in the CA domain were Cys44/226, with bonds Cys 352/408 and Cys 487/518 in the PTX domain.

The CA VI-PTX polypeptide contains four putative N-glycosylation sites (Asn210, Asn258, Asn339, and Asn394). Asn258 of the CA domain is linked to a core-fucosylated oligomannose type glycan GlcNAc2(Fuc)Man3, as shown in Figure 11. Asn339 of the PTX domain carries an oligomannose type glycan, GlcNAc2Man3. Verification of these glycosylation sites and glycan structures was further carried out using CID-MS/MS on glycopeptides [248–266] and [331–347]. The corresponding masses of the peptides (3416.5084 Da and 2819.3059 Da) suggested the glycan structures shown in Figure 11.

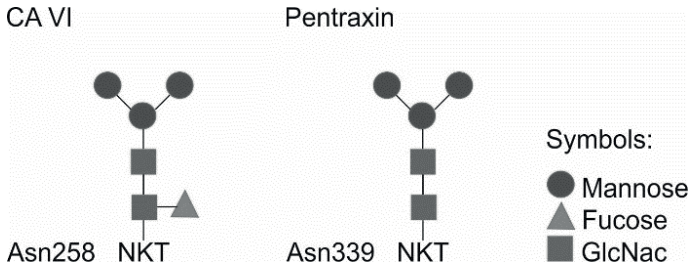

Figure 11. Glycans observed in the CA VI domain (left) and pentraxin domain (right).

The tryptic peptide [191–216] involving Asn210 showed a molecular mass of 2988.4744 Da. The result suggested that this particular Asn carries no glycan, and thus the peptide is non-glycosylated. There was also a non-glycosylated peptide spanning the Asn394 residue of the PTX domain. Hence, the glycosylation of the recombinant zebrafish CA VI-PTX produced in insect cells is rather homogenous because no other glycan variants were observed among the peptides. The confirmed N-glycosylation sites (different colors indicating the occupied and unoccupied sites) are shown in Figure 12.

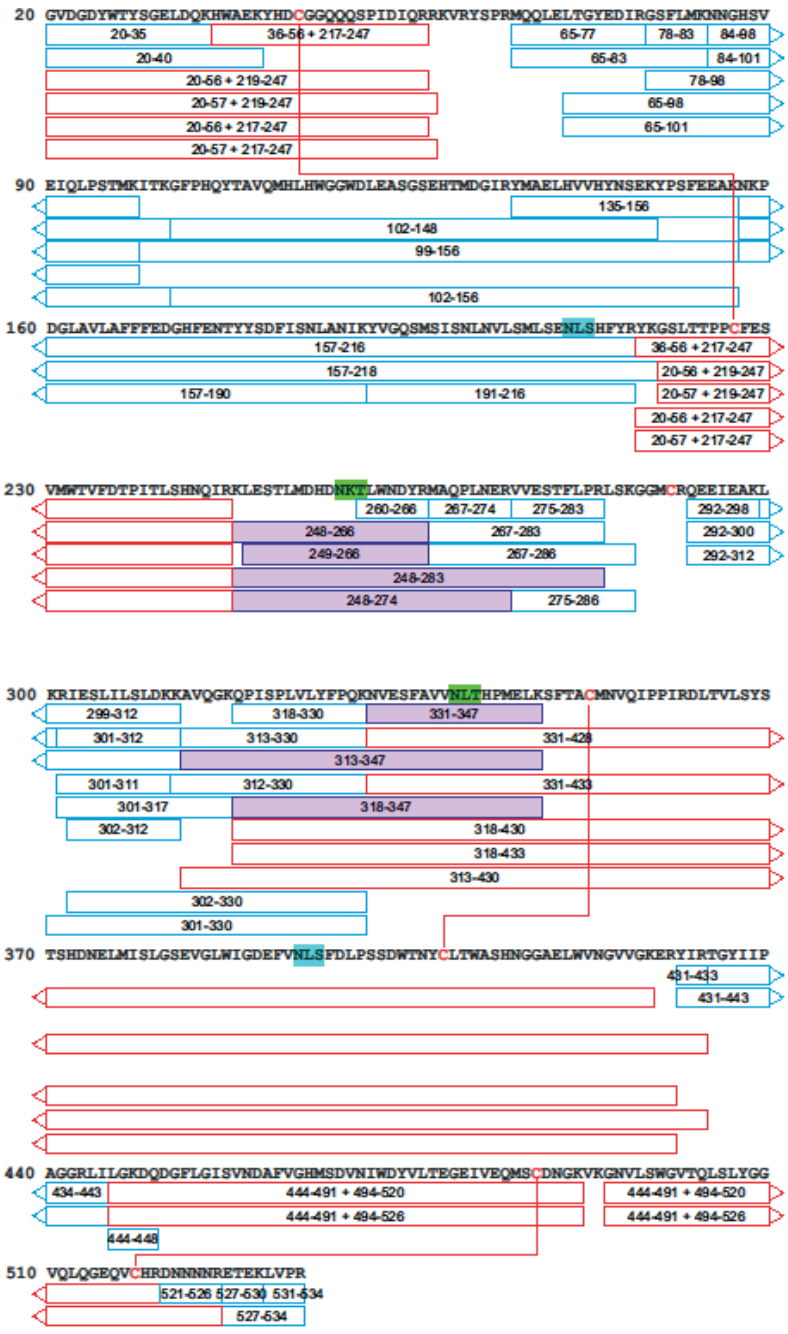

Figure 12. Tryptic peptide map of zebrafish CA VI-PTX. The blue boxes indicate identified tryptic peptides, including numbers of start and end residues. Red lines show the confirmed disulfide bonds and the corresponding peptides indicated with red boxes. Purple boxes indicate the glycopeptides, and different background colors indicate potential N-glycosylation sites (green: an occupied N-glycosylation site; blue: an unoccupied N-glycosylation site).

5.5.5 3D model of zebrafish CA VI-PTX

A homology-based model was built to combine the results obtained from bioinformatics, LC-SLS-DLS, and mass spectrometry analysis. Figure 13A shows the CA (top) and PTX (bottom) domains combined with an alpha helical model of the predicted APH region (pink). Variable residues 281 to 292 and 311 to 317 (yellow dotted lines), and residues 20 to 31 from the N-terminus, for which no template was available, were not modelled. In the figure, the active site cavity opens upwards, showing zinc-binding histidines (zinc not shown) of the CA domain as yellow sticks. Orange stick models show disulfide-forming cysteines, one in the CA domain and another in the PTX domain beta sheet, both being present in the templates. The third cysteine pair ended in close proximity, suggesting that one more disulfide bond would possibly lock the C-terminus of the PTX domain. This arrangement could be constructed by refining the model. The Cys290 is presumably unpaired and located in the un-modelled region.

The sites of conserved sequences in the multiple sequence alignment of 78 non-mammalian CA VI species predicted cysteine pairs to form disulfide bonds in both the CA domain (between amino acids 44/226) and in the PTX domain (352/408 and 487/518) (Article III, supplementary data 1: MSA column 51/234, 390/453, and 532/564, respectively). These disulfides were structurally verified. The disulfide in the CA domain is seen in all structures of extracellular CAs, [e.g., human CA VI in PDB 3FE4 (Pilka et al, 2012)], and the second disulfide in the PTX domain is homologous to that in short PTXs [e.g., human CRP in PDB 3PVN (Guillon et al, 2014)]. The third disulfide was supported by the proximity of cysteines in the 3D molecular model (Figure 13). Most importantly, the presence of these tree disulfides was also confirmed by mass spectrometry results, as described in the previous chapter. The N-glycosylated Asn residues (in the motif Asn-X-Ser/Thr) are all on the surface of the monomer, shown as spheres: unoccupied N-glycosylation sites are shown with blue color, and occupied N-glycosylation sites are shown with green color in Figure 13 A-C.

The pentameric model of zebrafish CA VI was introduced based on the pentamerization tendency of mammalian PTX domains. Five copies of the monomer were superimposed by the pentraxin domains following the pentameric structure of SAP (PDB 4AVS (Kolstoe et al, 2014)). Figure 13 B-C shows the individual monomers in different surface colors. The resulting shape of the modelled pentamer is a flat, planar five-point star (Figure 13 B-C).

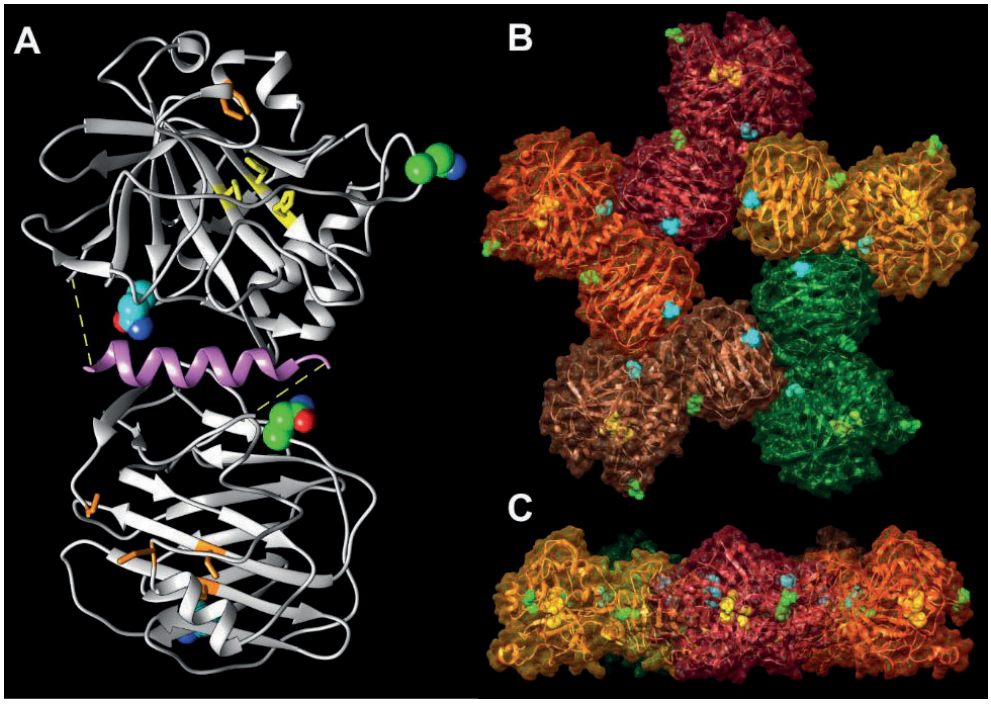

Figure 13. Molecular characterization of zebrafish CA VI-PTX. (A) A protomer of two domains, the CA VI on the top and the PTX at the bottom. Potential glycosylation site Asn residues are shown as spheres [unoccupied N-glycosylation site (blue), occupied N-glycosylation site (green)], active-site histidines (yellow), and assumed disulfide cysteines as sticks (orange). Region missing from the model is shown as a yellow dotted line and amphipathic helix in pink. (B) A front view of the pentamer model, showing protomers in different colors. (C) A side view of the pentamer structure.

6 DISCUSSION

6.1 *Car6* knockout mice showed differences in bitter taste perception and gene expression in the trachea and lung

Previous studies associating CA VI to taste modalities (Henkin et al, 1999a, Padiglia et al, 2010, Calo et al, 2011) inspired our group to test taste perception with our *Car6*$^{-/-}$ mouse model. The results of the behavioral tests demonstrated an interesting difference in bitter taste perception in the CA VI-deficient mice. These mice showed significant preference for 0.003 mM quinine solution, whereas the WT mice preferred water. Based on this result, we propose a role for CA VI in bitter taste perception, which contributes to the avoidance of bitter, potentially harmful, substances. In fact, PROP bitterness has been shown to associate with bitter taste receptor gene *TAS2R38* haplotypes, while the *CA6* gene SNP rs2274333 (A/G) genotype is associated with the density of fungiform papillae (Barbarossa et al, 2015). Our results on the comparison of WT and *Car6*$^{-/-}$ mice tongue samples neither showed differences in the von Ebner´s glands nor did they identify morphological changes in circumvallate, fungiform, or filiform papillae. It is important to note that we collected specimens from both WT mice and *Car6*$^{-/-}$ mice and predicted that complete deficiency of the CA VI enzyme could contribute to a different, more severe phenotype than that observed for a single polymorphism of the gene. Our material represented FFPE sections – not the whole tongue, which could be considered one potential source of error. Since we analyzed multiple serial sections from several mice per group, we consider the results reliable.

The bitter taste receptor, T2R38, is expressed in the human Sinonasal epithelial cells, making it an integral component of early response defense in the upper respiratory tract (Lee, R. J. et al, 2012). T2R38 triggers a rapid inter-kingdom signaling response, bridging functionally both the microbial secretions and the human upper airway epithelium. Following exposure to gram-negative quorum-sensing molecules, upper airway epithelial cells from individuals with one or two nonfunctional *T2R38* alleles have significantly blunted calcium dependent nitric oxide and ciliary responses. These individuals are more likely to be infected with gram-negative bacteria, such as *P. aeruginosa*, than those with two functional receptor

alleles. Experimental evidence about the possible functional association between T2R38 and CA VI proteins is still very limited. Findings of a recent study concluded that the *TAS2R38* diplotype, *CA6* rs2274333 and their combined genotype have no significant influence on dietary and alcohol intake, but genetic variations (AVI/AVI (A49, V262, I296) diplotype of the *TAS2R38* and *CA6* rs2274333 G allele) were a significant risk factor for colorectal cancer (Choi et al, 2017).

CA VI is expressed in the lower respiratory tract of rodents. Leinonen and coworkers showed positive signals for CA VI in the serous acinar and duct cells of the rat tracheobronchial glands as well as in the ciliated cells of the tracheobronchial surface epithelium and Clara cells (Leinonen, J. S. et al, 2004). Our results showed only faint signal for CA VI in the mouse trachea and lung, which may reflect different staining techniques and poor interspecies cross-reactivity of the antibody, but it may also be attributed to lower expression levels of CA VI in mouse compared to those in rat tissues. A previous study based on murine expression data showed *Car6* mRNA expression in an occasional lung sample after the mice were infected with influenza virus (Altboum et al, 2014). Positive expression was observed in several bronco-alveolar lavage samples and to a lesser extent in interferon-treated lung tissue.

The results presented in this thesis did not reveal any morphological changes of the tongue, trachea or lung attributable to CA VI deficiency, indicating that the enzyme has no major role in the developmental regulation of these tissues. The VLAD results, in contrast, suggested that CA VI might be involved in the regulation of several unforeseen biological processes associated with immune functions. Based on the observed alterations of gene expression, CA VI may contribute by some mechanism to the M-cell driven antigen transfer to mucosal-associated lymphoid tissue (MALT). M cells, which initiate antigen-specific immune responses, are specialized epithelial cells overlying the luminal lymphoid tissues, such as bronchus-associated lymphoid tissue (BALT) and other MALT follicles. MALT sites are anatomically separated: gut-associated lymphoid tissue (GALT), NALT, BALT, conjunctiva-associated lymphoid tissue (CALT), lacrimal duct-associated lymphoid tissue (LDALT), larynx-associated lymphoid tissue (LALT), and salivary duct-associated lymphoid tissue (DALT) have also been described (Cesta, 2006). These tissue sites are functionally connected, forming a common mucosal immune system, where antigen presentation and B-cell activation at one mucosal site can result in IgA secretion at the mucosal sites of different organs. The mucosal immune system can act independently of the systemic immune system, making the evaluation of MALT an important target in immunopathology (Kawanishi & Kiely, 1989, Cesta, 2006)

The lactating mammary glands and antibodies of milk, an integrated mucosal immune system, reflect the antigenic stimulation of MALT in the gut as well as in the airways. Indeed, it has been documented that secretory IgA (sIgA) from breast milk exhibits specificity for an array of common intestinal and respiratory pathogens in human (Brandtzaeg, 2002). The sIgA dimer is complexed to a glycosylated polypeptide chain, a secretory component (SC), which is derived from the polymeric Ig receptor (pIgR) expressed on the surface of epithelial cells. The sIgA-pIgR complex is secreted through the epithelia onto the mucosal lining of the GI and respiratory tract and into tears, saliva, and milk (Norderhaug et al, 1999). The SC glycans anchor sIgA to the mucosal lining of the epithelium, where bacteria can be sequestered (Phalipon et al, 2002, Arnold et al, 2007). Notably, immunoglobulins may also interact physically with CA VI. There is one previous study in which the association between IgG and CA VI was detected in human serum (Kivela et al, 1997). Although we do not yet understand the physiological significance of this interaction, it places CA VI into a strategic site of immune defense.

The GO enrichment analysis suggested that CA VI might have a role in enhancing cilial motion due to the coexpression of *Car6* and genes related to cilial function. The surface liquid covering the airway epithelium consists of a layer that is composed of viscous antimicrobial-rich mucus and contains glycosylated mucin proteins produced by the mucous cells of the submucosal exocrine glands and epithelial goblet cells (Jackson, 2001). Below the viscous mucus is a periciliary fluid layer that permits the submerged cilia of the airway epithelial cells to beat rapidly. The rapid and coordinated ciliary beating (frequency of ~8–15 Hz) enables the transport of the overlying mucus layer, along with any trapped debris and pathogens, to the oropharynx, where they are then swallowed or expectorated (Hariri & Cohen, 2016).

The observed changes in gene expression in *Car6*$^{-/-}$ mice suggest some role for CA VI in the lymphoid tissues, especially in the regulation of mucosal immune functions. Indeed, these findings have already opened new avenues for our future research. We have started to investigate whether CA VI deficiency induces any changes in the gut microbiome in our *Car6*$^{-/-}$ model. Future studies will hopefully increase our understanding of the role of CA VI and pinpoint a specific role for it in the regulation of immunity.

6.2 Characterization and localization of the novel zebrafish CA VI-PTX enzyme

Bioinformatics studies and experimental analyses built a coherent picture of the structure of a unique CA VI form containing a PTX domain. The PTX domain found to be associated with non-mammalian CA VI is a novel member of the PTX family. Our observations suggest that this novel pentraxin domain is most closely related to the short PTXs, CRP and SAP (Figure 2 of original article III), which are acute phase proteins in the innate immune system of mammals. The association of the CA domain with the PTX domain is a new observation for both the PTX and CA families. SAP and CRP are more closely similar to each other than either is to the CA-associated PTX domain. Based on a comparison of the exons coding for the spacer region after the CA domain of non-mammalian CA VI enzymes, we suggest that the exon coding for the cytoplasmic domain in ancestral CA VI was replaced by an exon coding for the PTX domain, and the TM helix transformed into an APH. The PTX domain was lost presumably in the therian mammalian lineage, leaving the APH in the C-terminus of CA VI.

Considering the mass of the glycosylated CA VI-PTX monomer (58.1 kDa), the LC–SLS–DLS results clearly confirmed that zebrafish CA VI is an oligomer. The obtained molecular mass (280 ± 11 kDa) by LC–SLS is slightly less than the estimated mass for a pentamer (290.5 kDa) calculated from the sequence. The molecular mass obtained from the gel filtration analysis was slightly lower (214 ± 10 kDa). This deviation may be attributable to the non-globular shape of the molecule or column interactions. The Rh calculated from light scattering (7.69 ± 0.29 nm; diameter 15.38 ± 0.58 nm) suggested a mass in the range of 364–434 kDa for a globular particle. The predicted 3D model of CA VI–PTX as a pentamer revealed the shape of a flat and roughly planar five-pointed star with a thickness of 4–5 nm and an approximate diameter of 15 nm (Figure 13B-C). The pentamer model is supported by the known pentamerization of related PTXs (CRP and SAP) (Hutchinson et al, 2000, Thompson et al, 1999).

The immunostaining results located CA VI-PTX mainly on cell surfaces in the zebrafish skin, heart, gills, and swim bladder. Mucosal homeostasis in teleosts is maintained by the gut, skin and gills, which contain mucosa-associated lymphoid tissue. According to the anatomical location, the MALT in teleost fish is subdivided into GALT, skin-associated lymphoid tissue (SALT), and gill-associated lymphoid tissue (GIALT) (Salinas et al, 2011). The GALT of higher vertebrates consists of both scattered and organized lymphoid tissue, whereas teleost fish have a lower level

of GALT organization. They also have a more diffusely organized immune system in their gut, containing lymphoid cells, macrophages, and eosinophilic and neutrophilic granulocytes (Rombout et al, 2011). The SALT in fish is mucosal lymphoid tissue in which the outermost layer of cells retains the capacity to divide (Salinas et al, 2011). Abundant secretory cells, including Malpighian cells, goblet cells, sacciform cells, and club cells, together produce mucus covering the epithelial cells, with ample biological functions (Zaccone et al, 2001). The gills, with large mucosal surfaces, constitute an important pathogen entry portal in fish but also represent a tissue that is capable of mounting an immune response (Campos-Perez et al, 2000, Holzer et al, 2003, Grove et al, 2006). Our results suggest that CA VI-PTX could participate in the mucosal defense systems in fish based on its distribution pattern and putative role in immune regulation.

6.3 *ca6* knockdown altered swimming pattern of zebrafish larvae

Our results showed the presence of CA VI in an adult zebrafish swim bladder. The major phenotype of *ca6* KD zebrafish larvae was a swim bladder deficiency and abnormal swimming pattern at the age of four dpf. We suggest that CA VI is required either for swim bladder development or swim bladder function. The fact that the fish gained the ability to swim normally at five dpf, and we observed no differences in swim bladder morphology between the WT and MO-injected fish at that age, suggest the phenotype is probably caused by delayed inflation of the swim bladder. The swim bladder is a hydrostatic organ in fish postulated as a homolog of the tetrapod lung, for which development can be divided into three phases: epithelial budding between 36 and 48 hpf, growth with the formation of two additional mesodermal layers up to 4.5 dpf, and inflation of the posterior and anterior chambers at 4.5 and 21 dpf, respectively (Winata et al, 2009).

6.4 Potential role of CA VI-PTX in immunity

Pentraxins are a type of pattern recognition protein and are part of the lectin family. The immune system of teleost fish includes most of the elements of the mammalian innate immune system (Magnadottir, 2006, Magor & Magor, 2001). Teleost fish lack lymph nodes and bone marrow; thus, the anterior part of the fish kidney is considered a functional ortholog of mammalian bone marrow, and it represents the

main hematopoietic lymphoid tissue of teleosts (Zapata & Amemiya, 2000). The innate immune molecules of lower vertebrates and invertebrates, such as teleost fish, include lectins and complement as well as natural killer cells, and the system is more diverse than the corresponding system in mammals (Sunyer et al, 1998, Vasta et al, 2011, Yoder & Litman, 2011).

Larvae can survive after fertilization with the help of innate immune responses because the adaptive immune system does not mature morphologically and functionally until 4–6 weeks post fertilization (Novoa & Figueras, 2012). This temporal separation system provides a way to study the vertebrate innate immune response *in vivo*, separately from the adaptive immune response (Stockhammer et al, 2009, Novoa & Figueras, 2012). Based on the findings in this thesis, we propose that CA VI–PTX in zebrafish may be needed for a novel type of membrane anchoring and immune function, and thus we decided to pursue these studies further. We have created a *ca6* mutation line using the CRISPR/cas9 technique to generate a full KO model to study the potential immune function of CA VI-PTX in zebrafish. We hope that these studies will finally provide convincing evidence for the role of CA VI in the regulation of immunity.

7 SUMMARY AND CONCLUSIONS

The general aim of this thesis was to reveal new information on the expression and function of secretory CA VI, the only secreted isozyme of the α-CA enzyme family and one of the major protein constituents of human saliva and milk. A *Car6*$^{-/-}$ KO mouse model was used to evaluate the role of CA VI for the taste modalities and gene expression levels in the trachea and lung. Over the course of our studies, a novel CA VI-PTX protein was identified in several species, including fish, frogs, and birds. Identification of this protein led us to produce recombinant zebrafish CA VI-PTX protein in insect cells, and the protein was biochemically and biophysically characterized by enzyme kinetic assays, SLS and DLS analyses, and mass spectrometry. Glycosylation sites and glycan structures were verified using CID-MS/MS. Sequence analyses were performed using bioinformatics tools, and the 3D structure was predicted by molecular modeling. The physiological significance of CA VI-PTX in zebrafish was studied by knocking down the corresponding gene, and the effect on the phenotype was examined. The localization of CA VI-PTX in different fish organs and tissues was investigated using immunofluorescence staining. The main results of this thesis are summarized below:

- CA VI is involved in bitter taste perception in mouse. CA VI may be one of the factors that contributes to the avoidance of bitter, potentially harmful, substances.
- CA VI deficiency leads to several significant alterations in gene expression levels in the trachea and lung.
- Several fish, frog, and bird species have a unique form of CA VI that is attached to a pentraxin domain.
- Phylogenetics and sequence analyses of CA VI-PTX provide a consistent hypothesis of the evolutionary history of domains associated with CA VI in mammals and non-mammals.
- A recombinant zebrafish CA VI-PTX protein shows high CA enzyme activity.

- The CA VI-PTX protein is probably pentameric in solution, contains three disulfide bonds, and is N-glycosylated through two of the four potential glycosylation sites.
- The KD of CA VI expression results in decreased buoyancy and swim bladder deflation in 4 dpf zebrafish larvae.

8 REFERENCES

Adler E, Hoon MA, Mueller KL, Chandrashekar J, Ryba NJ, & Zuker CS (2000) A novel family of mammalian taste receptors. *Cell* **100**: 693-702

Aidar M, Marques MR, Valjakka J, Mononen N, LehtimÃ¤ki T, Parkkila S, de Souza AP, & Line SRP (2013) Effect of Genetic Polymorphisms in ***CA6*** Gene on the Expression and Catalytic Activity of Human Salivary Carbonic Anhydrase VI. *Caries Res* **47**: 414-420

Alber BE & Ferry JG (1994) A carbonic anhydrase from the archaeon Methanosarcina thermophila. *Proc Natl Acad Sci U S A* **91**: 6909-6913

Aldred P, Fu P, Barrett G, Penschow JD, Wright RD, Coghlan JP, & Fernley RT (1991) Human secreted carbonic anhydrase: cDNA cloning, nucleotide sequence, and hybridization histochemistry. *Biochemistry* **30**: 569-575

Ali ES, Hua J, Wilson CH, Tallis GA, Zhou FH, Rychkov GY, & Barritt GJ (2016) The glucagon-like peptide-1 analogue exendin-4 reverses impaired intracellular Ca(2+) signalling in steatotic hepatocytes. *Biochim Biophys Acta* **1863**: 2135-2146

Altboum Z, Steuerman Y, David E, Barnett-Itzhaki Z, Valadarsky L, Keren-Shaul H, Meningher T, Mendelson E, Mandelboim M, Gat-Viks I, & Amit I (2014) Digital cell quantification identifies global immune cell dynamics during influenza infection. *Mol Syst Biol* **10**: 720

Alterio V, Pan P, Parkkila S, Buonanno M, Supuran CT, Monti SM, & De Simone G (2014) The structural comparison between membrane-associated human carbonic anhydrases provides insights into drug design of selective inhibitors. *Biopolymers* **101**: 769-778

Altschul SF, Gish W, Miller W, Myers EW, & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403-410

Arnold JN, Wormald MR, Sim RB, Rudd PM, & Dwek RA (2007) The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu Rev Immunol* **25**: 21-50

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, & Sherlock G (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**: 25-29

Aspatwar A, Tolvanen M, & Parkkila S (2010) Phylogeny and expression of carbonic anhydrase-related proteins. *BMC Molecular Biology* **11**: 25

Aspatwar A, Tolvanen ME, Jokitalo E, Parikka M, Ortutay C, Harjula SK, Ramet M, Vihinen M, & Parkkila S (2013) Abnormal cerebellar development and ataxia in CARP VIII morphant zebrafish. *Hum Mol Genet* **22**: 417-432

Aspatwar A, Tolvanen ME, Ojanen MJ, Barker HR, Saralahti AK, Bauerlein CA, Ortutay C, Pan P, Kuuslahti M, Parikka M, Ramet M, & Parkkila S (2015) Inactivation of ca10a and ca10b Genes Leads to Abnormal Embryonic Development and Alters Movement Pattern in Zebrafish. *PLoS One* **10**: e0134263

Aspatwar A, Tolvanen ME, & Parkkila S (2013) An update on carbonic anhydrase-related proteins VIII, X and XI. *J Enzyme Inhib Med Chem* **28**: 1129-1142

Avenet P & Lindemann B (1988) Amiloride-blockable sodium currents in isolated taste receptor cells. *J Membr Biol* **105**: 245-255

Baldini C, Giusti L, Ciregia F, Da Valle Y, Giacomelli C, Donadio E, Sernissi F, Bazzichi L, Giannaccini G, Bombardieri S, & Lucacchini A (2011) Proteomic analysis of saliva: a unique tool to distinguish primary Sjogren's syndrome from secondary Sjogren's syndrome and other sicca syndromes. *Arthritis Res Ther* **13**: R194

Barbarossa IT, Melis M, Mattes MZ, Calo C, Muroni P, Crnjar R, & Tepper BJ (2015) The gustin (CA6) gene polymorphism, rs2274333 (A/G), is associated with fungiform papilla density, whereas PROP bitterness is mostly due to TAS2R38 in an ethnically-mixed population. *Physiol Behav* **138**: 6-12

Bardow A, Moe D, Nyvad B, & Nauntofte B (2000) The buffer capacity and buffer systems of human whole saliva measured without loss of CO2. *Arch Oral Biol* **45**: 1-12

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, & Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Research* **28**: 235-242

Borghi GN, Rodrigues LP, Lopes LM, Parisotto TM, Steiner-Oliveira C, & Nobre-Dos-Santos M (2017) Relationship among alpha amylase and carbonic anhydrase VI in saliva, visible biofilm, and early childhood caries: a longitudinal study. *Int J Paediatr Dent* **27**: 174-182

Bottazzi B, Doni A, Garlanda C, & Mantovani A (2010) An integrated view of humoral innate immunity: pentraxins as a paradigm. *Annu Rev Immunol* **28**: 157-183

Bottazzi B, Inforzato A, Messa M, Barbagallo M, Magrini E, Garlanda C, & Mantovani A (2016) The pentraxins PTX3 and SAP in innate immunity, regulation of inflammation and tissue remodelling. *J Hepatol* **64**: 1416-1427

Brandtzaeg P (2002) The secretory immunoglobulin system: regulation and biological significance. Focusing on human mammary glands. *Adv Exp Med Biol* **503**: 1-16

Calo C, Padiglia A, Zonza A, Corrias L, Contu P, Tepper BJ, & Barbarossa IT (2011) Polymorphisms in TAS2R38 and the taste bud trophic factor, gustin gene co-operate in modulating PROP taste phenotype. *Physiol Behav* **104**: 1065-1071

Campos-Perez JJ, Ward M, Grabowski PS, Ellis AE, & Secombes CJ (2000) The gills are an important site of iNOS expression in rainbow trout Oncorhynchus mykiss after challenge with the gram-positive pathogen Renibacterium salmoninarum. *Immunology* **99**: 153-161

Carter N, Shiels A, & Tashian R (1978) Carbonic anhydrase III isoenzyme from human and bovine muscle [proceedings. *Biochem Soc Trans* **6**: 552-553

Carter ND, Fryer A, Grant AG, Hume R, Strange RG, & Wistrand PJ (1990) Membrane specific carbonic anhydrase (CAIV) expression in human tissues. *Biochim Biophys Acta* **1026**: 113-116

Cesta MF (2006) Normal structure, function, and histology of mucosa-associated lymphoid tissue. *Toxicol Pathol* **34**: 599-608

Chandrashekar J, Hoon MA, Ryba NJ, & Zuker CS (2006) The receptors and cells for mammalian taste. *Nature* **444**: 288-294

Chegwidden WR & Carter ND (2000) Introduction to the carbonic anhydrases. *EXS* **(90)**: 14-28

Chen J, Hu L, Zhang F, Wang J, Chen J, & Wang Y (2017) Downregulation of carbonic anhydrase IV contributes to promotion of cell proliferation and is associated with poor prognosis in non-small cell lung cancer. *Oncol Lett* **14**: 5046-5050

Choi JH, Lee J, Oh JH, Chang HJ, Sohn DK, Shin A, & Kim J (2017) Variations in the bitterness perception-related genes TAS2R38 and CA6 modify the risk for colorectal cancer in Koreans. *Oncotarget* **8**: 21253-21265

Crocetti L, Maresca A, Temperini C, Hall RA, Scozzafava A, Muhlschlegel FA, & Supuran CT (2009) A thiabendazole sulfonamide shows potent inhibitory activity against mammalian and nematode alpha-carbonic anhydrases. *Bioorg Med Chem Lett* **19**: 1371-1375

Culp DJ, Robinson B, Parkkila S, Pan PW, Cash MN, Truong HN, Hussey TW, & Gullett SL (2011) Oral colonization by Streptococcus mutans and caries development is reduced upon deletion of carbonic anhydrase VI expression in saliva. *Biochim Biophys Acta* **1812**: 1567-1576

De Langhe E, Bossuyt X, Shen L, Malyavantham K, Ambrus JL, & Suresh L (2017) Evaluation of Autoantibodies in Patients with Primary and Secondary Sjogren's Syndrome. *Open Rheumatol J* **11**: 10-15

de Vries SJ & Bonvin AM (2011) CPORT: a consensus interface predictor and its performance in prediction-driven docking with HADDOCK. *PLoS One* **6**: e17695

de Vries SJ, van Dijk AD, Krzeminski M, van Dijk M, Thureau A, Hsu V, Wassenaar T, & Bonvin AM (2007) HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. *Proteins* **69**: 726-733

Del Prete S, Vullo D, Fisher GM, Andrews KT, Poulsen SA, Capasso C, & Supuran CT (2014) Discovery of a new family of carbonic anhydrases in the malaria pathogen Plasmodium falciparum--the eta-carbonic anhydrases. *Bioorg Med Chem Lett* **24**: 4389-4396

Dennis G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, & Lempicki RA (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* **4**: R60

DeSimone JA, Lyall V, Heck GL, & Feldman GM (2001) Acid detection by taste receptor cells. *Respir Physiol* **129**: 231-245

Edgar WM, Higham SM, & Manning RH (1994) Saliva stimulation and caries prevention. *Adv Dent Res* **8**: 239-245

Esberg A, Haworth S, Brunius C, Lif Holgerson P, & Johansson I (2019) Carbonic Anhydrase 6 Gene Variation influences Oral Microbiota Composition and Caries Risk in Swedish adolescents. *Sci Rep* **9**: 452-018-36832-z

Etzel KR, Hempel JD, & Koepsel RR (1997) Identification of zinc proteins in rat parotid saliva. *Arch Oral Biol* **42**: 173-179

Feldstein JB & Silverman DN (1984) Purification and characterization of carbonic anhydrase from the saliva of the rat. *J Biol Chem* **259**: 5447-5453

Fernley RT, Farthing J, & Cooper EJ (1995) Radioimmunoassay for salivary carbonic anhydrase in human parotid saliva. *Arch Oral Biol* **40**: 567-569

Fernley RT, Wright RD, & Coghlan JP (1991) Radioimmunoassay of carbonic anhydrase VI in saliva and sheep tissues. *Biochem J* **274 (Pt 2)**: 313-316

Fernley RT, Wright RD, & Coghlan JP (1979) A novel carbonic anhydrase from the ovine parotid gland. *FEBS Lett* **105**: 299-302

Ferry JG (2010) The gamma class of carbonic anhydrases. *Biochim Biophys Acta* **1804**: 374-381

Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, & Bateman A (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* **44**: D279-85

Fisher SZ, Maupin CM, Budayova-Spano M, Govindasamy L, Tu C, Agbandje-McKenna M, Silverman DN, Voth GA, & McKenna R (2007) Atomic crystal and molecular dynamics simulation structures of human carbonic anhydrase II: insights into the proton transfer mechanism. *Biochemistry* **46**: 2930-2937

Fleming RE, Crouch EC, Ruzicka CA, & Sly WS (1993) Pulmonary carbonic anhydrase IV: developmental regulation and cell-specific expression in the capillary endothelium. *Am J Physiol* **265**: L627-35

Fleming RE, Parkkila S, Parkkila AK, Rajaniemi H, Waheed A, & Sly WS (1995) Carbonic anhydrase IV expression in rat and human gastrointestinal tract regional, cellular, and subcellular localization. *J Clin Invest* **96**: 2907-2913

Flicek P, Amode MR, Barrell D, Beal K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fairley S, Fitzgerald S, Gil L, Gordon L, Hendrix M, Hourlier T, Johnson N, Kahari AK, Keefe D, Keenan S, Kinsella R, Komorowska M et al (2012) Ensembl 2012. *Nucleic Acids Res* **40**: D84-90

Forster R (2000) Remarks on the discovery of carbonic anhydrase. In Chegwidden WR, Carter N, Edwards Y (eds) pp 1-11. BirkhÃ¤user Basel

Fujikawa-Adachi K, Nishimori I, Taguchi T, & Onishi S (1999a) Human carbonic anhydrase XIV (CA14): cDNA cloning, mRNA expression, and mapping to chromosome 1. *Genomics* **61**: 74-81

Fujikawa-Adachi K, Nishimori I, Taguchi T, & Onishi S (1999b) Human mitochondrial carbonic anhydrase VB. cDNA cloning, mRNA expression, subcellular localization, and mapping to chromosome x. *J Biol Chem* **274**: 21228-21233

Garlanda C, Bottazzi B, Bastone A, & Mantovani A (2005) Pentraxins at the crossroads between innate immunity, inflammation, matrix deposition, and female fertility. *Annu Rev Immunol* **23**: 337-366

Gasperskaja E & Kucinskas V (2017) The most common technologies and tools for functional genome analysis. *Acta Med Litu* **24**: 1-11

Gene Ontology Consortium (2015) Gene Ontology Consortium: going forward. *Nucleic Acids Res* **43**: D1049-56

Geurts AM, Wilber A, Carlson CM, Lobitz PD, Clark KJ, Hackett PB, McIvor RS, & Largaespada DA (2006) Conditional gene expression in the mouse using a Sleeping Beauty gene-trap transposon. *BMC Biotechnol* **6**: 30-6750-6-30

Gilmour KM (2010) Perspectives on carbonic anhydrase. *Comp Biochem Physiol A Mol Integr Physiol* **157**: 193-197

Giusti L, Baldini C, Bazzichi L, Ciregia F, Tonazzini I, Mascia G, Giannaccini G, Bombardieri S, & Lucacchini A (2007) Proteome analysis of whole saliva: a new tool for rheumatic diseases--the example of Sjogren's syndrome. *Proteomics* **7**: 1634-1643

Grove S, Johansen R, Reitan LJ, & Press CM (2006) Immune- and enzyme histochemical characterisation of leukocyte populations within lymphoid and mucosal tissues of Atlantic halibut (Hippoglossus hippoglossus). *Fish Shellfish Immunol* **20**: 693-708

Guillon C, Bigouagou UM, Folio C, Jeannin P, Delneste Y, & Gouet P (2014) A staggered decameric assembly of human C-reactive protein stabilized by zinc ions revealed by X-ray crystallography. *Protein Pept Lett* **22**: 248-255

Gut MO, Parkkila S, Vernerova Z, Rohde E, Zavada J, Hocker M, Pastorek J, Karttunen T, Gibadulinova A, Zavadova Z, Knobeloch KP, Wiedenmann B, Svoboda J, Horak I, & Pastorekova S (2002) Gastric hyperplasia in mice with targeted disruption of the carbonic anhydrase gene Car9. *Gastroenterology* **123**: 1889-1903

Hageman GS, Zhu XL, Waheed A, & Sly WS (1991) Localization of carbonic anhydrase IV in a specific capillary bed of the human eye. *Proc Natl Acad Sci U S A* **88**: 2716-2720

Hariri BM & Cohen NA (2016) New insights into upper airway innate immunity. *Am J Rhinol Allergy* **30**: 319-323

Headings VE & Tashian RE (1971) Isoenzymes of carbonic anhydrase I from primate red blood cells. *Biochim Biophys Acta* **236**: 353-359

Heck GL, Mierson S, & DeSimone JA (1984) Salt taste transduction occurs through an amiloride-sensitive sodium transport pathway. *Science* **223**: 403-405

Henkin RI, Lippoldt RE, Bilstad J, & Edelhoch H (1975) A zinc protein isolated from human parotid saliva. *Proceedings of the National Academy of Sciences* **72**: 488-492

Henkin RI, Martin BM, & Agarwal RP (1999a) Decreased parotid saliva gustin/carbonic anhydrase VI secretion: an enzyme disorder manifested by gustatory and olfactory dysfunction. *Am J Med Sci* **318**: 380-391

Henkin RI, Martin BM, & Agarwal RP (1999b) Efficacy of exogenous oral zinc in treatment of patients with carbonic anhydrase VI deficiency. *Am J Med Sci* **318**: 392-405

Henkin RI, Velicu I, & Papathanassiu A (2007) cAMP and cGMP in human parotid saliva: relationships to taste and smell dysfunction, gender, and age. *Am J Med Sci* **334**: 431-440

Hilvo M, Innocenti A, Monti SM, De Simone G, Supuran CT, & Parkkila S (2008) Recent Advances in Research on the Most Novel Carbonic Anhydrases, CA XIII and XV. *Curr Pharm Des* **14**: 672-678

Hilvo M, Tolvanen M, Clark A, Shen B, Shah GN, Waheed A, Halmi P, Hanninen M, Hamalainen JM, Vihinen M, Sly WS, & Parkkila S (2005) Characterization of CA XV, a new GPI-anchored form of carbonic anhydrase. *Biochem J* **392**: 83-92

Holzer AS, Sommerville C, & Wootten R (2003) Tracing the route of Sphaerospora truttae from the entry locus to the target organ of the host, Salmo salar L., using an optimized and specific in situ hybridization technique. *J Fish Dis* **26**: 647-655

Hoon MA, Adler E, Lindemeier J, Battey JF, Ryba NJ, & Zuker CS (1999) Putative mammalian taste receptors: a class of taste-specific GPCRs with distinct topographic selectivity. *Cell* **96**: 541-551

Hooper LV, Beranek MC, Manzella SM, & Baenziger JU (1995) Differential expression of GalNAc-4-sulfotransferase and GalNAc-transferase results in distinct glycoforms of carbonic anhydrase VI in parotid and submaxillary glands. *J Biol Chem* **270**: 5985-5993

Hsieh MS, Jeng YM, Jhuang YL, Chou YH, & Lin CY (2016) Carbonic anhydrase VI: a novel marker for salivary serous acinar differentiation and its application to discriminate acinic cell carcinoma from mammary analogue secretory carcinoma of the salivary gland. *Histopathology* **68**: 641-647

Huang da W, Sherman BT, & Lempicki RA (2009a) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**: 1-13

Huang da W, Sherman BT, & Lempicki RA (2009b) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**: 44-57

Huang AL, Chen X, Hoon MA, Chandrashekar J, Guo W, Trankner D, Ryba NJ, & Zuker CS (2006) The cells and logic for mammalian sour taste detection. *Nature* **442**: 934-938

Hunter JD (2007) Matplotlib: A 2D Graphics Environment. *Computing in Science and Engg.* **9**: 90-95

Hutchinson WL, Hohenester E, & Pepys MB (2000) Human serum amyloid P component is a single uncomplexed pentamer in whole serum. *Mol Med* **6**: 482-493

Ichihara N, Tsukamoto A, Kasuya T, Shibata S, Nishita T, Murakami M, Amasaki H, & Asari M (2007) Gene expression of secretory carbonic anhydrase isozymes in striated ducts of canine salivary glands using laser microdissection system. *Anat Histol Embryol* **36**: 357-360

Innocenti A, Leewattanapasuk W, Muhlschlegel FA, Mastrolorenzo A, & Supuran CT (2009) Carbonic anhydrase inhibitors. Inhibition of the beta-class enzyme from the pathogenic yeast Candida glabrata with anions. *Bioorg Med Chem Lett* **19**: 4802-4805

Ivanov SV, Kuzmin I, Wei MH, Pack S, Geil L, Johnson BE, Stanbridge EJ, & Lerman MI (1998) Down-regulation of transmembrane carbonic anhydrases in renal cell carcinoma cell lines by wild-type von Hippel-Lindau transgenes. *Proc Natl Acad Sci U S A* **95**: 12596-12601

Iverson TM, Alber BE, Kisker C, Ferry JG, & Rees DC (2000) A closer look at the active site of gamma-class carbonic anhydrases: high-resolution crystallographic studies of the carbonic anhydrase from Methanosarcina thermophila. *Biochemistry* **39**: 9222-9231

J. D. Hunter (2007) Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering* **9**: 90-95

Jackson AD (2001) Airway goblet-cell mucus secretion. *Trends Pharmacol Sci* **22**: 39-45

Jiang W & Gupta D (1999) Structure of the carbonic anhydrase VI (CA6) gene: evidence for two distinct groups within the alpha-CA gene family. *Biochem J* **344 Pt 2**: 385-390

Jiao Y, Yan J, Zhao Y, Donahue LR, Beamer WG, Li X, Roe BA, Ledoux MS, & Gu W (2005) Carbonic anhydrase-related protein VIII deficiency is associated with a distinctive lifelong gait disorder in waddles mice. *Genetics* **171**: 1239-1246

Juvonen T, Parkkila S, Parkkila AK, Niemela O, Lajunen LH, Kairaluoma MI, Peramaki P, & Rajaniemi H (1994) High-activity carbonic anhydrase isoenzyme (CA II) in human gallbladder epithelium. *J Histochem Cytochem* **42**: 1393-1397

Kadoya Y, Kuwahara H, Shimazaki M, Ogawa Y, & Yagi T (1987) Isolation of a novel carbonic anhydrase from human saliva and immunohistochemical demonstration of its related isozymes in salivary gland. *Osaka City Med J* **33**: 99-109

Kallio MA, Tuimala JT, Hupponen T, Klemela P, Gentile M, Scheinin I, Koski M, Kaki J, & Korpelainen EI (2011) Chipster: user-friendly analysis software for microarray and other high-throughput data. *BMC Genomics* **12**: 507-2164-12-507

Karhumaa P, Leinonen J, Parkkila S, Kaunisto K, Tapanainen J, & Rajaniemi H (2001) The identification of secreted carbonic anhydrase VI as a constitutive glycoprotein of human and rat milk. *Proc Natl Acad Sci U S A* **98**: 11604-11608

Kaseda M, Ichihara N, Nishita T, Amasaki H, & Asari M (2006) Immunohistochemistry of the bovine secretory carbonic anhydrase isozyme (CA-VI) in bovine alimentary canal and major salivary glands. *J Vet Med Sci* **68**: 131-135

Kasuya T, Shibata S, Kaseda M, Ichihara N, Nishita T, Murakami M, & Asari M (2007) Immunohistolocalization and gene expression of the secretory carbonic anhydrase isozymes (CA-VI) in canine oral mucosa, salivary glands and oesophagus. *Anat Histol Embryol* **36**: 53-57

Kawakami K, Largaespada DA, & Ivics Z (2017) Transposons As Tools for Functional Genomics in Vertebrate Models. *Trends Genet* **33**: 784-801

Kawanishi H & Kiely J (1989) Immune-related alterations in aged gut-associated lymphoid tissues in mice. *Dig Dis Sci* **34**: 175-184

Kelly C, Nogradi A, Walker R, Caddy K, Peters J, & Carter N (1994) Lurching, reeling, waddling and staggering in mice--is carbonic anhydrase (CA) VIII a candidate gene? *Biochem Soc Trans* **22**: 359S

Kikutani S, Nakajima K, Nagasato C, Tsuji Y, Miyatake A, & Matsuda Y (2016) Thylakoid luminal theta-carbonic anhydrase critical for growth and photosynthesis in the marine diatom Phaeodactylum tricornutum. *Proc Natl Acad Sci U S A* **113**: 9828-9833

Kilpinen S, Autio R, Ojala K, Iljin K, Bucher E, Sara H, Pisto T, Saarela M, Skotheim RI, Bjorkman M, Mpindi JP, Haapa-Paananen S, Vainio P, Edgren H, Wolf M, Astola J, Nees M, Hautaniemi S, & Kallioniemi O (2008) Systematic bioinformatic analysis of expression levels of 17,330 human genes across 9,783 samples from 175 types of healthy and pathological tissues. *Genome Biol* **9**: R139-2008-9-9-r139. Epub 2008 Sep 19

Kim G, Lee TH, Wetzel P, Geers C, Robinson MA, Myers TG, Owens JW, Wehr NB, Eckhaus MW, Gros G, Wynshaw-Boris A, & Levine RL (2004) Carbonic anhydrase III

is not required in the mouse for normal growth, development, and life span. *Mol Cell Biol* **24**: 9942-9947

Kim JH, Parkkila S, Shibata S, Fujimiya M, Murakami G, & Cho BH (2013) Expression of carbonic anhydrase IX in human fetal joints, ligaments and tendons: a potential marker of mechanical stress in fetal development? *Anat Cell Biol* **46**: 272-284

Kim MR, Kusakabe Y, Miura H, Shindo Y, Ninomiya Y, & Hino A (2003) Regional expression patterns of taste receptors and gustducin in the mouse tongue. *Biochem Biophys Res Commun* **312**: 500-506

Kimoto M, Iwai S, Maeda T, Yura Y, Fernley RT, & Ogawa Y (2004) Carbonic anhydrase VI in the mouse nasal gland. *J Histochem Cytochem* **52**: 1057-1062

Kitade K, Nishita T, Yamato M, Sakamoto K, Hagino A, Katoh K, & Obara Y (2003) Expression and localization of carbonic anhydrase in bovine mammary gland and secretion in milk. *Comp Biochem Physiol A Mol Integr Physiol* **134**: 349-354

Kivela J, Laine M, Parkkila S, & Rajaniemi H (2003) Salivary carbonic anhydrase VI and its relation to salivary flow rate and buffer capacity in pregnant and non-pregnant women. *Arch Oral Biol* **48**: 547-551

Kivela J, Parkkila S, Metteri J, Parkkila AK, Toivanen A, & Rajaniemi H (1997) Salivary carbonic anhydrase VI concentration and its relation to basic characteristics of saliva in young men. *Acta Physiol Scand* **161**: 221-225

Kivela J, Parkkila S, Parkkila AK, Leinonen J, & Rajaniemi H (1999a) Salivary carbonic anhydrase isoenzyme VI. *J Physiol* **520 Pt 2**: 315-320

Kivela J, Parkkila S, Parkkila AK, & Rajaniemi H (1999b) A low concentration of carbonic anhydrase isoenzyme VI in whole saliva is associated with caries prevalence. *Caries Res* **33**: 178-184

Kohler K, Hillebrecht A, Schulze Wischeler J, Innocenti A, Heine A, Supuran CT, & Klebe G (2007) Saccharin inhibits carbonic anhydrases: possible explanation for its unpleasant metallic aftertaste. *Angew Chem Int Ed Engl* **46**: 7697-7699

Kolesnikov N, Hastings E, Keays M, Melnichuk O, Tang YA, Williams E, Dylag M, Kurbatova N, Brandizi M, Burdett T, Megy K, Pilicheva E, Rustici G, Tikhonov A,

Parkinson H, Petryszak R, Sarkans U, & Brazma A (2015) ArrayExpress update--simplifying data submissions. *Nucleic Acids Res* **43**: D1113-6

Kolstoe SE, Jenvey MC, Purvis A, Light ME, Thompson D, Hughes P, Pepys MB, & Wood SP (2014) Interaction of serum amyloid P component with hexanoyl bis(D-proline) (CPHPC). *Acta Crystallogr D Biol Crystallogr* **70**: 2232-2240

Krishnan D, Pan W, Beggs MR, Trepiccione F, Chambrey R, Eladari D, Cordat E, Dimke H, & Alexander RT (2018) Deficiency of Carbonic Anhydrase II Results in a Urinary Concentrating Defect. *Front Physiol* **8**: 1108

Lagerlof F & Oliveby A (1994) Caries-protective factors in saliva. *Adv Dent Res* **8**: 229-238

Lamkin MS & Oppenheim FG (1993) Structural features of salivary function. *Crit Rev Oral Biol Med* **4**: 251-259

Lane TW, Saito MA, George GN, Pickering IJ, Prince RC, & Morel FM (2005) Biochemistry: a cadmium enzyme from a marine diatom. *Nature* **435**: 42

Langmead B, Trapnell C, Pop M, & Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25-2009-10-3-r25. Epub 2009 Mar 4

Law JS, Nelson N, Watanabe K, & Henkin RI (1987) Human salivary gustin is a potent activator of calmodulin-dependent brain phosphodiesterase. *Proceedings of the National Academy of Sciences* **84**: 1674-1678

Lee RB, Smith JA, & Rickaby RE (2013) Cloning, Expression and Characterization of the delta-carbonic Anhydrase of Thalassiosira weissflogii (Bacillariophyceae). *J Phycol* **49**: 170-177

Lee RJ, Xiong G, Kofonow JM, Chen B, Lysenko A, Jiang P, Abraham V, Doghramji L, Adappa ND, Palmer JN, Kennedy DW, Beauchamp GK, Doulias PT, Ischiropoulos H, Kreindler JL, Reed DR, & Cohen NA (2012) T2R38 taste receptor polymorphisms underlie susceptibility to upper respiratory infection. *J Clin Invest* **122**: 4145-4159

Lehenkari P, Hentunen TA, Laitala-Leinonen T, Tuukkanen J, & Vaananen HK (1998) Carbonic anhydrase II plays a major role in osteoclast differentiation and bone

resorption by effecting the steady state intracellular pH and Ca2+. *Exp Cell Res* **242**: 128-137

Lehtonen J, Shen B, Vihinen M, Casini A, Scozzafava A, Supuran CT, Parkkila AK, Saarnio J, Kivela AJ, Waheed A, Sly WS, & Parkkila S (2004) Characterization of CA XIII, a novel member of the carbonic anhydrase isozyme family. *J Biol Chem* **279**: 2719-2727

Leinonen J, Kivela J, Parkkila S, Parkkila AK, & Rajaniemi H (1999) Salivary carbonic anhydrase isoenzyme VI is located in the human enamel pellicle. *Caries Res* **33**: 185-190

Leinonen JS, Saari KA, Seppanen JM, Myllyla HM, & Rajaniemi HJ (2004) Immunohistochemical demonstration of carbonic anhydrase isoenzyme VI (CA VI) expression in rat lower airways and lung. *J Histochem Cytochem* **52**: 1107-1112

Leinonen J, Parkkila S, Kaunisto K, Koivunen P, & Rajaniemi H (2001) Secretion of Carbonic Anhydrase Isoenzyme VI (CA VI) from Human and Rat Lingual Serous von Ebner's Glands. *Journal of Histochemistry & Cytochemistry* **49**: 657-662

Leppilampi M, Karttunen TJ, Kivela J, Gut MO, Pastorekova S, Pastorek J, & Parkkila S (2005) Gastric pit cell hyperplasia and glandular atrophy in carbonic anhydrase IX knockout mice: studies on two strains C57/BL6 and BALB/C. *Transgenic Res* **14**: 655-663

Leppilampi M, Parkkila S, Karttunen T, Gut MO, Gros G, & Sjoblom M (2005) Carbonic anhydrase isozyme-II-deficient mice lack the duodenal bicarbonate secretory response to prostaglandin E2. *Proc Natl Acad Sci U S A* **102**: 15247-15252

Lewis SE, Erickson RP, Barnett LB, Venta PJ, & Tashian RE (1988) N-ethyl-N-nitrosourea-induced null mutation at the mouse Car-2 locus: an animal model for human carbonic anhydrase II deficiency syndrome. *Proc Natl Acad Sci U S A* **85**: 1962-1966

Li SP & Goldman ND (1996) Regulation of human C-reactive protein gene expression by two synergistic IL-6 responsive elements. *Biochemistry* **35**: 9060-9068

Li X, Staszewski L, Xu H, Durick K, Zoller M, & Adler E (2002) Human receptors for sweet and umami taste. *Proc Natl Acad Sci U S A* **99**: 4692-4696

Li ZQ, Hu XP, Zhou JY, Xie XD, & Zhang JM (2015) Genetic polymorphisms in the carbonic anhydrase VI gene and dental caries susceptibility. *Genet Mol Res* **14**: 5986-5993

Lindskog S & Silverman D, N. (2000) The catalytic mechanism of mammalian carbonic anhydrases. In *The Carbonic Anhydrases: New Horizons,* W.R. Chegwidden, N.D. Carter and Y. H. Edwards (ed) pp 175. Birkhäuser Verlag, Basel

Liu M, Walter GA, Pathare NC, Forster RE, & Vandenborne K (2007) A quantitative study of bioenergetics in skeletal muscle lacking carbonic anhydrase III using 31P magnetic resonance spectroscopy. *Proc Natl Acad Sci U S A* **104**: 371-376

Lizio M, Harshbarger J, Abugessaisa I, Noguchi S, Kondo A, Severin J, Mungall C, Arenillas D, Mathelier A, Medvedeva YA, Lennartsson A, Drablos F, Ramilowski JA, Rackham O, Gough J, Andersson R, Sandelin A, Ienasescu H, Ono H, Bono H et al (2017) Update of the FANTOM web resource: high resolution transcriptome of diverse cell types in mammals. *Nucleic Acids Res* **45**: D737-D743

Lizio M, Harshbarger J, Shimoji H, Severin J, Kasukawa T, Sahin S, Abugessaisa I, Fukuda S, Hori F, Ishikawa-Kato S, Mungall CJ, Arner E, Baillie JK, Bertin N, Bono H, de Hoon M, Diehl AD, Dimont E, Freeman TC, Fujieda K et al (2015) Gateways to the FANTOM5 promoter level mammalian expression atlas. *Genome Biol* **16**: 22

Lo CM, Ma YS, Wei YH, Hsieh BYT, & Hsieh M (2018) Promoter analysis and transcriptional regulation of human carbonic anhydrase VIII gene in a MERRF disease cell model. *Arch Biochem Biophys* **641**: 50-61

Magnadottir B (2006) Innate immunity of fish (overview). *Fish Shellfish Immunol* **20**: 137-151

Magor BG & Magor KE (2001) Evolution of effectors and receptors of innate immunity. *Dev Comp Immunol* **25**: 651-682

Makani S, Chen HY, Esquenazi S, Shah GN, Waheed A, Sly WS, & Chesler M (2012) NMDA receptor-dependent afterdepolarizations are curtailed by carbonic anhydrase 14: regulation of a short-term postsynaptic potentiation. *J Neurosci* **32**: 16754-16762

Manoussakis MN & Moutsopoulos HM (2000) Sjogren's syndrome: autoimmune epithelitis. *Baillieres Best Pract Res Clin Rheumatol* **14**: 73-95

Maresca A, Temperini C, Vu H, Pham NB, Poulsen SA, Scozzafava A, Quinn RJ, & Supuran CT (2009) Non-zinc mediated inhibition of carbonic anhydrases: coumarins are a new class of suicide inhibitors. *J Am Chem Soc* **131**: 3057-3062

Marsh PD (2006) Dental plaque as a biofilm and a microbial community - implications for health and disease. *BMC Oral Health* **6 Suppl 1**: S14-6831-6-S1-S14

Matsunami H, Montmayeur JP, & Buck LB (2000) A family of candidate taste receptors in human and mouse. *Nature* **404**: 601-604

Matthews TA, Abel A, Demme C, Sherman T, Pan PW, Halterman MW, Parkkila S, & Nehrke K (2014) Expression of the CHOP-inducible carbonic anhydrase CAVI-b is required for BDNF-mediated protection from hypoxia. *Brain Res* **1543**: 28-37

Mboge MY, Mahon BP, McKenna R, & Frost SC (2018) Carbonic Anhydrases: Role in pH Control and Cancer. *Metabolites* **8**: 10.3390/metabo8010019

Medzhitov R (2007) Recognition of microorganisms and activation of the immune response. *Nature* **449**: 819-826

Meijering E, Dzyubachyk O, & Smal I (2012) Methods for cell and particle tracking. *Methods Enzymol* **504**: 183-200

Melis M, Atzori E, Cabras S, Zonza A, Calo C, Muroni P, Nieddu M, Padiglia A, Sogos V, Tepper BJ, & Tomassini Barbarossa I (2013) The Gustin (CA6) Gene Polymorphism, rs2274333 (A/G), as a Mechanistic Link between PROP Tasting and Fungiform Taste Papilla Density and Maintenance. *PLoS One* **8**: e74151

Meng EC, Pettersen EF, Couch GS, Huang CC, & Ferrin TE (2006) Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinformatics* **7**: 339

Miller E, Fiete D, Blake NM, Beranek M, Oates EL, Mi Y, Roseman DS, & Baenziger JU (2008) A necessary and sufficient determinant for protein-selective glycosylation in vivo. *J Biol Chem* **283**: 1985-1991

Mitani S (2017) Comprehensive functional genomics using Caenorhabditis elegans as a model organism. *Proc Jpn Acad Ser B Phys Biol Sci* **93**: 561-577

Montgomery JC, Venta PJ, Eddy RL, Fukushima YS, Shows TB, & Tashian RE (1991) Characterization of the human gene for a newly discovered carbonic anhydrase, CA VII, and its localization to chromosome 16. *Genomics* **11**: 835-848

Monti DM, De Simone G, Langella E, Supuran CT, Di Fiore A, & Monti SM (2017) Insights into the role of reactive sulfhydryl groups of Carbonic Anhydrase III and VII during oxidative damage. *J Enzyme Inhib Med Chem* **32**: 5-12

Morey JS, Ryan JC, & Van Dolah FM (2006) Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biol Proced Online* **8**: 175-193

Murakami H & Sly WS (1987) Purification and characterization of human salivary carbonic anhydrase. *J Biol Chem* **262**: 1382-1388

Murakami M, Kasuya T, Matsuba C, Ichihara N, Nishita T, Fujitani H, & Asari M (2003) Nucleotide sequence and expression of a cDNA encoding canine carbonic anhydrase VI (CA-VI). *DNA Seq* **14**: 195-198

NCBI Resource Coordinators (2016) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **44**: D7-19

Nelson G, Chandrashekar J, Hoon MA, Feng L, Zhao G, Ryba NJ, & Zuker CS (2002) An amino-acid taste receptor. *Nature* **416**: 199-202

Nelson G, Hoon MA, Chandrashekar J, Zhang Y, Ryba NJ, & Zuker CS (2001) Mammalian sweet taste receptors. *Cell* **106**: 381-390

Nicholas KB, Nicholas HB, & Deerfield DW (1997) GeneDoc: Analysis and Visualization of Genetic Variation. *EMBNEWNEWS* **4**: 14

Nienaber L, Cave-Freeman E, Cross M, Mason L, Bailey UM, Amani P, A Davis R, Taylor P, & Hofmann A (2015) Chemical probing suggests redox-regulation of the carbonic anhydrase activity of mycobacterial Rv1284. *FEBS J* **282**: 2708-2721

Nishita T, Itoh S, Arai S, Ichihara N, & Arishima K (2011) Measurement of carbonic anhydrase isozyme VI (CA-VI) in swine sera, colostrums, saliva, bile, seminal plasma and tissues. *Anim Sci J* **82**: 673-678

Nishita T, Tanaka Y, Wada Y, Murakami M, Kasuya T, Ichihara N, Matsui K, & Asari M (2007) Measurement of carbonic anhydrase isozyme VI (CA-VI) in bovine sera, saliva, milk and tissues. *Vet Res Commun* **31**: 83-92

Nishita T, Yatsu J, Murakami M, Kamoshida S, Orito K, Ichihara N, Arishima K, & Ochiai H (2014) Isolation and sequencing of swine carbonic anhydrase VI, an enzyme expressed in the swine kidney. *BMC Res Notes* **7**: 116-0500-7-116

Norderhaug IN, Johansen FE, Schjerven H, & Brandtzaeg P (1999) Regulation of the formation and external transport of secretory immunoglobulins. *Crit Rev Immunol* **19**: 481-508

Noursadeghi M, Bickerstaff MC, Gallimore JR, Herbert J, Cohen J, & Pepys MB (2000) Role of serum amyloid P component in bacterial infection: protection of the host or protection of the pathogen. *Proc Natl Acad Sci U S A* **97**: 14584-14589

Novoa B & Figueras A (2012) Zebrafish: model for the study of inflammation and the innate immune response to infectious diseases. *Adv Exp Med Biol* **946**: 253-275

Ochiai H, Kanemaki N, Kamoshida S, Murakami M, Ichihara N, Asari M, & Nishita T (2009) Determination of full-length cDNA nucleotide sequence of equine carbonic anhydrase VI and its expression in various tissues. *J Vet Med Sci* **71**: 1233-1237

Ogawa Y, Matsumoto K, Maeda T, Tamai R, Suzuki T, Sasano H, & Fernley RT (2002) Characterization of lacrimal gland carbonic anhydrase VI. *J Histochem Cytochem* **50**: 821-827

Ogilvie JM, Ohlemiller KK, Shah GN, Ulmasov B, Becker TA, Waheed A, Hennig AK, Lukasiewicz PD, & Sly WS (2007) Carbonic anhydrase XIV deficiency produces a functional defect in the retinal light response. *Proc Natl Acad Sci U S A* **104**: 8514-8519

Opavsky R, Pastorekova S, Zelnik V, Gibadulinova A, Stanbridge EJ, Zavada J, Kettmann R, & Pastorek J (1996) Human MN/CA9 gene, a novel member of the carbonic anhydrase family: structure and exon to protein domain relationships. *Genomics* **33**: 480-487

Padiglia A, Zonza A, Atzori E, Chillotti C, Calo C, Tepper BJ, & Barbarossa IT (2010) Sensitivity to 6-n-propylthiouracil is associated with gustin (carbonic anhydrase VI) gene polymorphism, salivary zinc, and body mass index in humans. *Am J Clin Nutr* **92**: 539-545

Pan PW, Kayra K, Leinonen J, Nissinen M, Parkkila S, & Rajaniemi H (2011) Gene expression profiling in the submandibular gland, stomach, and duodenum of CAVI-deficient mice. *Transgenic Res* **20**: 675-698

Pan PW, Parkkila AK, Autio S, Hilvo M, Sormunen R, Pastorekova S, Pastorek J, Haapasalo H, & Parkkila S (2012) Brain phenotype of carbonic anhydrase IX-deficient mice. *Transgenic Res* **21**: 163-176

Paradis E, Claude J, & Strimmer K (2004) APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* **20**: 289-290

Parkkila S, Kaunisto K, Kellokumpu S, & Rajaniemi H (1991) A high activity carbonic anhydrase isoenzyme (CA II) is present in mammalian spermatozoa. *Histochemistry* **95**: 477-482

Parkkila S, Kaunisto K, Rajaniemi L, Kumpulainen T, Jokinen K, & Rajaniemi H (1990) Immunohistochemical localization of carbonic anhydrase isoenzymes VI, II, and I in human parotid and submandibular glands. *J Histochem Cytochem* **38**: 941-947

Parkkila S & Parkkila AK (1996) Carbonic anhydrase in the alimentary tract. Roles of the different isozymes and salivary factors in the maintenance of optimal conditions in the gastrointestinal canal. *Scand J Gastroenterol* **31**: 305-317

Parkkila S, Parkkila AK, Juvonen T, & Rajaniemi H (1994) Distribution of the carbonic anhydrase isoenzymes I, II, and VI in the human alimentary tract. *Gut* **35**: 646-650

Parkkila S, Parkkila AK, & Rajaniemi H (1995) Circadian periodicity in salivary carbonic anhydrase VI concentration. *Acta Physiol Scand* **154**: 205-211

Parkkila S, Parkkila AK, Vierjoki T, Stahlberg T, & Rajaniemi H (1993) Competitive time-resolved immunofluorometric assay for quantifying carbonic anhydrase VI in saliva. *Clin Chem* **39**: 2154-2157

Pastorekova S, Parkkila S, Parkkila AK, Opavsky R, Zelnik V, Saarnio J, & 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**: 398-408

Penschow JD, Giles ME, Coghlan JP, & Fernley RT (1997) Redistribution of carbonic anhydrase VI expression from ducts to acini during development of ovine parotid and submandibular glands. *Histochem Cell Biol* **107**: 417-422

Peres RC, Camargo G, Mofatto LS, Cortellazzi KL, Santos MC, Nobre-dos-Santos M, Bergamaschi CC, & Line SR (2010) Association of polymorphisms in the carbonic

anhydrase 6 gene with salivary buffer capacity, dental plaque pH, and caries index in children aged 7-9 years. *Pharmacogenomics J* **10**: 114-119

Peri Y, Agmon-Levin N, Theodor E, & Shoenfeld Y (2012) Sjogren's syndrome, the old and the new. *Best Pract Res Clin Rheumatol* **26**: 105-117

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, & Ferrin TE (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**: 1605-1612

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29**: e45

Phalipon A, Cardona A, Kraehenbuhl JP, Edelman L, Sansonetti PJ, & Corthesy B (2002) Secretory component: a new role in secretory IgA-mediated immune exclusion in vivo. *Immunity* **17**: 107-115

Picco DCR, Lopes LM, Rocha Marques M, Line SRP, Parisotto TM, & Nobre Dos Santos M (2017) Children with a Higher Activity of Carbonic Anhydrase VI in Saliva Are More Likely to Develop Dental Caries. *Caries Res* **51**: 394-401

Pilka ES, Kochan G, Oppermann U, & Yue WW (2012) Crystal structure of the secretory isozyme of mammalian carbonic anhydrases CA VI: implications for biological assembly and inhibitor development. *Biochem Biophys Res Commun* **419**: 485-489

Postel R & Sonnenberg A (2012) Carbonic anhydrase 5 regulates acid-base homeostasis in zebrafish. *PLoS One* **7**: e39881

Raisanen SR, Lehenkari P, Tasanen M, Rahkila P, Harkonen PL, & Vaananen HK (1999) Carbonic anhydrase III protects cells from hydrogen peroxide-induced apoptosis. *FASEB J* **13**: 513-522

Ramji DP, Vitelli A, Tronche F, Cortese R, & Ciliberto G (1993) The two C/EBP isoforms, IL-6DBP/NF-IL6 and C/EBP delta/NF-IL6 beta, are induced by IL-6 to promote acute phase gene transcription via different mechanisms. *Nucleic Acids Res* **21**: 289-294

Reibring CG, El Shahawy M, Hallberg K, Kannius-Janson M, Nilsson J, Parkkila S, Sly WS, Waheed A, Linde A, & Gritli-Linde A (2014) Expression patterns and subcellular

localization of carbonic anhydrases are developmentally regulated during tooth formation. *PLoS One* **9**: e96007

Rice P, Longden I, & Bleasby A (2000) EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **16**: 276-277

Richardson JE & Bult CJ (2015) Visual annotation display (VLAD): a tool for finding functional themes in lists of genes. *Mamm Genome*

Roberts SB, Lane TW, & Morel FMM (1997) CARBONIC ANHYDRASE IN THE MARINE DIATOM THALASSIOSIRA WEISSFLOGII (BACILLARIOPHYCEAE)1. *J Phycol* **33**: 845-850

Rombout JH, Abelli L, Picchietti S, Scapigliati G, & Kiron V (2011) Teleost intestinal immunology. *Fish Shellfish Immunol* **31**: 616-626

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, & Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* **61**: 539-542

Ross MH & Pawlina W (2011) *Histology: a text and atlas: with correlated cell and molecular biology*. Lippincott Williams & Wilkins, a Wolters Kluwer business: Baltimore, MD 21201

Roy A, Kucukural A, & Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* **5**: 725-738

Rubio N, Sharp PM, Rits M, Zahedi K, & Whitehead AS (1993) Structure, expression, and evolution of guinea pig serum amyloid P component and C-reactive protein. *J Biochem* **113**: 277-284

Ruggiero MA, Gordon DP, Orrell TM, Bailly N, Bourgoin T, Brusca RC, Cavalier-Smith T, Guiry MD, & Kirk PM (2015) A higher level classification of all living organisms. *PLoS One* **10**: e0119248

Ruusuvuori E, Huebner AK, Kirilkin I, Yukin AY, Blaesse P, Helmy M, Kang HJ, El Muayed M, Hennings JC, Voipio J, Sestan N, Hubner CA, & Kaila K (2013) Neuronal carbonic anhydrase VII provides GABAergic excitatory drive to exacerbate febrile seizures. *EMBO J* **32**: 2275-2286

S. van der Walt, S. C. Colbert, & G. Varoquaux (2011) The NumPy Array: A Structure for Efficient Numerical Computation. *Computing in Science & Engineering* **13**: 22-30

Saari S, Hilvo M, Pan P, Gros G, Hanke N, Waheed A, Sly WS, & Parkkila S (2010) The most recently discovered carbonic anhydrase, CA XV, is expressed in the thick ascending limb of Henle and in the collecting ducts of mouse kidney. *PLoS One* **5**: e9624

Salinas I, Zhang YA, & Sunyer JO (2011) Mucosal immunoglobulins and B cells of Teleost fish. *Dev Comp Immunol* **35**: 1346-1365

Schneider CA, Rasband WS, & Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**: 671-675

Sengul F, Kilic M, Gurbuz T, & Tasdemir S (2016) Carbonic Anhydrase VI Gene Polymorphism rs2274327 Relationship Between Salivary Parameters and Dental-Oral Health Status in Children. *Biochem Genet* **54**: 467-475

Shah GN, Rubbelke TS, Hendin J, Nguyen H, Waheed A, Shoemaker JD, & Sly WS (2013) Targeted mutagenesis of mitochondrial carbonic anhydrases VA and VB implicates both enzymes in ammonia detoxification and glucose metabolism. *Proc Natl Acad Sci U S A* **110**: 7423-7428

Shah GN, Ulmasov B, Waheed A, Becker T, Makani S, Svichar N, Chesler M, & Sly WS (2005) Carbonic anhydrase IV and XIV knockout mice: roles of the respective carbonic anhydrases in buffering the extracellular space in brain. *Proc Natl Acad Sci U S A* **102**: 16771-16776

Sharma RK & Kalra J (1994) Molecular interaction between cAMP and calcium in calmodulin-dependent cyclic nucleotide phosphodiesterase system. *Clin Invest Med* **17**: 374-382

Shatzman AR & Henkin RI (1981) Gustin concentration changes relative to salivary zinc and taste in humans. *Proceedings of the National Academy of Sciences* **78**: 3867-3871

Shen L, Suresh L, Lindemann M, Xuan J, Kowal P, Malyavantham K, & Ambrus JL,Jr (2012) Novel autoantibodies in Sjogren's syndrome. *Clin Immunol* **145**: 251-255

Shi C, Uda Y, Dedic C, Azab E, Sun N, Hussein AI, Petty CA, Fulzele K, Mitterberger-Vogt MC, Zwerschke W, Pereira R, Wang K, & Pajevic PD (2018) Carbonic anhydrase III protects osteocytes from oxidative stress. *FASEB J* **32**: 440-452

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, & Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**: 539

Sly WS, Hewett-Emmett D, Whyte MP, Yu YS, & Tashian RE (1983) Carbonic anhydrase II deficiency identified as the primary defect in the autosomal recessive syndrome of osteopetrosis with renal tubular acidosis and cerebral calcification. *Proc Natl Acad Sci U S A* **80**: 2752-2756

Smith LM, Sanders JZ, Kaiser RJ, Hughes P, Dodd C, Connell CR, Heiner C, Kent SB, & Hood LE (1986) Fluorescence detection in automated DNA sequence analysis. *Nature* **321**: 674-679

Sok J, Wang XZ, Batchvarova N, Kuroda M, Harding H, & Ron D (1999) CHOP-Dependent stress-inducible expression of a novel form of carbonic anhydrase VI. *Mol Cell Biol* **19**: 495-504

Spicer SS, Lewis SE, Tashian RE, & Schulte BA (1989) Mice carrying a CAR-2 null allele lack carbonic anhydrase II immunohistochemically and show vascular calcification. *Am J Pathol* **134**: 947-954

Stockhammer OW, Zakrzewska A, Hegedus Z, Spaink HP, & Meijer AH (2009) Transcriptome profiling and functional analyses of the zebrafish embryonic innate immune response to Salmonella infection. *J Immunol* **182**: 5641-5653

Sugiura Y, Ichihara N, Nishita T, Murakami M, Amasaki H, & Asari M (2008) Immunohistolocalization and gene expression of secretory carbonic anhydrase isoenzyme CA-VI in canine nasal cavity. *J Vet Med Sci* **70**: 1037-1041

Sunyer JO, Zarkadis IK, & Lambris JD (1998) Complement diversity: a mechanism for generating immune diversity? *Immunol Today* **19**: 519-523

Supuran CT (2010) Carbonic anhydrase inhibitors. *Bioorg Med Chem Lett* **20**: 3467-3474

Supuran CT (2018) Carbonic Anhydrases and Metabolism. *Metabolites* **8**: 10.3390/metabo8020025

Supuran CT (2008a) Carbonic anhydrases: novel therapeutic applications for inhibitors and activators. *Nat Rev Drug Discov* **7**: 168-181

Supuran CT (2008b) Diuretics: from classical carbonic anhydrase inhibitors to novel applications of the sulfonamides. *Curr Pharm Des* **14**: 641-648

Supuran CT & De Simone G (2015) Chapter 1 - Carbonic Anhydrases: An Overview. In *Carbonic Anhydrases as Biocatalysts,* Simone CTSD (ed) pp 3-13. Elsevier: Amsterdam

Sutherland GR, Baker E, Fernandez KE, Callen DF, Aldred P, Coghlan JP, Wright RD, & Fernley RT (1989) The gene for human carbonic anhydrase VI(CA6) is on the tip of the short arm of chromosome 1. *Cytogenet Cell Genet* **50**: 149-150

Suyama M, Torrents D, & Bork P (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res* **34**: W609-12

Tashian R, Carter ND, Bergenhem N, & (ed) (2000) Carbonic anhydrase (CA)-related proteins (CA-RPs), and trans-membrane proteins with CA or CA RP membrane proteins with CA or CA RP domains.

Temperini C, Innocenti A, Guerri A, Scozzafava A, Rusconi S, & Supuran CT (2007) Phosph(on)ate as a zinc-binding group in metalloenzyme inhibitors: X-ray crystal structure of the antiviral drug foscarnet complexed to human carbonic anhydrase I. *Bioorg Med Chem Lett* **17**: 2210-2215

Thatcher BJ, Doherty AE, Orvisky E, Martin BM, & Henkin RI (1998) Gustin from human parotid saliva is carbonic anhydrase VI. *Biochem Biophys Res Commun* **250**: 635-641

The UniProt Consortium (2010) The Universal Protein Resource (UniProt) in 2010. *Nucleic Acids Research* **38**: D142-D148

Thompson D, Pepys MB, & Wood SP (1999) The physiological structure of human C-reactive protein and its complex with phosphocholine. *Structure* **7**: 169-177

Tolvanen ME, Ortutay C, Barker HR, Aspatwar A, Patrikainen M, & Parkkila S (2013) Analysis of evolution of carbonic anhydrases IV and XV reveals a rich history of gene duplications and a new group of isozymes. *Bioorg Med Chem* **21**: 1503-1510

Toniatti C, Demartis A, Monaci P, Nicosia A, & Ciliberto G (1990) Synergistic trans-activation of the human C-reactive protein promoter by transcription factor HNF-1 binding at two distinct sites. *EMBO J* **9**: 4467-4475

Torres BV, McCrumb DK, & Smith DF (1988) Glycolipid-lectin interactions: reactivity of lectins from Helix pomatia, Wisteria floribunda, and Dolichos biflorus with glycolipids containing N-acetylgalactosamine. *Arch Biochem Biophys* **262**: 1-11

Trapnell C, Pachter L, & Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**: 1105-1111

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, & Pachter L (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**: 511-515

Tureci O, Sahin U, Vollmar E, Siemer S, Gottert E, Seitz G, Parkkila AK, Shah GN, Grubb JH, Pfreundschuh M, & Sly WS (1998) Human carbonic anhydrase XII: cDNA cloning, expression, and chromosomal localization of a carbonic anhydrase gene that is overexpressed in some renal cell cancers. *Proc Natl Acad Sci U S A* **95**: 7608-7613

Vaananen HK & Parvinen EK (1983) High active isoenzyme of carbonic anhydrase in rat calvaria osteoclasts. Immunohistochemical study. *Histochemistry* **78**: 481-485

Vasta GR, Nita-Lazar M, Giomarelli B, Ahmed H, Du S, Cammarata M, Parrinello N, Bianchet MA, & Amzel LM (2011) Structural and functional diversity of the lectin repertoire in teleost fish: relevance to innate and adaptive immunity. *Dev Comp Immunol* **35**: 1388-1399

Vullo D, Del Prete S, Di Fonzo P, Carginale V, Donald WA, Supuran CT, & Capasso C (2017) Comparison of the Sulfonamide Inhibition Profiles of the beta- and gamma-Carbonic Anhydrases from the Pathogenic Bacterium Burkholderia pseudomallei. *Molecules* **22**: 10.3390/molecules22030421

Waheed A & Sly WS (2017) Carbonic anhydrase XII functions in health and disease. *Gene* **623**: 33-40

Warren WC, Hillier LW, Marshall Graves JA, Birney E, Ponting CP, Grutzner F, Belov K, Miller W, Clarke L, Chinwalla AT, Yang SP, Heger A, Locke DP, Miethke P, Waters PD, Veyrunes F, Fulton L, Fulton B, Graves T, Wallis J et al (2008) Genome analysis of the platypus reveals unique signatures of evolution. *Nature* **453**: 175-183

Webb B & Sali A (2002) Comparative Protein Structure Modeling Using MODELLER. In *Current Protocols in Bioinformatics,* Anonymous John Wiley & Sons, Inc.

Westerfield M (2007) *The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (Danio rerio).* University of Oregon Press: Eugene

Wienholds E, van Eeden F, Kosters M, Mudde J, Plasterk RH, & Cuppen E (2003) Efficient target-selected mutagenesis in zebrafish. *Genome Res* **13**: 2700-2707

Winata CL, Korzh S, Kondrychyn I, Zheng W, Korzh V, & Gong Z (2009) Development of zebrafish swimbladder: The requirement of Hedgehog signaling in specification and organization of the three tissue layers. *Dev Biol* **331**: 222-236

Winum JY, Montero JL, Vullo D, & Supuran CT (2009) Carbonic anhydrase inhibitors: glycosylsulfanilamides act as subnanomolar inhibitors of the human secreted isoform VI. *Chem Biol Drug Des* **74**: 636-639

Xu J, Xu X, Wang B, Ma Y, Zhang L, Xu H, Hu Y, Wu J, & Cao X (2017) Nuclear carbonic anhydrase 6B associates with PRMT5 to epigenetically promote IL-12 expression in innate response. *Proc Natl Acad Sci U S A* **114**: 8620-8625

Xu Y, Feng L, Jeffrey PD, Shi Y, & Morel FM (2008) Structure and metal exchange in the cadmium carbonic anhydrase of marine diatoms. *Nature* **452**: 56-61

Yamaguchi M & Yoshida H (2018) Drosophila as a Model Organism. *Adv Exp Med Biol* **1076**: 1-10

Yoder JA & Litman GW (2011) The phylogenetic origins of natural killer receptors and recognition: relationships, possibilities, and realities. *Immunogenetics* **63**: 123-141

Zaccone G, Kapoor BG, Fasulo S, & Ainis L (2001) Structural, histochemical and functional aspects of the epidermis of fishes. *Advances in Marine Biology* **40**: 253-348

Zahedi K (1997) Characterization of the binding of serum amyloid P to laminin. *J Biol Chem* **272**: 2143-2148

Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, Billis K, Cummins C, Gall A, Giron CG, Gil L, Gordon L, Haggerty L, Haskell E, Hourlier T, Izuogu OG, Janacek SH, Juettemann T, To JK, Laird MR et al (2018) Ensembl 2018. *Nucleic Acids Res* **46**: D754-D761

Zhao GQ, Zhang Y, Hoon MA, Chandrashekar J, Erlenbach I, Ryba NJ, & Zuker CS (2003) The receptors for mammalian sweet and umami taste. *Cell* **115**: 255-266

Zolfaghari Emameh R, Barker H, Tolvanen ME, Ortutay C, & Parkkila S (2014) Bioinformatic analysis of beta carbonic anhydrase sequences from protozoans and metazoans. *Parasit Vectors* **7**: 38-3305-7-38

9 ORIGINAL COMMUNICATIONS

PUBLICATION I

The role of carbonic anhydrase VI in bitter taste perception: evidence from the *Car6*$^{-/-}$ mouse model

Maarit Patrikainen, Peiwen Pan, Natalia Kulesskaya, Vootele Voikar, Seppo Parkkila

Journal of Biomedical Science, 2014 Aug 19;21:82
https://doi.org/10.1186/s12929-014-0082-2

Patrikainen et al. Journal of Biomedical Science 2014, 21:82
http://www.jbiomedsci.com/content/21/1/82

科 技 部 Ministry of Science and Technology
The cost of publication in Journal of Biomedical Science is borne by the Ministry of Science and Technology, Taiwan.

JBS JOURNAL OF BIOMEDICAL SCIENCE

RESEARCH Open Access

The role of carbonic anhydrase VI in bitter taste perception: evidence from the *Car6*$^{-/-}$ mouse model

Maarit Patrikainen[1], Peiwen Pan[1], Natalia Kulesskaya[2], Vootele Voikar[2] and Seppo Parkkila[1,3*]

Abstract

Background: Carbonic anhydrase VI (CA VI) is a secretory isozyme of the α-CA gene family. It is highly expressed in the salivary and mammary glands and secreted into saliva and milk. Although CA VI was first described as a gustatory protein, its exact functional roles have remained enigmatic. Interestingly, polymorphism of the *CA6* gene was recently linked to bitter taste perception in humans. In this study, we compared the preference of *Car6*$^{-/-}$ and wild-type mice for different taste modalities in an IntelliCage monitoring environment. Morphologies of taste buds, tongue papillae, and von Ebner's glands were evaluated by light microscopy. Cell proliferation and rate of apoptosis in tongue specimens were examined by Ki67 immunostaining and fluorescent DNA fragmentation staining, respectively.

Results: The behavioral follow up of the mice in an IntelliCage system revealed that *Car6*$^{-/-}$ mice preferred 3 μM quinine (bitter) solution, whereas wild type mice preferred water. When the quinine concentration increased, both groups preferentially selected water. Histological analysis, Ki67 immunostaining and detection of apoptosis did not reveal any significant changes between tongue specimens of the knockout and wild type mice.

Conclusions: Our knockout mouse model confirms that CA VI is involved in bitter taste perception. CA VI may be one of the factors which contribute to avoidance of bitter, potentially harmful, substances.

Keywords: Bitter, Carbonic anhydrase, Gustin, Mouse, Saliva, Taste

Background

Carbonic anhydrase VI (CA VI) is the only secretory isozyme in the α-CA enzyme family. In the first report on CA VI Henkin's group identified a novel protein, gustin, from human saliva [1], which was later shown to be identical to CA VI [2]. Independently, this same protein was described as a distinct CA enzyme by Fernley's team, who identified a novel high molecular weight form of CA in the sheep parotid gland and saliva [3]. Later, CA VI was also isolated from rat [4] and human saliva [5]. Immunohistochemical studies have indicated that CA VI is highly expressed in the serous acinar cells of the parotid and submandibular glands [6,7]. In fact, it is one of the major protein constituents of human saliva. The mean concentration of CA VI in paraffin-stimulated saliva is 6.8 +/− 4.3 mg/L and the secretion rate is 10.2 +/− 7.9 μg/min [8]. Secretion of CA VI is tightly regulated by circadian periodicity; the levels are low during the night and increase rapidly after awakening [9]. The expression of CA VI is not restricted only to the salivary glands. Ogawa's group demonstrated CA VI expression in the lacrimal gland [10], and Karhumaa and coworkers found that milk contains high levels of secretory CA VI [11]. The 42-kDa polypeptide purified from human milk by CA inhibitor affinity chromatography shared 100% homology with salivary CA VI according to a protein sequence analysis. A time-resolved immunofluorometric assay showed that colostrum contained an eight times higher concentration of CA VI than mature milk, the latter containing concentrations comparable to the mean levels in saliva.

Even though CA VI was originally discovered more than 30 years ago, its physiological role has remained unclear. As an enzymatically active CA it could maintain optimal pH homeostasis within the oral cavity and upper

* Correspondence: seppo.parkkila@uta.fi
[1]School of Medicine and BioMediTech, University of Tampere, Tampere FI-33014, Finland
[3]Fimlab Ltd and Tampere University Hospital, Tampere FI-33520, Finland
Full list of author information is available at the end of the article

alimentary tract [12,13]. The high concentrations in milk and colostrum suggested that it could participate in the developmental processes of the gastrointestinal canal during the postnatal period [11]. Earlier studies also suggested that CA VI present in saliva could play a protective role against cariogenesis [14]. These and other possible physiological roles can now be studied using the recently described knockout mouse model for CA VI deficiency [15]. The knockout mice are viable, fertile, and have shown a normal life span. Surprisingly, an *in vivo* cariogenesis model revealed that the CA VI knockout mice had a lower rate of cariogenesis and oral colonization of *Streptococcus mutans* than the wild type (WT) controls [16]. Histological analyses have indicated a greater number of lymphoid follicles in the small intestinal Peyer's patches of the knockout mice, suggesting an immunological phenotype for CA VI deficiency [15]. This was further supported by functional clustering of differentially expressed genes, which revealed a number of altered biological processes in the knockout mice. Importantly, these included a Gene Ontology (GO) term for a biological process called "immune system process" in the duodenum.

Henkin's group linked gustin (CA VI) to the regulation of taste function in 1981 [17]. They found that the biochemical characteristics of CA VI were similar in protein isolated from both subjects with normal taste acuity and from patients with hypogeusia. Interestingly, they reported that hypogeusic subjects had salivary CA VI concentration as low as 20% that of normal subjects, but they did not link CA VI to any specific taste modality. Recently, Barbarossa's group has shown a link between bitter taste modality and CA VI by finding polymorphism in the *CA6* gene (rs2274333 (A/G)) which contributes to 6-*n*-propylthiouracil taster status [18]. Later, they elegantly showed that alterations of bitter taste function are due to polymorphic changes in both bitter receptor gene (*TAS2R38*) and *CA6* gene, while also requiring contributions from other still unknown factors [19]. Most recently, Barbarossa's group demonstrated that rs2274333 polymorphic change in the *CA6* gene affects 6-*n*-propylthiouracil sensitivity by acting on fungiform papilla development and maintenance [20]. There is still another important piece of evidence that CAs indeed contribute to taste perception. Chandrashekar and coworkers demonstrated by targeted genetic ablation and neurophysiological measurements, that sour-sensing cells act as taste sensors for carbonation, and also showed that membrane-associated CA IV functions as the principal CO_2 taste sensor [21].

In this study, we investigated the role of CA VI in taste function by utilizing the CA VI-deficient mouse model. These mice were placed in an IntelliCage system [22] for automated behavioral screening and their preferences for various taste modalities were analyzed. The results provide a mouse model confirmation that CA VI deficiency leads to abnormal bitter taste perception.

Methods

Ethical approval

The production of the knockout mouse line was approved by the Animal Experimentation Committee of the University of Oulu, Finland. Behavioural experiments were carried out in accordance with the Guidelines laid down with the European Communities Council Directive of 24 November 1986 (86/609/EEC) and were approved by the County Administrative Board of Southern Finland (license number ESAVI-2010-09011/Ym-23).

Car6$^{-/-}$ mice

Generation and phenotypic characterization of *Car6*$^{-/-}$ mice have been described previously by Pan et al. [15]. The mice with the targeted allele were backcrossed more than 10 generations to obtain mice with a pure C57BL/6 strain background. *Car6*$^{-/-}$ mice proved to be fertile and showed normal life-span. The absence of CA VI protein in saliva and salivary glands was confirmed by immunohistochemistry and western blotting.

Behavioral monitoring

Animals

Nine WT female and nine knock-out female mice at the age of 12 weeks, at the beginning of experiment, were used for behavioral analysis. One week before onset of testing in an IntelliCage, RFID transponders (Planet ID GmbH, Essen, Germany) were injected subcutaneously in the dorso-cervical region under isoflurane inhalation anesthesia. Throughout the experiment the mice were maintained under 12/12 h light cycle (lights on at 06:00) at controlled temperature (21 ± 1°C) and humidity (50–60%).

Apparatus and procedure

The IntelliCage apparatus (NewBehavior AG, Zurich, Switzerland, www.newbehavior.com) is placed in a polycarbonate cage (20.5 cm high, 58 × 40 cm top, 55 × 37.5 cm bottom, Tecniplast, 2000P, Buguggiate, Italy) and accommodates up to 16 mice. Its aluminum top contains a freely accessible food rack filled with standard mouse chow (Teklad 2016, Harlan). The floor is covered with bedding (aspen chips 5 × 5 × 1 mm, Tapvei Oy, Finland) and provides four central red shelters (Tecniplast, Buguggiate, Italy). Four triangular conditioning chambers (15 × 15 × 21 cm) are fitted in the cage corners and each provide room for one mouse at a time. Each chamber contains two drinking bottles, accessible via two round openings (13 mm diameter) with motorized doors. Three multicolor LEDs are mounted above each door and the chamber ceiling contains a motorized valve for delivery of air puffs. Mice that access a chamber are identified by a circular RIFD antenna

at its entrance (30 mm inner diameter) and the duration of their visit is determined by both the antenna reading and a temperature sensor that detects the presence of the animal inside the corner. During a visit, number and duration of individual nosepokes at each door are recorded using IR-beam sensors. Licking episodes at each bottle are monitored using lickometers (duration of the episode, number of licks, total contact time). IntelliCages have individual controllers and are connected to a central PC running the IntelliCage Plus software that permits designing and running experiments, as well as analysis of the recorded data (IntelliCage Plus, NewBehavior AG).

The animals were habituated to the new environment for 8 days. During this period, every visit to the chamber opened both doors for 7 sec allowing access to drinking bottles filled with tap water. For testing spontaneous taste preference, one corner contained two bottles with tap water. The other three corners contained bottles with different concentrations of taste solution (one concentration per corner). Each taste modality was tested over 4 days, and the position of bottles was changed every day in order to avoid development of place preference or avoidance. Four taste modalities were tested in the following order and concentrations: 1) Sweet – saccharin 0.1, 5, 50 g/l; 2) Salty – NaCl 50, 150, 450 mM; 3) Sour – citric acid 0.1, 10, 100 mM; 4) Bitter – quinine 0.003, 0.03, 0.3 mM. Two-day wash-out periods with tap water in all corners and bottles were applied before subsequent tastants. The number of licks in each corner was recorded and used for calculating the preference expressed as a percentage of licks at different concentrations. The data were compared between genotypes by repeated measures ANOVA with preference to tastant as a within-subject factor, followed by Newman-Keuls post hoc test.

Morphological analyses and immunohistochemistry

Whole tongue samples were taken from the wild-type (n = 9) and *Car6*$^{-/-}$ (n = 9) adult mice. The samples were dissected, fixed in 4% formaldehyde overnight, embedded in paraffin, and sectioned to a thickness of 5 μm. For morphological analysis paraffin was removed with xylene, rehydrated by a descending series of ethanol, and the sections were stained with hematoxylin and eosin. Morphologies of different papillae (circumvallate, fungiform, and filiform) and von Ebner's glands were evaluated by light microscopy. The numbers of fungiform papillae were evaluated in 27 sections/group (three sections/mouse). The anterior half of the tongue was examined for the presence of fungiform papillae. The number of fungiform papillae from each section were counted. Because of the small sample size, and potential non-normal distribution of section cell counts, the Mann–Whitney test was used for comparing the two groups. P value < 0.05 was considered statistically significant.

Immunohistochemistry for the proliferation of marker Ki67 was performed using a Vectastain Elite ABC kit (Vector Laboratories, Burlingame, CA, USA) following the manufacturer's instructions. Prior to staining, the sections were boiled in 0.01 M sodium citrate buffer, pH 6.0, for 20 min. Endogenous peroxidase activity was blocked by treatment with 3% H_2O_2 solution. Nonspecific binding of the primary antibody was prevented using 10% normal rabbit serum (NRS) as a blocking agent. The sections were incubated overnight at + 4°C in a primary antibody solution containing monoclonal rat anti-mouse Ki67 (Dako Denmark A/S, Glostrup, Denmark) (diluted 1:100), 1% NRS, and 0.1% Tween-20 (Sigma-Aldrich, St. Louis, MO, USA) in phosphate buffered saline (PBS). Biotinylated rabbit anti-rat serum (Vector Laboratories) (diluted 1:1000) was used as a secondary antibody, and the sections were incubated 30 min at room temperature. DAB (3, 3'-diaminobenzidine, Invitrogen, Camarillo, CA, USA) was used for precipitating the substrate, and the sections were finally counterstained with Mayer's hematoxylin. Positively stained cells were counted from one field/tongue photographed with 100× magnification (Nikon Microphot-FXA, Nikon Instruments Europe B.V., Amsterdam, Netherlands). To determine the difference between the two groups, 21 wild type and 16 knockout mice fields were analyzed. Statistical analysis was performed using Mann–Whitney test.

For studying the apoptotic level of tongue epithelial cells we used FragEL DNA Fragmentation Detection Kit, Fluorescent-TdT Enzyme (Calbiochem, EMD Chemicals, Inc., San Diego, CA, USA) according to manufacturer's instructions. After deparaffinisation and rehydration the sections were incubated with proteinase K (20 μg/ml) at room temperature for 20 min. Prior to the equilibration step positive control sections were treated with 1 μg/μl DNase I in 1 mM $MgSO_4$/1 × tris-buffered saline (TBS) and incubated for 20 min at room temperature. After 15 min incubation in 1 × TdT Equilibration buffer, the sections were covered with TdT Labeling Reaction Mixture. Negative control was prepared at labeling step using dH_2O instead of enzyme in TdT Labeling Reaction mixture. Samples were incubated at + 37°C for 1.5 h. Sections were mounted with Fluorescein-FragEL mounting medium and photographed using Nikon Mikrophot-FXA microscope with 450–490 nm filter using 100 × magnification (Nikon Instruments Europe BV). Labeled nuclei were analyzed from 7 *Car6*$^{-/-}$ and 7 WT mice tongues (two fields/tongue). Statistical comparison between the two groups was performed using Mann–Whitney test.

Results

Behavioral study

There was no difference in body weight between the WT and *Car6*$^{-/-}$ mice. Moreover, the activity (number

of corner visits, Figure 1A) and total liquid consumption (number of licks, Figure 1B) were similar between the groups in all phases of the experiment.

The mice showed similar preference for saccharin 5 g/l, whereas 50 g/l was avoided by both groups (Effect of concentration $F(3,48) = 126.9$, $p < 0.0001$; no significant

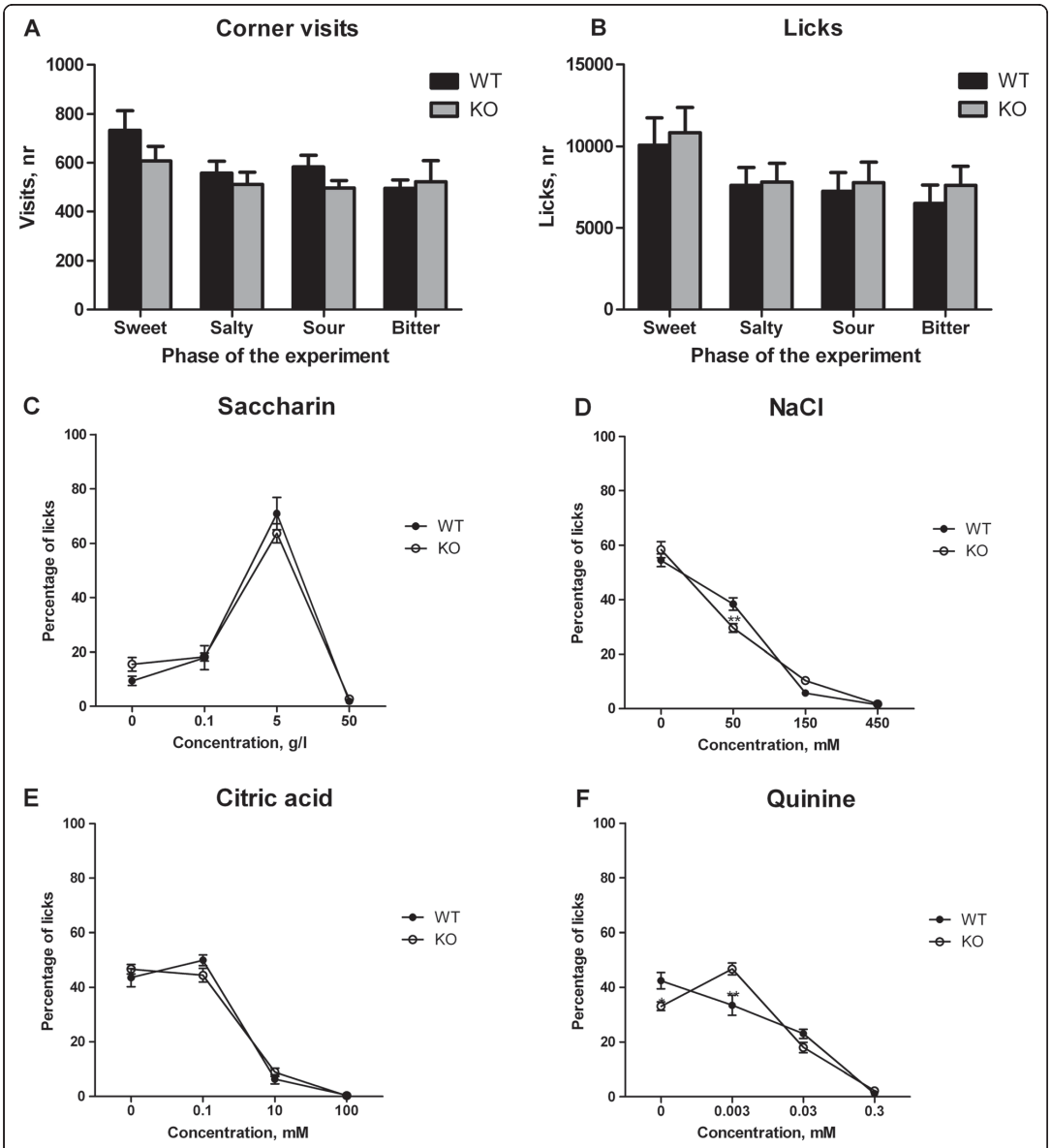

Figure 1 Assessment of taste preference in the IntelliCage. A. Total number of corner visits during 4 days of each phase of the experiment. **B**. Total number of licks during 4 days of each phase of the experiment. **C**. Percentage of licks of sweet (saccharin) solution at different concentrations. **D**. Percentage of licks of salty (NaCl) solution at different concentrations. **E**. Percentage of licks of sour (citric acid) solution at different concentrations. **F**. Percentage of licks of bitter (quinine) solution at different concentrations. * - $p < 0.05$, ** - $p < 0.01$ between the genotypes; # - $p < 0.01$ compared to water.

interaction between genotype and concentration F(3,48) = 1.1, p = 0.34) (Figure 1C). Both groups displayed a similar preference for water and avoidance of 150 and 450 mM of NaCl. However, Car6$^{-/-}$ mice showed slightly lower percentage of licks at 50 mM NaCl as compared with WT mice (effect of concentration F(3,48) = 300.2, p < 0.0001; significant interaction between genotype and concentration F(3,48) = 4.5, p = 0.0075) (Figure 1D). Both groups showed equal preference for water and 0.1 mM of citric acid whereas 10 mM and 100 mM concentrations of citric acid were avoided (effect of concentration F(3,48) = 246.6, p < 0.0001; no significant interaction between genotype and concentration F(3,48) = 1.6, p = 0.21) (Figure 1E). When testing bitter taste, *Car6*$^{-/-}$ mice showed preference for 0.003 mM quinine solution, whereas WT mice preferred water (Figure 1F). Higher concentrations of quinine were clearly avoided by both WT and knockout mice (effect of concentration F(3,48) = 103.4, p < 0.0001; significant interaction between genotype and concentration F(3,48) = 7.8, p = 0.0002).

Tongue morphology, cell proliferation and apoptosis

Tongue histologies were analyzed from *Car6*$^{-/-}$ and WT mice, including the morphology and number of taste buds and different types of papillae. Figure 2 shows representative images of morphology. The morphologies of three different papillae (circumvallate, fungiform and filiform) were not significantly different between the two groups of mice, nor did we find any changes in the von Ebner's glands (data not shown). The number of fungiform papillae showed no change in *Car6*$^{-/-}$ mice as compared to the control mice (p = 0.58, Mann–Whitney test).

Immunohistochemistry for the proliferation marker Ki67 showed positive for nuclei in the tongue epithelium of all specimens, as expected. Mean counts for cell proliferation were 105.3/section (range 35–159) and 100.2/section (range 8–168) in the *Car6*$^{-/-}$ and WT mice, respectively. No statistically significant difference was observed between the *Car6*$^{-/-}$ and WT mice (p = 0.63, Mann–Whitney test) (Figure 3). Rate of apoptosis in the tongue specimens was analyzed using a DNA fragmentation detection kit. Mean values for the apoptosis count were 2.9/section (range 0–9) and 2.0/section (range 0–5) in the *Car6*$^{-/-}$ and WT mice, respectively. No statistically significant difference was found in the rate of apoptosis between the *Car6*$^{-/-}$ and WT mice (p = 0.69, Mann–Whitney test).

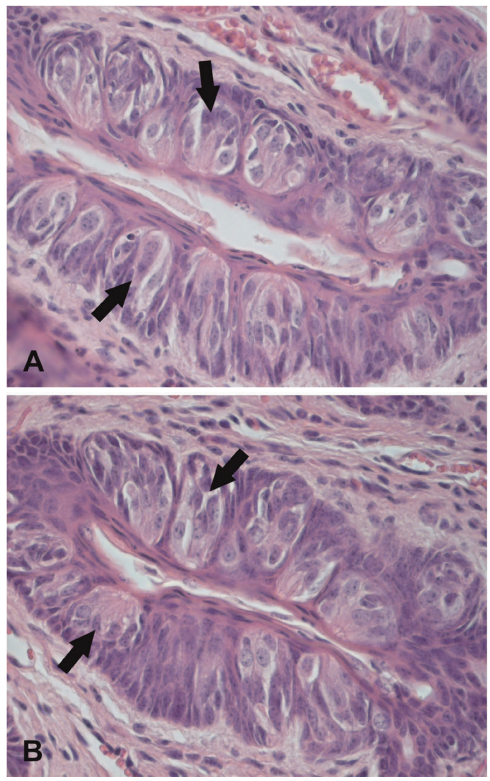

Figure 2 Histological images of tongue morphology in *Car6*$^{-/-}$ (A) and WT (B) mice. Arrows point to some examples of taste buds. No significant differences are visible in morphology. Original magnifications x 400.

Discussion

The present data shows that CA VI deficiency results in an altered behavior in mice for preferred taste. The most significant change was observed in bitter taste, which is well in line with the recent observations from Barbarossa's group, showing an altered bitter taste perception in subjects with *CA6* gene polymorphism [18-20]. The previous study indicated that both *TAS2R38* and *CA6* loci are important for discriminating low concentrations of the bitter tasting chemical, 6-n-propylthiouracil [19]. In our study, the CA VI deficient mice showed a lower percentage of licks at 50 mM NaCl than the WT mice, suggesting that CA VI is somehow involved in perception of salty taste. This finding has not been reported earlier and warrants further investigations. There were no other significant behavioral changes in terms of the taste preference.

It is commonly known that oral and perioral infections can cause taste dysfunction [23]. Henkin's group demonstrated that in human patients suffering from taste dysfunction and influenza-like symptoms, the taste buds in circumvallate papillae exhibited apoptotic features, such as severe vacuolization and cellular degeneration [24]. These patients also had lowered concentrations of CA VI in their saliva. Treatment with zinc normalized the CA VI concentrations and the senses of taste and smell in some cases. Interestingly, the taste bud morphology was normalized in

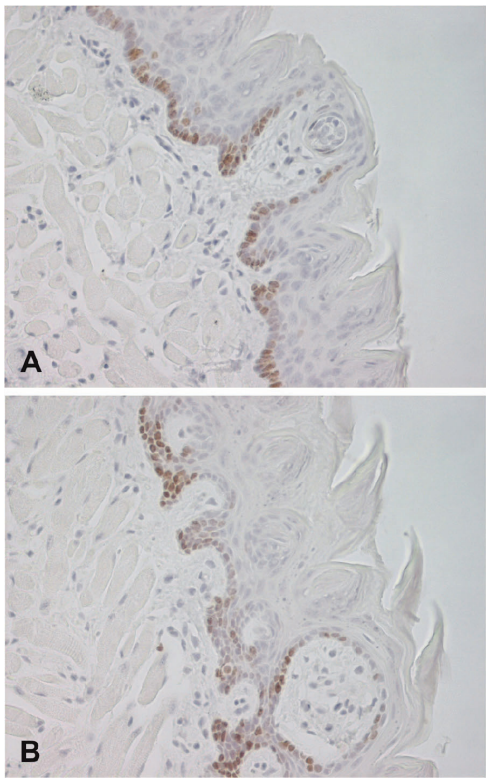

Figure 3 Immunohistochemical staining of Ki67 proliferation marker in tongue specimens from *Car6*$^{-/-}$ (A) and WT (B) mice. Both samples demonstrate a number of positive nuclei mainly located in the basal layer of the stratified epithelium. Original magnifications × 400.

these patients, suggesting that CA VI may function as a trophic factor for the taste bud stem cells. Our analysis by light microscopy revealed no apparent changes in taste bud morphology in CA VI deficient mice nor did we observe any difference in the rate of apoptosis. This might suggest that the apoptotic changes observed earlier [24] are primarily due to the viral infection itself [25] rather than a result of CA VI deficiency. Similarly, the normalized taste bud morphology after the zinc treatment can be attributed to the effects of zinc on multiple cellular processes, other than those regulated by CA VI.

The present morphological analyses showed no significant change in the number of fungiform papillae in *Car6*$^{-/-}$ mice compared to controls. This finding potentially contradicts the previous observations which showed an abnormal development of fungiform papillae in humans with *CA6* gene polymorphism [20]. The previous study also showed that CA VI protein induces proliferation of tongue epithelial cells. Even though our results were not able to confirm the role of CA VI in morphogenesis of fungiform papillae, nor did they show any effect in proliferation, the recent findings from Barbarossa's group are attractive and warrant confirmation in different cell lines. Importantly, our *in vivo* studies were performed using a mouse model in which only a single salivary factor, CA VI, had been knocked out. There are a number of other growth factors present in saliva [26], and deficiency of CA VI alone may not be sufficient to affect significantly the cell proliferation rate. In spite of these negative results regarding the role of CA VI in cell proliferation, it is still attractive to speculate whether addition of exogenous CA VI into various products, such as mouth rinse, tooth paste, or chewing gum, could have some beneficial effects on oral health.

Conclusions

Our results on *Car6*$^{-/-}$ mice confirm that CA VI deficiency leads to abnormal bitter taste perception. This finding agrees well with the recent observations on the role of CA VI in human taste function. CA VI may be one of the factors, which can contribute to avoidance of bitter substances in mammals.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors were involved in design of the study. PP generated and characterized the *Car6*$^{-/-}$ mouse colony. MP carried out the histochemical and immunohistochemical analyses. NK and VV performed the taste function tests by IntelliCage apparatus and analyzed the data. SP and VV drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This work was supported by the competitive Research Funding of the Tampere University Hospital (Grant 9 N054) and grants from the Academy of Finland, Biocenter Finland, Jane and Aatos Erkko Foundation, Sigrid Juselius Foundation, and Tampere Tuberculosis Foundation. We thank Marianne Kuuslahti and Aulikki Lehmus for their skillful technical assistance and Harlan Barker for revising the manuscript.

Author details

[1]School of Medicine and BioMediTech, University of Tampere, Tampere FI-33014, Finland. [2]Neuroscience Center, University of Helsinki, Helsinki FI-00014, Finland. [3]Fimlab Ltd and Tampere University Hospital, Tampere FI-33520, Finland.

Received: 24 March 2014 Accepted: 13 August 2014
Published: 19 August 2014

References

1. Henkin RI, Lippoldt RE, Bilstad J, Edelhoch H: **A zinc protein isolated from human parotid saliva.** *Proc Natl Acad Sci U S A* 1975, **72:**488–492.
2. Thatcher BJ, Doherty AE, Orvisky E, Martin BM, Henkin RI: **Gustin from human parotid saliva is carbonic anhydrase VI.** *Biochem Biophys Res Commun* 1998, **250:**635–641.
3. Fernley RT, Wright RD, Coghlan JP: **A novel carbonic anhydrase from the ovine parotid gland.** *FEBS Lett* 1979, **105:**299–302.
4. Feldstein JB, Silverman DN: **Purification and characterization of carbonic anhydrase from the saliva of the rat.** *J Biol Chem* 1984, **259:**5447–5453.
5. Murakami H, Sly WS: **Purification and characterization of human salivary carbonic anhydrase.** *J Biol Chem* 1987, **262:**1382–1388.
6. Kadoya Y, Kuwahara H, Shimazaki M, Ogawa Y, Yagi T: **Isolation of a novel carbonic anhydrase from human saliva and immunohistochemical**

demonstration of its related isozymes in salivary gland. *Osaka City Med J* 1987, **33:**99–109.
7. Parkkila S, Kaunisto K, Rajaniemi L, Kumpulainen T, Jokinen K, Rajaniemi H: **Immunohistochemical localization of carbonic anhydrase isoenzymes VI, II, and I in human parotid and submandibular glands.** *J Histochem Cytochem* 1990, **38:**941–947.
8. Parkkila S, Parkkila AK, Vierjoki T, Stahlberg T, Rajaniemi H: **Competitive time-resolved immunofluorometric assay for quantifying carbonic anhydrase VI in saliva.** *Clin Chem* 1993, **39:**2154–2157.
9. Parkkila S, Parkkila AK, Rajaniemi H: **Circadian periodicity in salivary carbonic anhydrase VI concentration.** *Acta Physiol Scand* 1995, **154:**205–211.
10. Ogawa Y, Matsumoto K, Maeda T, Tamai R, Suzuki T, Sasano H, Fernley RT: **Characterization of lacrimal gland carbonic anhydrase VI.** *J Histochem Cytochem* 2002, **50:**821–827.
11. Karhumaa P, Leinonen J, Parkkila S, Kaunisto K, Tapanainen J, Rajaniemi H: **The identification of secreted carbonic anhydrase VI as a constitutive glycoprotein of human and rat milk.** *Proc Natl Acad Sci U S A* 2001, **98:**11604–11608.
12. Kivela J, Parkkila S, Parkkila AK, Leinonen J, Rajaniemi H: **Salivary carbonic anhydrase isoenzyme VI.** *J Physiol* 1999, **520**(Pt 2):315–320.
13. Parkkila S, Parkkila AK: **Carbonic anhydrase in the alimentary tract. Roles of the different isozymes and salivary factors in the maintenance of optimal conditions in the gastrointestinal canal.** *Scand J Gastroenterol* 1996, **31:**305–317.
14. Kivela J, Parkkila S, Parkkila AK, Rajaniemi H: **A low concentration of carbonic anhydrase isoenzyme VI in whole saliva is associated with caries prevalence.** *Caries Res* 1999, **33:**178–184.
15. Pan PW, Kayra K, Leinonen J, Nissinen M, Parkkila S, Rajaniemi H: **Gene expression profiling in the submandibular gland, stomach, and duodenum of CAVI-deficient mice.** *Transgenic Res* 2011, **20:**675–698.
16. Culp DJ, Robinson B, Parkkila S, Pan PW, Cash MN, Truong HN, Hussey TW, Gullett SL: **Oral colonization by Streptococcus mutans and caries development is reduced upon deletion of carbonic anhydrase VI expression in saliva.** *Biochim Biophys Acta* 1812, **2011:**1567–1576.
17. Shatzman AR, Henkin RI: **Gustin concentration changes relative to salivary zinc and taste in humans.** *Proc Natl Acad Sci U S A* 1981, **78:**3867–3871.
18. Padiglia A, Zonza A, Atzori E, Chillotti C, Calo C, Tepper BJ, Barbarossa IT: **Sensitivity to 6-n-propylthiouracil is associated with gustin (carbonic anhydrase VI) gene polymorphism, salivary zinc, and body mass index in humans.** *Am J Clin Nutr* 2010, **92:**539–545.
19. Calo C, Padiglia A, Zonza A, Corrias L, Contu P, Tepper BJ, Barbarossa IT: **Polymorphisms in TAS2R38 and the taste bud trophic factor, gustin gene co-operate in modulating PROP taste phenotype.** *Physiol Behav* 2011, **104:**1065–1071.
20. Melis M, Aragoni MC, Arca M, Cabras T, Caltagirone C, Castagnola M, Crnjar R, Messana I, Tepper BJ, Barbarossa IT: **Marked increase in PROP taste responsiveness following oral supplementation with selected salivary proteins or their related free amino acids.** *PLoS One* 2013, **8:**e59810.
21. Chandrashekar J, Yarmolinsky D, von Buchholtz L, Oka Y, Sly W, Ryba NJ, Zuker CS: **The taste of carbonation.** *Science* 2009, **326:**443–445.
22. Krackow S, Vannoni E, Codita A, Mohammed AH, Cirulli F, Branchi I, Alleva E, Reichelt A, Willuweit A, Voikar V, Colacicco G, Wolfer DP, Buschmann JU, Safi K, Lipp HP: **Consistent behavioral phenotype differences between inbred mouse strains in the IntelliCage.** *Genes Brain Behav* 2010, **9:**722–731.
23. Bromley SM: **Smell and taste disorders: a primary care approach.** *Am Fam Physician* 2000, **61:**427–436. 438.
24. Henkin RI, Martin BM, Agarwal RP: **Decreased parotid saliva gustin/carbonic anhydrase VI secretion: an enzyme disorder manifested by gustatory and olfactory dysfunction.** *Am J Med Sci* 1999, **318:**380–391.
25. Iwai A, Shiozaki T, Miyazaki T: **Relevance of signaling molecules for apoptosis induction on influenza A virus replication.** *Biochem Biophys Res Commun* 2013, **441:**531–537.
26. Kagami H, Hiramatsu Y, Hishida S, Okazaki Y, Horie K, Oda Y, Ueda M: **Salivary growth factors in health and disease.** *Adv Dent Res* 2000, **14:**99–102.

doi:10.1186/s12929-014-0082-2
Cite this article as: Patrikainen *et al.*: The role of carbonic anhydrase VI in bitter taste perception: evidence from the *Car6*$^{-/-}$ mouse model. *Journal of Biomedical Science* 2014 **21**:82.

PUBLICATION II

Altered gene expression in the lower respiratory tract of *Car6* (-/-) mice

Maarit Patrikainen, Peiwen Pan, Harlan Barker, Seppo Parkkila

Transgenic Research, 2016 Oct;25(5):649-64
https://doi.org/10.1007/s11248-016-9961-5

ALTERED GENE EXPRESSION IN THE LOWER RESPIRATORY TRACT OF *Car6*$^{-/-}$ MICE

Maarit S. Patrikainen[a], Peiwen Pan[a], Harlan R. Barker[a], Seppo Parkkila[a,b]

[a]School of Medicine, FI-33014 University of Tampere, Finland; [b]Fimlab Ltd and Tampere University Hospital, FI-33520 Tampere, Finland

Running title: Gene expression in *Car6*$^{-/-}$ deficient mice

Key words: Carbonic anhydrase, cDNA microarray, Trachea, Lung, Immune response

Correspondence to:

Maarit Patrikainen, MSc

School of Medicine,

Medisiinarinkatu 3, FI-33014 University of Tampere,

Finland

e-mail: Maarit.Patrikainen@uta.fi

tel: +358-50-3186325

ABSTRACT

From birth, the respiratory tract mucosa is exposed to various chemical, physical, and microbiological stress factors. Efficient defense mechanisms and strictly regulated renewal systems in the mucosa are thus required. Carbonic anhydrase VI (CA VI) is the only secreted isoenzyme of the α-CA gene family. It is transported in high concentrations in saliva and milk into the alimentary tract where it contributes to optimal pH homeostasis. Earlier study of transcriptomic responses of *Car6*$^{-/-}$ mice has shown changes in the response to oxidative stress and brown fat cell differentiation in the submandibular gland. It has been suggested that CA VI delivered to the mucosal surface of the bronchiolar epithelium is an essential factor in defense and renewal of the lining epithelium. In this study, the transcriptional effects of CA VI deficiency were investigated in both trachea and lung of *Car6*$^{-/-}$ mice using a cDNA microarray analysis. Functional clustering of the results indicated significant changes of gene transcription in the lower airways. The altered biological processes included antigen transport by M-cells, potassium transport, muscle contraction, and thyroid hormone synthesis. Immunohistochemical staining confirmed the absence of CA VI in the submandibular gland of *Car6*$^{-/-}$ mice. Immunostaining of the trachea and lung samples revealed no differences between the knockout and wild type groups nor were any morphological changes observed. The present findings can help us to recognize novel functions for CA VI – one of the major protein constituents of saliva and milk.

Abbreviations

BAL	Bronchial alveolar lavage
BDNF	Brain derived neurotrophic factor
CA	Carbonic anhydrase
CHOP	CCAAT/Enhancer-Binding Protein Homologous Protein
DAVID	Database for annotation, visualization and integrated discovery tool
FC	Fold change
FPKM	Fragments per kilobase of transcript per million mapped reads
GO	Gene ontology
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
UPR^{ER}	Endoplasmic reticulum unfolded protein stress response pathway
VLAD	Visual annotation display
WT	Wild type

INTRODUCTION

Carbonic anhydrase VI (CA VI) (CA; EC4.2.1.1), a glycoprotein with molecular weight of 42 kDa, is the only secreted isozyme in the α-CA enzyme family (Fernley et al. 1979; Murakami and Sly 1987). In 1970s it was first named as gustin, a novel taste-associated protein which is present in human saliva (Henkin et al. 1975). Two decades later, Thatcher et al. (Thatcher et al. 1998) showed that gustin was indeed identical to CA VI, which had been independently identified as a CA enzyme present in the sheep parotid gland and saliva (Fernley et al. 1979), and later isolated from rat (Feldstein and Silverman 1984) and human saliva (Murakami and Sly 1987). The human *CA6* gene is located at chromosome 1p36.22 and has 8 exons and 7 introns (Jiang and Gupta 1999; Sutherland et al. 1989). Immunohistochemical studies have indicated that CA VI is highly expressed in the serous acinar cells of the parotid and submandibular glands (Kadoya et al. 1987; Parkkila et al. 1990) as well as minor salivary glands of the tongue, including von Ebner's glands (Leinonen et al. 2001). It is one of the major protein constituents of human saliva (Parkkila et al. 1993). CA VI expression has also been found in the lacrimal glands (Ogawa et al. 1995; Ogawa et al. 2002), tracheobronchial glands of developing and adult rats (Leinonen et al. 2004), and nasal glands where it may function in olfaction (Kimoto et al. 2004). High levels of secretory CA VI is produced in human milk, where the purified 42-kDa polypeptide shared 100% homology with salivary CA VI according to a protein sequence analysis (Karhumaa et al. 2001). Circadian periodicity tightly regulates CA VI secretion into saliva: the levels are low during the night and increase rapidly after awakening (Parkkila et al. 1995).

CA VI is an enzymatically active CA and as such it could maintain optimal pH homeostasis within the oral cavity and upper alimentary tract (Kivela et al. 1999a; Parkkila and Parkkila

1996). It could participate in the developmental processes of the alimentary tract during the postnatal period (Karhumaa et al. 2001). Earlier studies suggested that CA VI, as a component of the enamel pellicle (Leinonen et al. 1999) and saliva, could play a protective role against cariogenesis (Kivela et al. 1999a). CA VI activity can be modulated by both CA inhibitors and activators. Supuran´s group has shown that several anions and sulfonamides efficiently inhibit the activity of the enzyme (Nishimori et al. 2007b; Nishimori et al. 2007a). They also showed that certain amino acids and amines, such as L-Phe, D-DOPA, L-Tyr, 4-amino-L-Phe, and histamine, are excellent CA VI activators (Nishimori et al. 2007c). Thus, there are appropriate chemical tools available to modulate CA VI activity at least *in vitro*.

The roles of CA VI can be now studied using a mouse model in which targeted deletion of the gene encoding CA VI, *Car6*, causes total absence of the enzyme (Pan et al. 2011). In 2011 Culp et al. surprisingly showed by an *in vivo* cariogenesis model that *Car6*$^{-/-}$ mice had a lower rate of cariogenesis and oral colonization of *Streptococcus mutans* than the wild type (WT) controls (Culp et al. 2011). In several studies, CA VI has been linked to the regulation of taste function. Henkin's group found that hypogeusic subjects had salivary CA VI concentration as low as 20% that of normal subjects (Shatzman and Henkin 1981). Recent studies also showed a link between bitter taste modality and CA VI by finding a polymorphism in the *CA6* gene (rs2274333 (A/G)) which contributes to 6-*n*-propylthiouracil taster status (Padiglia et al. 2010). It was also shown that alterations of bitter taste function are due to polymorphic changes in both bitter receptor (*TAS2R38*) and *CA6* genes (Calo et al. 2011). According to a recent study the rs2274333 polymorphic change of the *CA6* gene affects fungiform papilla development and maintenance (Melis et al. 2013). Our group recently demonstrated in CA VI deficient mice that absence of CA

VI indeed altered the behavior of the animals for preferred bitter taste perception (Patrikainen et al. 2014).

In 1999, Sok et al. introduced a stress-induced form of CA VI-b, encoded by the murine *Car6* gene (Sok et al. 1999). CA VI-b is expressed in the cytosol after activation of the endoplasmic reticulum unfolded protein stress response pathway (UPR^{ER}), and it is regulated by the CCAAT/Enhancer-Binding Protein Homologous Protein (CHOP). Later studies with $Car6^{-/-}$ mice have shown that CA VI-b can function as part of the pro-survival branch of the CHOP signaling cascade and is necessary for conveying the beneficial effects of brain derived neurotrophic factor (BDNF) on hypoxic cell survival (Matthews et al. 2014).

Histological analyses have indicated a greater number of lymphoid follicles in the small intestinal Peyer's patches of the $Car6^{-/-}$ mice. Transcriptomic responses and functional clustering results of $Car6^{-/-}$ mice indicated a number of altered biological processes, including a Gene Ontology (GO) term for a biological process called "immune system process" in the duodenum (Pan et al. 2011).

The aim of the present study was to investigate whether the gene expression profiles differ in the trachea and lung of CA VI deficient mice compared to WT mice. As far as we know this is the first investigation on respiratory tract phenotype associated with CA VI deficiency.

METHODS

Car6$^{-/-}$ mice

Generation and phenotypic characterization of *Car6*$^{-/-}$ mice have been previously described in detail by Pan et al. (Pan et al. 2011). To obtain mice with a pure C57BL/6 strain background, the mice with the targeted allele were backcrossed more than 10 generations. The production of the knockout mouse line was approved by the Animal Experimentation Committee of the University of Oulu. Genotyping of *Car6*$^{-/-}$ mice was performed by multiplex PCR (data not shown).

cDNA microarray analysis

All microarray data reported in the present article are described in accordance with MIAME guidelines, have been deposited in NCBI's Gene Expression Omnibus public repository (http://www.ncbi.nlm.nih.gov/geo/) (Edgar et al. 2002), and are accessible through GEO Series accession number GSE73755. Trachea and lung samples were collected from eight WT and six knockout *Car6*$^{-/-}$ (KO) male mice at the age of 8 weeks. Total RNA was purified using RNeasy Mini Kit (Qiagen, Basel, Switzerland), including DNase digestion according to manufacturer's instructions. Optical density (OD) was measured using ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, USA) to determine RNA concentrations and purity. All the samples had an OD260/OD280 ratio 2.0 or higher. The 32 RNA samples including 8 WT, and 6 KO for trachea; 8 WT, and 6 KO for lung, and 4 sample duplicates as internal controls were analyzed individually in Gen-Probe Life Sciences Ltd. at Oak Business Park in Manchester, UK. 150 ng of total RNA from each sample was converted to double-stranded cDNA, and amplified using Illumina TotalPrep-96 RNA Amplification Kit (Ambion, Inc., TX, USA) according to the

protocol recommended by the manufacturer in order to gain biotin-labeled cRNA. The cRNA corresponding to the polyadenylated mRNA fractions were normalized, and 750 ng of each sample were hybridized onto Illumina's Sentrix® Mouse Ref-8 v2 Bead Chips (Illumina, Inc., San Diego, CA) at + 58°C for 14-20 hours according to Illumina® Whole-Genome Gene Expression Direct Hybridization Assay manual (Illumina, Inc.). Arrays were washed and stained with Cy3-Streptavidin. The Illumina IScan (Illumina, Inc.) was used to scan the arrays according to manufacturer's instructions. Samples were analyzed using GenomeStudio (2010.3) Gene Expression Module (1.8.0) software (Illumina, Inc.).

Data analyses

Microarray data of the trachea and lung were normalized separately using the quantile normalization method and were analyzed with Chipster™ v2.9.0 (http://chipster.csc.fi/) (Kallio et al. 2011). Quality control of the data included non-metric multidimensional scaling, dendrograms, and hierarchical clustering (data not shown). The data were filtered according to the SD of the probes. The percentage of data that did not pass through the filter was adjusted to 99%, implicating a SD value of almost 3. At this point, statistical analysis was performed using the empirical Bayes t-test for the comparison of 2 groups. Filtering was applied to the data according to p-values. The remaining 750 probes in the trachea and 51 probes in lung were further filtered according to up- and down-regulated expression fold change with cut-off values of ±1.4, which have been considered to have high correlation between microarray and quantitative RT-PCR data (Morey et al. 2006).

A functional annotation analysis from the trachea data was performed using VLAD Gene List Analysis and Visualization tool (http://proto.informatics.jax.org/prototypes/vlad/). The VLAD tool performs a statistical analysis to test for the enrichment of ontology terms in a set of genes based on their annotations to bio-ontologies (Richardson and Bult 2015).

The functional annotation tool DAVID (Database for Annotation, Visualization and Integrated Discovery) (http://david.abcc.ncifcrf.gov/), is a Web-based application that provides a high-throughput and integrative gene functional annotation environment to systematically extract biological themes behind large gene lists (Huang da et al. 2009a; Huang da et al. 2009b). In this case, it was used to identify enriched biological categories among the regulated genes as compared to all the genes present in Illumina's Sentrix® Mouse Ref-8 v2 Bead Chips (Illumina, Inc.). The DAVID currently uses over 40 annotation categories, including GO terms, protein-protein interactions, protein functional domains, disease associations, bio-pathways, sequence general features, homologies, gene functional summaries, gene tissue expressions, and literatures (Dennis et al. 2003; Huang da et al. 2009a).

Quantitative real-time PCR (qRT-PCR)

The 10 µg RNA samples, extracted for microarray analysis from WT and *Car6*$^{-/-}$ mice from trachea and lung, were converted to cDNA using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Inc., CA, USA) with random hexamer primers according to manufacturer's instructions. Target genes used for validation of microarray analysis were chosen based on their fold change and known function. Primer sequences of the housekeeping gene, *mouse* β-*actin* (*Actb*), and the target genes, functions of the target genes, amplicon sizes and

annealing temperatures are shown in Table 1. The qRT-PCR reactions of duplicate samples were performed using Power SYBR® Green PCR Master Mix (Applied Biosystems, Inc.) on 96-Well Optical Reaction Plate (ABI PRISM™, Applied Biosystems, Inc.). A total of 15 µl of each qRT-PCR reaction contained 0.5 µl of first strand cDNA, 0.5 µM of each primer, and 1× SYBR Green Master Mix (Applied Biosystems, Inc.). Amplification and detection were carried out using ABI PRISM®7000 Detection System and Sequence Detection System v 1.2.3 (Applied Biosystems, Inc.). The relative expression levels of genes were assessed using *Actb* values as internal control using a mathematical model for relative quantification in real-time RT-PCR (Pfaffl 2001). The median value of wild type data for each gene was used as a control for normalization of the knockout and wildtype data for that particular gene. The final qRT-PCR results were presented as expression of the gene-of-interest relative to the housekeeping gene *Actb*. Statistical analyses were performed on IBM SPSS Statistics version 23.0.0 (IBM SPSS Statistics, IBM Corporation, Chicago, IL). Because of the small sample size, the Mann–Whitney test was used for evaluation of *Car6*$^{-/-}$ and WT mice relative expression values for the qRT-PCR. The relative expression levels of genes were assessed using *Actb* values as internal control using mathematical model for relative quantification in real-time RT-PCR (Pfaffl, 2001).

RNA-Seq analysis of *Car6* expression

RNA-Seq data from mouse lung (normal (E-GEOD-64027), influenza (E-GEOD-49933) (Altboum et al. 2014), interferon (E-GEOD-55480), bronchial alveolar lavage (BAL) (E-ERAD-397) and bronchiolar epithelium (E-GEOD-61627) (Lange et al. 2015)) was retrieved from the ArrayExpress database (Kolesnikov et al. 2015). RNA-Seq reads were mapped to CA genes in

the mouse genome (release GRCM38) using the Bowtie (Langmead et al. 2009) and Tophat (Trapnell et al. 2009) modules of the Tophat package. Mapped reads were compared to reference CA transcript structures from the Ensembl database using the Cuffdiff module of the Cufflinks package (Trapnell et al. 2010).

GO enrichment analysis of *CA6* in lung

The Medisapiens database (Kilpinen et al. 2008) (http://ist.medisapiens.com/) houses data from a large number of human microarray expression studies. For lung tissue there were between 34 and 182 microarray experiment samples for all genes present on a human genome U133 chip. We extracted all expression correlation values for *CA6* with all other genes. In total there were 208 genes which had a positive correlation of expression of 0.50 or greater. These gene names were then used to perform a biological processes gene ontology enrichment analysis using the (http://geneontology.org) webserver (Gene Ontology Consortium 2015; Ashburner et al. 2000). The gene ontology terms which were enriched at 2-fold or greater were then assessed for biological significance.

Immunohistochemical analysis

Tissue samples from the trachea, lung and submandibular gland were collected from adult WT, and *Car6*$^{-/-}$ mice. The samples were fixed in 4% paraformaldehyde for 24 hours, embedded in paraffin and cut to 5 μm thick sections. A polyclonal rabbit anti-rat CA VI serum (Leinonen et al. 2001) was used for immunohistochemical staining as described below. Prior to staining, the

slides were treated with xylene, and series of decreasing ethanol concentration baths to remove paraffin. Endogenous peroxidase was removed with 3% H_2O_2 in phosphate-buffered saline (PBS) for five minutes, and for blocking 10 % normal goat serum (NGS) was used. A Vectastain Elite ABC kit (Vector Laboratories, Burlingame, CA, USA) was used following to the manufacturer's instructions. Sections were incubated with primary antibody (rabbit anti-rat CA VI in 2% NGS/0.1% Tween in PBS) at + 4 ˚C overnight. Biotinylated anti-rabbit IgG (Vector Laboratories) in 2% NGS/0.1 % Tween in PBS was used as the secondary antibody. Labels were detected using 3, 3'-diaminobenzidine (DAB), and sections were counterstained using Mayer's Hematoxylin.

RESULTS

cDNA microarray analysis

In order to identify genes with an altered transcription level in lower airways of CA VI deficient mice, RNAs of the trachea and lung from six *Car6*$^{-/-}$ and eight WT mice were analyzed by microarray. A total of 44 genes in the trachea and two genes in the lung were either up- or down-regulated. The fold change cut-off value was set at ± 1.4, a proposed level above which there is a high correlation between microarray and qRT-PCR data regardless of other factors, such as cycle threshold or spot intensity (Morey et al. 2006). All the genes above the fold change threshold ± 1.4 in our data set had significantly ($p<0.05$) altered expression. Five genes were up-regulated in the trachea of *Car6*$^{-/-}$ mice as compared to WT mice, including *Sftpc* with the highest FC (+2.13), *Nppa*, *Ucp1*, *Myl4*, and *Faim3*. The 39 down-regulated genes included *Calca* with the lowest FC (-3.48). In the lung, the FC threshold indicated differential expression of two genes: *Gdpd3* FC (+1.45) and *Dbp* FC (-1.45). A complete list of genes with altered expression levels, their symbols, gene descriptions, fold change, and *P*-value, are shown in Table 2.

Functional annotation analysis of genes with altered expression in the trachea of *Car6*$^{-/-}$ mice

A functional annotation analysis from the trachea data was performed using the VLAD Gene List Analysis and Visualization tool. The results, shown in ontology hierarchy graphs, for differentially expressed genes (n=40; Expi and NA were unannotated), revealed changes in metabolic process, biological regulation, single-organism process, and immune response in mucosal-associated lymphoid tissue (Fig. 1). Genes with induced expression were located to

single-organism process. The cellular component result, given in Online Resource 1, indicated expression level changes in the genes which encode cytoskeletal, cytoplasmic and extracellular proteins. The results of molecular function, given in Online Resource 2, showed that the differentially expressed genes are linked to transmembrane transporter, antioxidant and catalytic activity, protein binding, and are structural constituents of cytoskeleton.

The DAVID Database for Annotation, Visualization and Integrated Discovery tool was used to identify enriched biological categories of the trachea data. The DAVID functional table of annotations reported 30 out of 42 genes with DAVID IDs which were categorized by different associated enriched terms. The following up-regulated genes in the trachea were categorized functionally by the associated terms: *Myl4* was categorized in “muscle protein”, “myosin”, and “motor protein”; *Nppa* in “secreted”, “signal”, “hormone”, and “disulfide bond”; *Sftpc* in “secreted”; *Faim3* in “signal”, and “disulfide bond”. Three most common functionally categorized terms among the down-regulated genes were “signal” (15 genes), “secreted” (10 genes), and “disulfide bond” (10 genes). Many of the down-regulated genes were annotated with several functional category terms. A complete list of terms and different functional annotation categories are given in Online Resource 3.

qRT-PCR

Based on Chipster ^{TM}, DAVID, and VLAD analyses nine differentially expressed genes were selected for qRT-PCR. The same RNA samples used for the microarray analysis were also used for the qRT-PCR. Statistical comparison between relative expression ratios of *Car6*$^{-/-}$ and WT groups revealed statistically significant differences ($p < 0.05$) in *Faim3* (p-value 0.013), *Sftpc* (p-

value 0.020), *Dcpp1* (p-value 0.043), and *Dbp* (p-value 0.043) expression. The fold change was counted using median values of gene expression ratio of genes-of-interest. The results of statistical analyses between *Car6*$^{-/-}$ and WT mice groups' relative gene expression ratios, p-values, and fold change are shown in Table 3.

RNA-Seq analysis of *Car6* expression

The RNA-Seq data showed that *Car6* is not expressed in mouse normal lung tissues. However, when the mice were infected with influenza virus there was a marked *Car6* expression in an occasional lung sample. Results are shown as Fragments per kilobase of transcript per million mapped reads (FPKM) in Fig. 2. Additionally, BAL of influenza virus infected mice also showed *Car6* expression in several samples. To a lesser extent we also saw *Car6* expression in interferon treated lung tissue.

GO enrichment analysis of *CA6* in lung

The results of the GO enrichment analysis of genes significantly coexpressed with human *CA6* showed three over-represented biological terms: axonemal dynein complex assembly (GO:0070286) over-represented more than 26-fold, cilium assembly (GO:0042384) over-represented more than 7-fold, and cilium organization (GO:0044782) over-represented more than 6-fold. All three of these terms are related to cellular cilia.

Immunohistochemistry

Tissue samples stained with anti-rat CA VI serum showed an absence of CA VI in submandibular gland acinar cells of *Car6*$^{-/-}$ mice (b in Fig. 3). The corresponding tissue samples from the WT mice presented a clear positive CA VI stain in the acinar cells. Both trachea and lung tissue samples showed only a faint background staining. There were no morphological differences between *Car6*$^{-/-}$ and WT mice in the submandibular gland, trachea, and lung tissues (Fig. 3).

DISCUSSION

The present data demonstrates that CA VI deficiency in mice alters the expression of several genes in the lower respiratory tract; these changes do not cause any marked morphological changes at light microcopy level. A previous study of *Car6*$^{-/-}$ mice showed altered gene expression, given as FC, in the submandibular gland (-2.70 to +5.18), stomach (-9.40 to +2.78), and duodenum (-15.56 to +5.31) (Pan et al. 2011). In the present study, we observed differentially expressed genes, given as FC, in the trachea (-3.48 to +2.13) and lung (-1.45 to +1.45). The FC levels indicate that CA VI deficiency changes the gene expression levels less drastically in the lower airways compared to the alimentary tract tissues. Even though the present FC values are considered moderate at the most, all genes above or below the FC threshold were statistically significantly altered ($p<0.05$). Interestingly, there were both induced and repressed genes in the trachea that were also differently expressed in the submandibular gland (*UCP1, BC048546, Myl1, 2310057J18Rik, Dcpp1, Dcpp2 and Dcpp3)*, stomach *(Tpm2, Mylpf, Myl1, Expi*), and duodenum (*BC048546, Lyz1, 2310057J18Rik, Expi*). *Ucp1* was up-regulated in the trachea but down-regulated in the submandibular gland; *Myl1* was repressed in the trachea and submandibular gland but induced in the stomach; *2310057J18Rik* was repressed in the trachea and duodenum but induced in the submandibular gland; *BC048546* was repressed in the trachea but induced in the submandibular gland and duodenum; *Lyz1* was up-regulated in the duodenum but down-regulated in the trachea. *Dcpp1*, *Dcpp2*, and *Dcpp3* expressions in the trachea were repressed, whereas all of these genes were induced in the submandibular gland. *Expi* was the only gene which was repressed in the stomach and duodenum as well as in the trachea.

The resulting data of qRT-PCR analysis revealed statistically significant differences in relative expression ratios in the trachea of induced genes *Sftpc,* and *Faim3*, and repressed gene *Dcpp1*. The repressed *Dbp* shows statistically significant difference in the lung. Pulmonary surfactant is comprised of two types of specific surfactant proteins: hydrophilic SP-A and SP-D and hydrophobic SP-B and SP-C, which is expressed by the gene *Sftpc*. They have dual functions, i.e. lowering the surface tension and participation in innate immune defense of the lung and other mucosal surfaces (Haagsman and Diemel 2001). Both SP-B and SP-C enhance lipid insertion into the monolayer at the air/liquid interface in the peripheral air spaces, lowering the surface tension upon inhalation, and protects the surface film from non-surfactant protein contamination (Curstedt et al. 1987; Hawgood et al. 1987). SP-C is expressed only in type II cells (Keller et al. 1992) and surfactant protein B is expressed in alveolar type II epithelial cells and Club cells (originally known as Clara cells) (Cruz et al. 1995; Lin et al. 1996). Leinonen and coworkers previously suggested that CA VI is delivered to the mucosal surface from the serous acinar cells of the tracheobronchial glands and Club cells of the bronchiolar epithelium (Leinonen et al. 2004). In the airways it may be an essential factor for the defense system and also participate in renewal of the lining epithelium. Surfactant SP-C and SP-A are collagen-containing, (C-type) calcium-dependent lectins called collectins, which interact with glycoconjugates and lipids on the surface of microorganisms mostly through their carbohydrate recognition domains (CRDs), and are bound by protein deleted in malignant brain tumors 1 (DMBT1) (Kishore et al. 2006). A previous study on another CA isozyme suggested that CA IX deficiency can induce an immune process in the gastric mucosa, which is associated with upregulated expression of surfactant associated protein D and Dmbt1 (Kallio et al. 2010). In the trachea of *Car6*$^{-/-}$ mice, *Dmbt1* gene expression was repressed (FC -2.52). The human DMBT1 glycoprotein, expressed in the salivary

glands, airways, and genital tract, binds various bacterial pathogens and viruses, and its role in innate immune defense has been demonstrated in various studies (Prakobphol et al. 2000; Stoddard et al. 2007; White et al. 2005; Wu et al. 2003).

The presence of *Car6* transcripts in immune stimulated lung tissues, and lack in normal lung tissues, implies some role for *Car6* in immune function in mice. The GO enrichment analysis terms over-represented in our analysis were all related to cellular cilia. These organelles exist outside of the cell and can be either motile or primary (non-motile) (Yokoyama 2004). The strong over-representation of the axonemal dynein complex assembly biological process implies that *CA6* is involved in motile cilia, as dynein is involved in bending of cilial microtubules associated with movement. The bending of cilia is especially prominent in air passageways such as the trachea and bronchi where they contribute to the action of the innate immune system by moving mucus away from the lungs (Yokoyama 2004). Importantly, previous immunostaining of CA VI has shown that it localizes to the basal areas of cilia in rats (Leinonen et al. 2004) and dogs (Sugiura et al. 2008).

As we have demonstrated previously, *Car6* is involved in bitter taste perception in mouse (Patrikainen et al. 2014). A relatively new revelation in our understanding of bitter taste receptors (T2Rs) is that they occur not only in the oral cavity, where they contribute to detection of noxious compounds, but also in air passageways where they can help detect bacterial byproducts (Devillier et al. 2015; Krasteva-Christ et al. 2015). Like CA VI, T2Rs are found located in cilia and apical membrane of sinonasal passageways (Lee et al. 2012) as well as cilia of epithelial cells in lower respiratory tract. Activation of these receptors results in increased intracellular Ca^{2+} and subsequently an increase in the frequency of cilia beating (Shah et al. 2009; Braiman and Priel 2008). The most strongly down regulated protein in our study was

calcitonin/calcitonin related peptide, with a fold change of -3.48. This protein has been shown to be involved in positive intracellular Ca^{2+} changes in pre-synaptic taste cells (Huang and Wu 2015).

Our group showed earlier that CA VI secretion is tightly regulated by circadian periodicity, the levels are low during the night and increase rapidly after awakening (Parkkila et al. 1995). The only significantly repressed gene in the lung was *Dbp*, which protein product functions as a transcription factor, controlling many target genes with a circadian expression pattern (Ripperger et al. 2000). Even though we do not know yet the exact mechanisms by which CA VI deficiency leads to down-regulation of *Dbp* gene, this finding is of interest and perhaps stimulates more studies on diurnal changes of gene expression in the lower respiratory tract. These studies would be relevant to certain clinical conditions including asthma. The nocturnal worsening of asthma manifests as a reduction in lung function and an increase in bronchial hyper-reactivity (Lewis et al. 2006). These changes are associated to diurnal changes in both hormone concentrations and autonomic nervous system control (Busse 1988).

To study morphological differences between the *Car6*$^{-/-}$ and WT mice, we collected tissue samples from trachea, lung and submandibular gland and investigated them by light microscopy. Previous reports have shown positive CA VI staining in the tracheobronchial glands and the Club cells of the bronchiolar surface epithelium of rat, thus suggesting that CA VI has a mucosa-protective role not only in the gastrointestinal tract but also in the respiratory tract (Leinonen et al. 2004). The present investigations did not reveal any morphological changes attributable to *Car6* deficiency, indicating that CA VI has no major role in the developmental regulation of the lower respiratory tract tissues. In contrast, the VLAD results suggested that CA VI might be involved in the regulation of several unforeseen biological processes, such as antigen transfer to

mucosal-associated lymphoid tissue by M-cells and potassium ion transport. In the GO enrichment analysis, the coexpression of genes related to cilial function with CA VI and up-regulation of CA VI in experimental lung inflammation suggest to us that CA VI might have a role in enhancing cilial motion. This hypothesis is further supported by the combination of facts that a) T2R agonists increase the beating frequency of cilia (Shah et al. 2009), b) T2Rs, the primary receptors of bitter taste, are expressed in the lung (Yokoyama 2004), and c) CA VI is involved in the perception of bitter taste (Patrikainen et al. 2014). These facts lead us to speculate that CA VI might sensitize T2Rs to their agonists, leading to increased perception of bitter taste and enhanced cilial motion. Exploring the exact interactions remains an exciting topic for further studies. These are novel putative functions associated with CA VI, which clearly warrant further immunological and electrophysiological investigations.

FUNDING

This work was supported by grants from the Academy of Finland, Sigrid Jusélius Foundation, Tampere Tuberculosis Foundation, and Finnish Cultural Foundation.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

ETHICAL APPROVAL

The production of the knockout mouse line was approved by the Animal Experimentation Committee of the University of Oulu, Finland. The health status of the animals was monitored on a regular basis in accordance to the FELASA recommendations.

ACKNOWLEDGEMENTS

The authors thank Aulikki Lehmus and Marianne Kuuslahti for their skillful technical assistance, Tiina Luukkaala for performing statistical analyses, and Dr. Martti Tolvanen for the valuable help on the discussion section.

REFERENCES

Altboum Z, Steuerman Y, David E, Barnett-Itzhaki Z, Valadarsky L, Keren-Shaul H, Meningher T, Mendelson E, Mandelboim M, Gat-Viks I, Amit I (2014) Digital cell quantification identifies global immune cell dynamics during influenza infection. Mol Syst Biol 10:720

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25:25-29

Bekhor I, Wen Y, Shi S, Hsieh CH, Denny PA, Denny PC (1994) cDNA cloning, sequencing and in situ localization of a transcript specific to both sublingual demilune cells and parotid intercalated duct cells in mouse salivary glands. Arch Oral Biol 39:1011-1022

Braiman A, Priel Z (2008) Efficient mucociliary transport relies on efficient regulation of ciliary beating. Respir Physiol Neurobiol 163:202-207

Busse WW (1988) Pathogenesis and pathophysiology of nocturnal asthma. Am J Med 85:24-29

Calo C, Padiglia A, Zonza A, Corrias L, Contu P, Tepper BJ, Barbarossa IT (2011) Polymorphisms in TAS2R38 and the taste bud trophic factor, gustin gene co-operate in modulating PROP taste phenotype. Physiol Behav 104:1065-1071

Cannon B, Nedergaard J (2004) Brown adipose tissue: function and physiological significance. Physiol Rev 84:277-359

Cruz A, Casals C, Perez-Gil J (1995) Conformational flexibility of pulmonary surfactant proteins SP-B and SP-C, studied in aqueous organic solvents. Biochim Biophys Acta 1255:68-76

Culp DJ, Robinson B, Parkkila S, Pan PW, Cash MN, Truong HN, Hussey TW, Gullett SL (2011) Oral colonization by Streptococcus mutans and caries development is reduced upon deletion of carbonic anhydrase VI expression in saliva. Biochim Biophys Acta 1812:1567-1576

Curstedt T, Jornvall H, Robertson B, Bergman T, Berggren P (1987) Two hydrophobic low-molecular-mass protein fractions of pulmonary surfactant. Characterization and biophysical activity. Eur J Biochem 168:255-262

De Bold AJ (1985) Atrial natriuretic factor: a hormone produced by the heart. Science 230:767-770

Dennis G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 4:R60

Devillier P, Naline E, Grassin-Delyle S (2015) The pharmacology of bitter taste receptors and their role in human airways. Pharmacol Ther 155:11-21

Edgar R, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30:207-210

Feldstein JB, Silverman DN (1984) Purification and characterization of carbonic anhydrase from the saliva of the rat. J Biol Chem 259:5447-5453

Felsenfeld AJ, Levine BS (2015) Calcitonin, the forgotten hormone: does it deserve to be forgotten?. Clin Kidney J 8:180-187

Fernley RT, Wright RD, Coghlan JP (1979) A novel carbonic anhydrase from the ovine parotid gland. FEBS Lett 105:299-302

Gene Ontology Consortium (2015) Gene Ontology Consortium: going forward. Nucleic Acids Res 43:D1049-56

Gordon S, Todd J, Cohn ZA (1974) In vitro synthesis and secretion of lysozyme by mononuclear phagocytes. J Exp Med 139:1228-1248

Haagsman HP, Diemel RV (2001) Surfactant-associated proteins: functions and structural variation. Comp Biochem Physiol A Mol Integr Physiol 129:91-108

Hawgood S, Benson BJ, Schilling J, Damm D, Clements JA, White RT (1987) Nucleotide and amino acid sequences of pulmonary surfactant protein SP 18 and evidence for cooperation between SP 18 and SP 28-36 in surfactant lipid adsorption. Proc Natl Acad Sci U S A 84:66-70

Henkin RI, Lippoldt RE, Bilstad J, Edelhoch H (1975) A zinc protein isolated from human parotid saliva. Proceedings of the National Academy of Sciences 72:488-492

Huang da W, Sherman BT, Lempicki RA (2009a) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4:44-57

Huang da W, Sherman BT, Lempicki RA (2009b) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37:1-13

Huang AY, Wu SY (2015) Calcitonin Gene-Related Peptide Reduces Taste-Evoked ATP Secretion from Mouse Taste Buds. J Neurosci 35:12714-12724

Jiang W, Gupta D (1999) Structure of the carbonic anhydrase VI (CA6) gene: evidence for two distinct groups within the alpha-CA gene family. Biochem J 344 Pt 2:385-390

Kadoya Y, Kuwahara H, Shimazaki M, Ogawa Y, Yagi T (1987) Isolation of a novel carbonic anhydrase from human saliva and immunohistochemical demonstration of its related isozymes in salivary gland. Osaka City Med J 33:99-109

Kallio H, Hilvo M, Rodriguez A, Lappalainen EH, Lappalainen AM, Parkkila S (2010) Global transcriptional response to carbonic anhydrase IX deficiency in the mouse stomach. BMC Genomics 11:397-2164-11-397

Kallio MA, Tuimala JT, Hupponen T, Klemela P, Gentile M, Scheinin I, Koski M, Kaki J, Korpelainen EI (2011) Chipster: user-friendly analysis software for microarray and other high-throughput data. BMC Genomics 12:507-2164-12-507

Karhumaa P, Leinonen J, Parkkila S, Kaunisto K, Tapanainen J, Rajaniemi H (2001) The identification of secreted carbonic anhydrase VI as a constitutive glycoprotein of human and rat milk. Proc Natl Acad Sci U S A 98:11604-11608

Keller A, Steinhilber W, Schafer KP, Voss T (1992) The C-terminal domain of the pulmonary surfactant protein C precursor contains signals for intracellular targeting. Am J Respir Cell Mol Biol 6:601-608

Kilpinen S, Autio R, Ojala K, Iljin K, Bucher E, Sara H, Pisto T, Saarela M, Skotheim RI, Bjorkman M, Mpindi JP, Haapa-Paananen S, Vainio P, Edgren H, Wolf M, Astola J, Nees M, Hautaniemi S, Kallioniemi O (2008) Systematic bioinformatic analysis of expression levels of 17,330 human genes across 9,783 samples from 175 types of healthy and pathological tissues. Genome Biol 9:R139-2008-9-9-r139. Epub 2008 Sep 19

Kimoto M, Iwai S, Maeda T, Yura Y, Fernley RT, Ogawa Y (2004) Carbonic anhydrase VI in the mouse nasal gland. J Histochem Cytochem 52:1057-1062

Kishore U, Greenhough TJ, Waters P, Shrive AK, Ghai R, Kamran MF, Bernal AL, Reid KB, Madan T, Chakraborty T (2006) Surfactant proteins SP-A and SP-D: structure, function and receptors. Mol Immunol 43:1293-1315

Kivela J, Parkkila S, Parkkila AK, Leinonen J, Rajaniemi H (1999a) Salivary carbonic anhydrase isoenzyme VI. J Physiol 520 Pt 2:315-320

Kolesnikov N, Hastings E, Keays M, Melnichuk O, Tang YA, Williams E, Dylag M, Kurbatova N, Brandizi M, Burdett T, Megy K, Pilicheva E, Rustici G, Tikhonov A, Parkinson H, Petryszak R, Sarkans U, Brazma A (2015) ArrayExpress update--simplifying data submissions. Nucleic Acids Res 43:D1113-6

Kouadjo K, Nishida Y, Cadrin-Girard J, Yoshioka M, St-Amand J (2007) Housekeeping and tissue-specific genes in mouse tissues. BMC Genomics 8:127

Krasteva-Christ G, Soultanova A, Schutz B, Papadakis T, Weiss C, Deckmann K, Chubanov V, Gudermann T, Voigt A, Meyerhof W, Boehm U, Weihe E, Kummer W (2015) Identification of cholinergic chemosensory cells in mouse tracheal and laryngeal glandular ducts. Int Immunopharmacol 29:158-165

Lange AW, Sridharan A, Xu Y, Stripp BR, Perl AK, Whitsett JA (2015) Hippo/Yap signaling controls epithelial progenitor cell proliferation and differentiation in the embryonic and adult lung. J Mol Cell Biol 7:35-47

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25-2009-10-3-r25. Epub 2009 Mar 4

Lee RJ, Xiong G, Kofonow JM, Chen B, Lysenko A, Jiang P, Abraham V, Doghramji L, Adappa ND, Palmer JN, Kennedy DW, Beauchamp GK, Doulias PT, Ischiropoulos H, Kreindler JL, Reed DR, Cohen NA (2012) T2R38 taste receptor polymorphisms underlie susceptibility to upper respiratory infection. J Clin Invest 122:4145-4159

Leinonen J, Kivela J, Parkkila S, Parkkila AK, Rajaniemi H (1999) Salivary carbonic anhydrase isoenzyme VI is located in the human enamel pellicle. Caries Res 33:185-190

Leinonen JS, Saari KA, Seppanen JM, Myllyla HM, Rajaniemi HJ (2004) Immunohistochemical demonstration of carbonic anhydrase isoenzyme VI (CA VI) expression in rat lower airways and lung. J Histochem Cytochem 52:1107-1112

Leinonen J, Parkkila S, Kaunisto K, Koivunen P, Rajaniemi H (2001) Secretion of Carbonic Anhydrase Isoenzyme VI (CA VI) from Human and Rat Lingual Serous von Ebner's Glands. Journal of Histochemistry & Cytochemistry 49:657-662

Lewis MJ, Short AL, Lewis KE (2006) Autonomic nervous system control of the cardiovascular and respiratory systems in asthma. Respir Med 100:1688-1705

Lin S, Phillips KS, Wilder MR, Weaver TE (1996) Structural requirements for intracellular transport of pulmonary surfactant protein B (SP-B). Biochim Biophys Acta 1312:177-185

Markart P, Faust N, Graf T, Na CL, Weaver TE, Akinbi HT (2004) Comparison of the microbicidal and muramidase activities of mouse lysozyme M and P. Biochem J 380:385-392

Matthews TA, Abel A, Demme C, Sherman T, Pan PW, Halterman MW, Parkkila S, Nehrke K (2014) Expression of the CHOP-inducible carbonic anhydrase CAVI-b is required for BDNF-mediated protection from hypoxia. Brain Res 1543:28-37

Matthias A, Ohlson KB, Fredriksson JM, Jacobsson A, Nedergaard J, Cannon B (2000) Thermogenic responses in brown fat cells are fully UCP1-dependent. UCP2 or UCP3 do not substitute for UCP1 in adrenergically or fatty scid-induced thermogenesis. J Biol Chem 275:25073-25081

Melis M, Atzori E, Cabras S, Zonza A, Calo C, Muroni P, Nieddu M, Padiglia A, Sogos V, Tepper BJ, Tomassini Barbarossa I (2013) The Gustin (CA6) Gene Polymorphism, rs2274333 (A/G), as a Mechanistic Link between PROP Tasting and Fungiform Taste Papilla Density and Maintenance. PLoS One 8:e74151

Morey JS, Ryan JC, Van Dolah FM (2006) Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. Biol Proced Online 8:175-193

Mullins JJ, Mullins LJ, Dunbar DR, Brammar WJ, Gross KW, Morley SD (2006) Identification of a human ortholog of the mouse Dcpp gene locus, encoding a novel member of the CSP-1/Dcpp salivary protein family. Physiol Genomics 28:129-140

Murakami H, Sly WS (1987) Purification and characterization of human salivary carbonic anhydrase. J Biol Chem 262:1382-1388

Nishimori I, Innocenti A, Vullo D, Scozzafava A, Supuran CT (2007a) Carbonic anhydrase inhibitors. Inhibition studies of the human secretory isoform VI with anions. Bioorg Med Chem Lett 17:1037-1042

Nishimori I, Minakuchi T, Onishi S, Vullo D, Scozzafava A, Supuran CT (2007b) Carbonic anhydrase inhibitors. DNA cloning, characterization, and inhibition studies of the human secretory isoform VI, a new target for sulfonamide and sulfamate inhibitors. J Med Chem 50:381-388

Nishimori I, Onishi S, Vullo D, Innocenti A, Scozzafava A, Supuran CT (2007c) Carbonic anhydrase activators: the first activation study of the human secretory isoform VI with amino acids and amines. Bioorg Med Chem 15:5351-5357

Ogawa Y, Matsumoto K, Maeda T, Tamai R, Suzuki T, Sasano H, Fernley RT (2002) Characterization of lacrimal gland carbonic anhydrase VI. J Histochem Cytochem 50:821-827

Ogawa Y, Toyosawa S, Inagaki T, Hong SS, Ijuhin N (1995) Carbonic anhydrase isozyme VI in rat lacrimal gland. Histochem Cell Biol 103:387-394

Ohshima N, Kudo T, Yamashita Y, Mariggio S, Araki M, Honda A, Nagano T, Isaji C, Kato N, Corda D, Izumi T, Yanaka N (2015) New members of the mammalian glycerophosphodiester phosphodiesterase family: GDE4 and GDE7 produce lysophosphatidic acid by lysophospholipase D activity. J Biol Chem 290:4260-4271

Padiglia A, Zonza A, Atzori E, Chillotti C, Calo C, Tepper BJ, Barbarossa IT (2010) Sensitivity to 6-n-propylthiouracil is associated with gustin (carbonic anhydrase VI) gene polymorphism, salivary zinc, and body mass index in humans. Am J Clin Nutr 92:539-545

Pan PW, Kayra K, Leinonen J, Nissinen M, Parkkila S, Rajaniemi H (2011) Gene expression profiling in the submandibular gland, stomach, and duodenum of CAVI-deficient mice. Transgenic Res 20:675-698

Parkkila S, Parkkila AK (1996) Carbonic anhydrase in the alimentary tract. Roles of the different isozymes and salivary factors in the maintenance of optimal conditions in the gastrointestinal canal. Scand J Gastroenterol 31:305-317

Parkkila S, Parkkila AK, Rajaniemi H (1995) Circadian periodicity in salivary carbonic anhydrase VI concentration. Acta Physiol Scand 154:205-211

Parkkila S, Parkkila AK, Vierjoki T, Stahlberg T, Rajaniemi H (1993) Competitive time-resolved immunofluorometric assay for quantifying carbonic anhydrase VI in saliva. Clin Chem 39:2154-2157

Parkkila S, Kaunisto K, Rajaniemi L, Kumpulainen T, Jokinen K, Rajaniemi H (1990) Immunohistochemical localization of carbonic anhydrase isoenzymes VI, II, and I in human parotid and submandibular glands. J Histochem Cytochem 38:941-947

Patrikainen M, Pan P, Kulesskaya N, Voikar V, Parkkila S (2014) The role of carbonic anhydrase VI in bitter taste perception: evidence from the Car6(-)/(-) mouse model. J Biomed Sci 21:82-014-0082-2

Pearse AG (1966) The cytochemistry of the thyroid C cells and their relationship to calcitonin. Proc R Soc Lond B Biol Sci 164:478-487

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29:e45

Prakobphol A, Xu F, Hoang VM, Larsson T, Bergstrom J, Johansson I, Frangsmyr L, Holmskov U, Leffler H, Nilsson C, Boren T, Wright JR, Stromberg N, Fisher SJ (2000) Salivary agglutinin, which binds Streptococcus mutans and Helicobacter pylori, is the lung scavenger receptor cysteine-rich protein gp-340. J Biol Chem 275:39860-39866

Richardson JE, Bult CJ (2015) Visual annotation display (VLAD): a tool for finding functional themes in lists of genes. Mamm Genome

Ripperger JA, Shearman LP, Reppert SM, Schibler U (2000) CLOCK, an essential pacemaker component, controls expression of the circadian transcription factor DBP. Genes Dev 14:679-689

Shah AS, Ben-Shahar Y, Moninger TO, Kline JN, Welsh MJ (2009) Motile cilia of human airway epithelia are chemosensory. Science 325:1131-1134

Shatzman AR, Henkin RI (1981) Gustin concentration changes relative to salivary zinc and taste in humans. Proceedings of the National Academy of Sciences 78:3867-3871

Shima H, Takatsu H, Fukuda S, Ohmae M, Hase K, Kubagawa H, Wang JY, Ohno H (2010) Identification of TOSO/FAIM3 as an Fc receptor for IgM. Int Immunol 22:149-156

Sok J, Wang XZ, Batchvarova N, Kuroda M, Harding H, Ron D (1999) CHOP-Dependent stress-inducible expression of a novel form of carbonic anhydrase VI. Mol Cell Biol 19:495-504

Stoddard E, Cannon G, Ni H, Kariko K, Capodici J, Malamud D, Weissman D (2007) gp340 expressed on human genital epithelia binds HIV-1 envelope protein and facilitates viral transmission. J Immunol 179:3126-3132

Sugiura Y, Ichihara N, Nishita T, Murakami M, Amasaki H, Asari M (2008) Immunohistolocalization and gene expression of secretory carbonic anhydrase isoenzyme CA-VI in canine nasal cavity. J Vet Med Sci 70:1037-1041

Sutherland GR, Baker E, Fernandez KE, Callen DF, Aldred P, Coghlan JP, Wright RD, Fernley RT (1989) The gene for human carbonic anhydrase VI(CA6) is on the tip of the short arm of chromosome 1. Cytogenet Cell Genet 50:149-150

Thatcher BJ, Doherty AE, Orvisky E, Martin BM, Henkin RI (1998) Gustin from human parotid saliva is carbonic anhydrase VI. Biochem Biophys Res Commun 250:635-641

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 28:511-515

Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25:1105-1111

White MR, Crouch E, van Eijk M, Hartshorn M, Pemberton L, Tornoe I, Holmskov U, Hartshorn KL (2005) Cooperative anti-influenza activities of respiratory innate immune proteins and neuraminidase inhibitor. Am J Physiol Lung Cell Mol Physiol 288:L831-40

Wu Z, Van Ryk D, Davis C, Abrams WR, Chaiken I, Magnani J, Malamud D (2003) Salivary agglutinin inhibits HIV type 1 infectivity through interaction with viral glycoprotein 120. AIDS Res Hum Retroviruses 19:201-209

Wuarin J, Schibler U (1990) Expression of the liver-enriched transcriptional activator protein DBP follows a stringent circadian rhythm. Cell 63:1257-1266

Yokoyama T (2004) Motor or sensor: a new aspect of primary cilia function. Anat Sci Int 79:47-54

FIGURE LEGENDS

Figure 1. VLAD gene ontology "biological process" annotation results for differentially transcribed genes in trachea of *Car6*$^{-/-}$ mice.

Figure 2. RNA-Seq analysis of *Car6* expression levels in lung related tissues. RNA-Seq data from multiple experiments was retrieved from the ArrayExpress database and the reads separately mapped to mouse *Car* genes. The FPKM values (median with range) for *Car6* in each tissue have been shown. *Car6* expression is negligible in lung but expressed in some samples of bronchoalveolar lavage and lung treated with influenza virus.

Figure 3. Immunohistochemical staining of CA VI protein in the submandibular gland, trachea and lung of WT and *Car6*$^{-/-}$ mice. Serous acinar cells show distinct positive staining in the submandibular gland of WT mouse (a), whereas no staining is detectable in the submandibular gland of CA VI deficient mouse (b). Cartilage cells show weak background staining in both WT (c) and CA VI deficient (d) mice trachea sections. Lung sections show similar background staining for WT (e) and CA VI deficient (f) mice. Original magnification, ×400.

ESM_1.pdf. VLAD gene ontology "cellular component" annotation results for differentially transcribed genes in trachea of *Car6*$^{-/-}$ mice.

ESM_2.pdf. VLAD gene ontology "molecular function" annotation results for differentially transcribed genes in trachea of *Car6*$^{-/-}$ mice.

ESM_3.pdf. The functional annotation categories for the genes regulated in the trachea by CA VI deficiency.

PUBLICATION III

Identification and characterization of a novel zebrafish (*Danio rerio*) pentraxin-carbonic anhydrase

Maarit Patrikainen, Martti Tolvanen, Ashok Aspatwar, Harlan Barker, Chaba Ortutay, Janne Jänis, Mikko Laitaoja, Vesa Hytönen, Latifeh Azizi, Prajwol Manandhar, Edit Jáger, Daniela Vullo, Sampo Kukkurainen, Mika Hilvo, Claudiu Supuran, Seppo Parkkila

PeerJ. 2017 Dec 7;5:e4128
https://doi.org/10.7717/peerj.4128. eCollection 2017

Identification and characterization of a novel zebrafish (*Danio rerio*) pentraxin–carbonic anhydrase

Maarit S. Patrikainen[1,*], Martti E.E. Tolvanen[2,*], Ashok Aspatwar[1,3,*], Harlan R. Barker[1], Csaba Ortutay[4], Janne Jänis[5], Mikko Laitaoja[5], Vesa P. Hytönen[1,3], Latifeh Azizi[1], Prajwol Manandhar[1,6], Edit Jáger[7], Daniela Vullo[8], Sampo Kukkurainen[1], Mika Hilvo[1,9], Claudiu T. Supuran[8] and Seppo Parkkila[1,3]

[1] Faculty of Medicine and Life Sciences, University of Tampere, Tampere, Finland
[2] Department of Future Technologies, University of Turku, Turku, Finland
[3] Fimlab Ltd., Tampere University Hospital, Tampere, Finland
[4] HiDucator Ltd., Kangasala, Finland
[5] Department of Chemistry, University of Eastern Finland, Joensuu, Finland
[6] Center for Molecular Dynamics Nepal, Kathmandu, Nepal
[7] Department of Epidemiology, Faculty of Health Sciences, Semmelweis University, Budapest, Hungary
[8] Dipartimento Neurofarba, Sezione di Scienze Farmaceutiche e Nutraceutiche, Università degli Studi di Firenze, Florence, Italy
[9] Zora Biosciences Ltd., Espoo, Finland
[*] These authors contributed equally to this work.

ABSTRACT

Background: Carbonic anhydrases (CAs) are ubiquitous, essential enzymes which catalyze the conversion of carbon dioxide and water to bicarbonate and H^+ ions. Vertebrate genomes generally contain gene loci for 15–21 different CA isoforms, three of which are enzymatically inactive. CA VI is the only secretory protein of the enzymatically active isoforms. We discovered that non-mammalian CA VI contains a C-terminal pentraxin (PTX) domain, a novel combination for both CAs and PTXs.
Methods: We isolated and sequenced zebrafish (*Danio rerio*) CA VI cDNA, complete with the sequence coding for the PTX domain, and produced the recombinant CA VI–PTX protein. Enzymatic activity and kinetic parameters were measured with a stopped-flow instrument. Mass spectrometry, analytical gel filtration and dynamic light scattering were used for biophysical characterization. Sequence analyses and Bayesian phylogenetics were used in generating hypotheses of protein structure and CA VI gene evolution. A CA VI–PTX antiserum was produced, and the expression of CA VI protein was studied by immunohistochemistry.
A knock-down zebrafish model was constructed, and larvae were observed up to five days post-fertilization (dpf). The expression of *ca6* mRNA was quantitated by qRT-PCR in different developmental times in morphant and wild-type larvae and in different adult fish tissues. Finally, the swimming behavior of the morphant fish was compared to that of wild-type fish.
Results: The recombinant enzyme has a very high carbonate dehydratase activity. Sequencing confirms a 530-residue protein identical to one of the predicted proteins in the Ensembl database (ensembl.org). The protein is pentameric in solution, as studied by gel filtration and light scattering, presumably joined by the PTX domains.

Submitted 30 August 2017
Accepted 14 November 2017
Published 7 December 2017

Corresponding author
Martti E.E. Tolvanen,
martti.tolvanen@utu.fi

Academic editor
Ana Rojas

Additional Information and Declarations can be found on page 32

DOI **10.7717/peerj.4128**

How to cite this article Patrikainen et al. (2017), Identification and characterization of a novel zebrafish (*Danio rerio*) pentraxin–carbonic anhydrase. **PeerJ 5:e4128; DOI 10.7717/peerj.4128**

Mass spectrometry confirms the predicted signal peptide cleavage and disulfides, and N-glycosylation in two of the four observed glycosylation motifs. Molecular modeling of the pentamer is consistent with the modifications observed in mass spectrometry. Phylogenetics and sequence analyses provide a consistent hypothesis of the evolutionary history of domains associated with CA VI in mammals and non-mammals. Briefly, the evidence suggests that ancestral CA VI was a transmembrane protein, the exon coding for the cytoplasmic domain was replaced by one coding for PTX domain, and finally, in the therian lineage, the PTX-coding exon was lost. We knocked down CA VI expression in zebrafish embryos with antisense morpholino oligonucleotides, resulting in phenotype features of decreased buoyancy and swim bladder deflation in 4 dpf larvae.
Discussion: These findings provide novel insights into the evolution, structure, and function of this unique CA form.

Subjects Biochemistry, Bioinformatics, Evolutionary Studies
Keywords Phylogeny, Protein modeling, Pentraxin, Zebrafish, Carbonic anhydrase VI, Carbonic anhydrase, Mass spectrometry, Innate immunity, Knockdown

INTRODUCTION

Carbonic anhydrase VI (CA VI) is the only secretory CA enzyme in mammals. In its very first reporting, *Henkin et al. (1975)* described a novel protein, gustin, from human saliva, which was later shown to be CA VI (*Thatcher et al., 1998*). This protein was first described as a CA enzyme by *Fernley, Wright & Coghlan (1979)*, who identified a novel high molecular weight (MW) form of CA in the sheep parotid gland and saliva. The first immunohistochemical studies on human CA VI indicated that it is highly expressed in the serous acinar cells of the parotid and submandibular glands (*Parkkila et al., 1990*). It is one of the major protein constituents of human saliva (*Parkkila et al., 1993*), and also found in human and rat milk (*Karhumaa et al., 2001*).

The physiological role of CA VI has remained unclear, even though it was discovered three decades ago. Henkin's group linked gustin (CA VI) to the regulation of taste function (*Shatzman & Henkin, 1981*). Expression of CA VI in the von Ebner's glands implicate CA VI in the paracrine modulation of taste function and TRC apoptosis (*Leinonen et al., 2001*). Various studies have later shown a link between bitter taste perception and CA VI (*Melis et al., 2013*; *Patrikainen et al., 2014*). Two studies have shown a link between CA VI and immunological function in mouse and human. First, *Car6*$^{-/-}$ mice have a greater number of lymphoid follicles in the small intestinal Peyer's patches, suggesting an immunological phenotype (*Pan et al., 2011*). Second, the analysis of gene expression in the trachea and lung of *Car6*$^{-/-}$ mice showed alterations in biological processes such as antigen transfer to mucosal-associated lymphoid tissue (*Patrikainen et al., 2016*).

Innate immune systems, based on pattern recognition, exist in some form in all metazoan organisms (*Medzhitov, 2007*). The pattern-recognition molecules (PRMs) recognize conserved structures on the surface of pathogens and activate the innate immune response. Pentraxins (PTXs) are a superfamily of fluid phase PRMs conserved in

evolution and characterized by a cyclic multimeric structure with a regulatory role in inflammation (*Bottazzi et al., 2016*). They contain a characteristic ~200-residue-long domain at their C-terminus. Based on their primary subunit structures PTXs are divided into short PTXs and long PTXs. Short PTXs are classically represented by C-reactive protein (a.k.a. CRP, pentraxin-1, PTX-1) and serum amyloid P (a.k.a. APCS, SAP, pentraxin-2, PTX-2), whereas long PTXs comprise pentraxin-3 (PTX3), neuronal PTXs, and others (*Garlanda et al., 2005*).

We noted the presence of an additional PTX domain in some non-mammalian *CA6* gene predictions in 2007, but did not follow up this observation at that time. More recently, with more non-mammalian genomes available, we realized that the PTX domain is present in non-mammalian CA VI too consistently to be an annotation artifact, which inspired this study. We used zebrafish (*Danio rerio*) as a vertebrate model organism for functional and structural characterization of the PTX-associated CA VI.

MATERIALS AND METHODS

Sequence conservation

In order to compare conservation in the CA and PTX domains of CA VI–PTX proteins, non-mammalian CA VI sequences were retrieved from NCBI (*NCBI Resource Coordinators, 2016*) nr protein database as of December 5, 2015, using BLASTP (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) (*Altschul et al., 1990*), with human CA VI (ENSP00000366662 from Ensembl (*Flicek et al., 2012*) as query sequence, taxonomically filtered for non-mammalian vertebrates. Full-length or nearly full-length CA VI–PTX sequences were seen at extremely low e values, not higher than 2×10^{-80}, indicating very high similarity; CA VI sequence fragments were seen at e values from 10^{-79} to 10^{-71}; and the remaining matches, at e values of 10^{-68} and higher, were annotated as other CA isoforms and did not contain a PTX domain. Sequences with an e value of 2×10^{-80} or lower were taken for further quality control. We discarded sequences shorter than 485 residues and any with non-specific "X" characters. Furthermore, we rejected sequences with unaligned, unique insertions of at least 20 residues at exon boundaries, which we assume to be introns mispredicted as coding sequence. Likewise, sequences containing gaps in the alignment between exon boundaries were interpreted to miss data for internal exons and were discarded. Thus, all sequences which were incomplete in the CA domain were discarded, but sequences devoid of the signal peptide region were still kept. The final sequence set contained 78 sequences from 75 species, (sequence accession numbers shown in Fig. S1 and in Data S1). After inspection of the multiple sequence alignment, four sequences were edited for a more plausible initiation site (Table 1) deemed to be at the conserved M at the start of the signal peptide region. All sequences had complete PTX domains. Sequences were aligned with Clustal Omega (*Sievers et al., 2011*). In order to calculate conserved positions in each domain, the CA domain was defined to correspond to residues 24–280 in zebrafish CA VI (UniProt annotation in E9QB97_DANRE), and the PTX domain was defined as residues 317–518 (InterProScan at http://www.ebi.ac.uk/interpro/sequence-search (*Jones et al., 2014*), match to profile SM00159, PTX).

Table 1 Suggested corrections for predicted translation start sites.

RefSeq ID	Name	Organism	Number of N-terminal residues removed
XP_010721064.1	PREDICTED: carbonic anhydrase 6	*Meleagris gallopavo*	110
XP_005057921.1	PREDICTED: carbonic anhydrase 6	*Ficedula albicollis*	17
XP_002187446.1	PREDICTED: carbonic anhydrase 6	*Taeniopygia guttata*	6
XP_005143337.1	PREDICTED: carbonic anhydrase 6	*Melopsittacus undulatus*	31

Note:
The following database entries have N-terminal extensions which we assume mispredicted. We have shortened these sequences to start at a conserved initiating Met residues for use in this study.

Phylogenetic analyses

For the tree in Fig. 1, cDNA sequences and their protein translations were collected from the Ensembl database (release 67) for CAs 6, 9, 12, and 14, from selected species. Protein sequences were aligned with Clustal Omega. Codon aligned cDNA sequences were produced in the PAL2NAL web server v. 14 (http://www.bork.embl.de/pal2nal/) (*Suyama, Torrents & Bork, 2006*) using the protein alignment as a guide (protein alignment: Data S2; final codon alignment: Data S3). For the tree in Fig. S2, a second alignment was similarly made using catalytic domains of CA VI sequences only (protein alignment: Data S4; final codon alignment: Data S5). For the tree in Fig. 2, we made a third alignment of CA VI-associated PTX domains from selected species and human PTXs (codon aligned sequences: Data S6). The resulting codon (DNA) alignments and the program MrBayes v 3.2 (*Ronquist et al., 2012*) were used to estimate the phylogeny of the sequences by Bayesian inference. Bayesian estimation was run for at least 10,000 generations, with flat a priori distribution of base frequencies, substitution rates, proportion of invariable sites, and gamma shape parameter. The 50% majority rule consensus trees were saved and visualized using the APE R package (*Paradis, Claude & Strimmer, 2004*).

Run lengths, relevant parameters at the end of run, and rooting of the trees were as follows. For the first tree, the average standard deviation of split frequencies after 10,000 generations was 5.2×10^{-2} when the analysis was stopped. The arithmetic mean of the estimated marginal likelihoods for runs sampled was −17175.07. *Drosophila melanogaster* CAH1 sequence was used as an outgroup to root the tree. For the second tree, the average standard deviation of split frequencies after 20,000 generations was 1.2×10^{-1} when the analysis was stopped. The arithmetic mean of the estimated marginal likelihoods for runs sampled was −10596.5. Fish sequences were used as an outgroup to root the consensus tree. Branching points with lower than 50% consensus in the mammal branch are collapsed. Finally, for the third tree, the average standard deviation of split frequencies after 10,000 generations was 8.1×10^{-2} when the analysis was stopped. The arithmetic mean of the estimated marginal likelihoods for runs sampled was −10319.1.

BlastN search in platypus genome

In order to see if the orphan fragment Contig22468 of platypus genome (which contains the exon coding for a "CA VI-type" PTX domain) would have been somehow

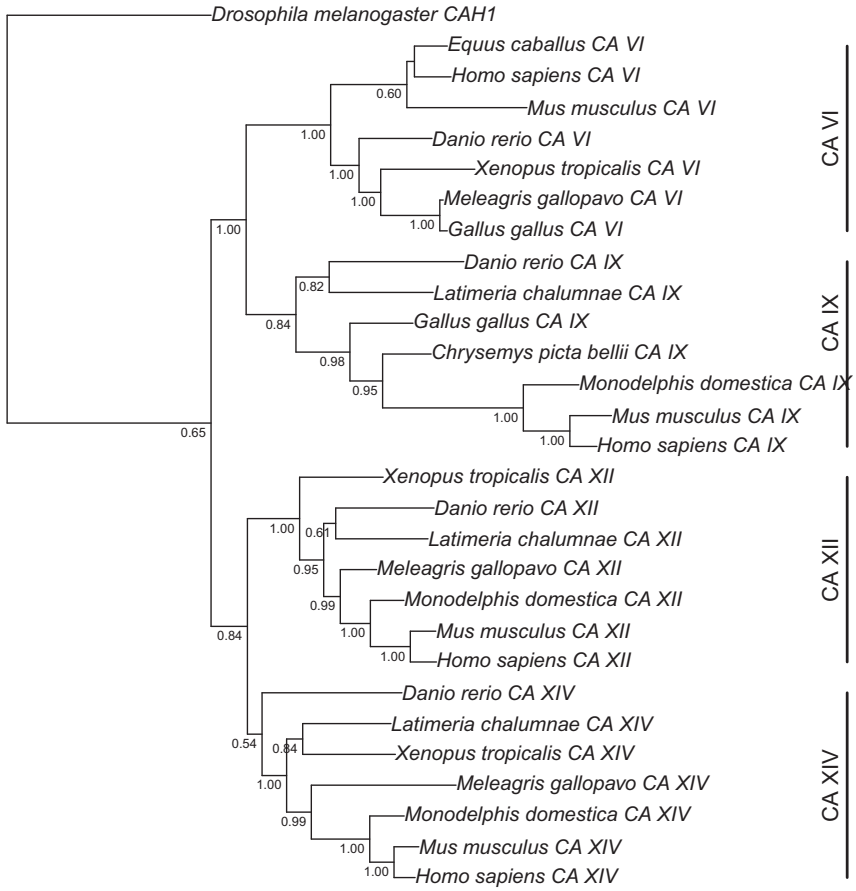

Figure 1 Bayesian phylogenetic tree of CA VI, CA IX, CA XII, and CA XIV. Analysis of protein alignment guided DNA alignments as detailed in "Materials and Methods." Sidebars indicate the groups of isoforms. The CA VI subtree with more species is shown in Fig. S2.

Full-size DOI: 10.7717/peerj.4128/fig-1

missed in the genome assembly, we performed a BlastN search in Ensembl (http://www.ensembl.org/Homo_sapiens/Tools/Blast?db=core). BlastN was run against the platypus genome with the full 11,311 nt sequence of supercontig:OANA5:Contig22468 as query sequence.

Exon length comparisons

Exon data was retrieved from Ensembl. Lengths of the exons that follow those coding for the CA domain were noted for Ensembl transcripts for human *CA6* (ENST00000377443), *CA9* (ENST00000378357), *CA12* (ENST00000178638 and ENST00000344366), and *CA14* (ENST00000369111), and zebrafish *ca6* (ENSDART00000132733). Similarly,

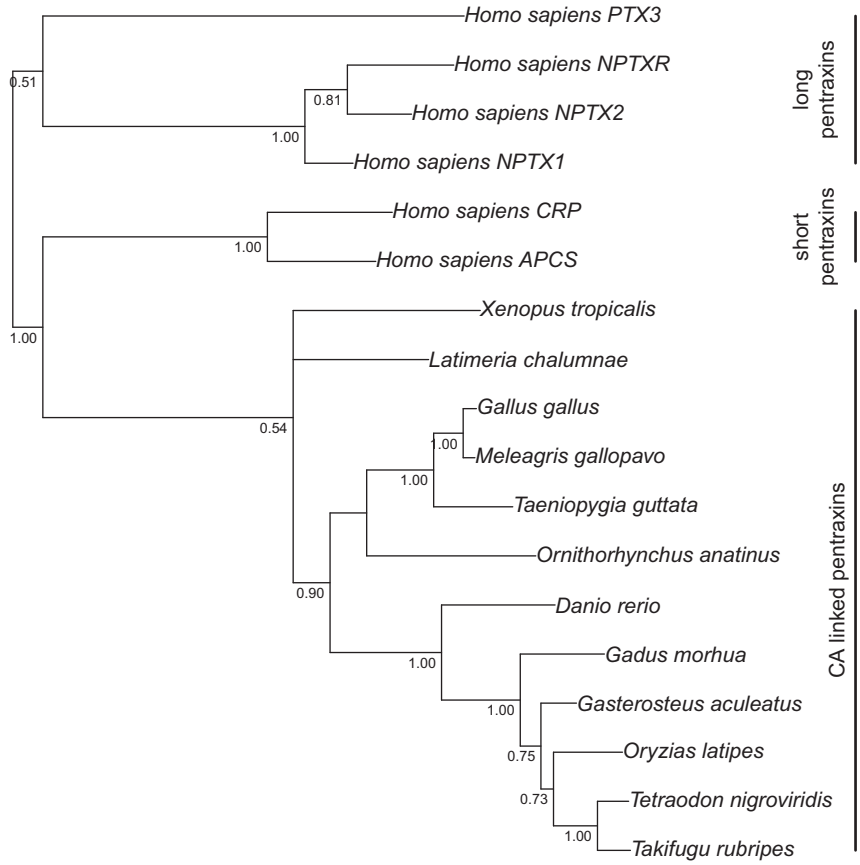

Figure 2 Bayesian phylogenetic tree of pentraxin domains. Analysis of protein alignment guided DNA alignments as detailed in "Materials and Methods." Sidebars indicate PTX domains extracted from non-mammalian CA VI sequences (bottom) and groups of human pentraxins.

Full-size DOI: 10.7717/peerj.4128/fig-2

lengths of the exons preceding the PTX domain exon were noted for Ensembl transcripts of human *CRP* (ENST00000255030), *APCS* (SAP, ENST00000255040).

Amphipathic helix prediction

A study of the region between the CA and PTX domains in the alignment of CA VI protein sequences showed little conservation except for five sites with hydrophobic residues spaced three or four residues apart, with mostly polar residues between them, suggestive of an amphipathic alpha helix. The subsequences from 287 to 303 and from 293 to 309 for human and zebrafish CA VI, respectively, were visualized as helical wheel diagrams, or end projections of a hypothetical alpha helix of 17 residues, using the PepWheel program of the EMBOSS suite (http://www.bioinformatics.nl/cgi-bin/emboss/help/pepwheel) (*Rice, Longden & Bleasby, 2000*).

Construction of recombinant baculoviruses

The 1,593 bp zebrafish *ca6* sequence encoding the full-length, PTX-containing CA VI polypeptide (CA VI–PTX) was amplified by PCR using the forward primer 5′-ATGGAGCAGCTGACTCTAGTC-3′ and reverse primer 5′-TTTCTCTGTTTCTCTATTATTATTAT-3′. PCR conditions consisted of an initial denaturation step at 98 °C for 30 s followed by 35 cycles at 98 °C for 10 s (denaturation), 55 °C for 30 s (annealing), and 72 °C for 25 s (elongation). The final extension step was carried out at 72 °C for 5 min. The expression construct was optimized for protein production in *Spodoptera frugiperda* insect cells (Sf9) by inserting into second round PCR primers restriction sites for *Bam*HI and *Sal*I plus sequences coding for C-terminal histidine tag for protein purification and a thrombin cleavage site for tag removal. The second round of PCR was carried out using the protocol described above, except that the temperature for annealing was 62 °C, and final extension step was carried out at 74 °C for 7 min. The baculoviral genomes encoding CA VI recombinant proteins were generated according to the Bac-to-Bac Baculovirus Expression System instructions (Invitrogen, Camarillo, CA, USA). The recombinant protein insert was sequenced using ABI PRISM BigDye® Terminator v3.1 Cycle Sequencing kit (Applied Biosystems, Inc., Foster City, CA, USA.) and pFASTBac primers (forward: 5′-AATGATAACCATCTGGCA-3′ and reverse: 5′-GGTATGGCTGATTATGAT-3′) in order to obtain the full-length insert sequence. The PCR conditions consisted of 35 cycles at 96 °C for 10 s (denaturation), 50 °C for 5 s (annealing), and 55 °C for 4 min (elongation) with final extension at 37 °C for 5 min.

Production and purification of recombinant CA VI–PTX

The Sf9 insect cells (Invitrogen) were maintained in HyQ SFX-Insect serum-free cell culture medium (HyClone, Logan, UT, USA). The cells were centrifuged (2,000×*g*, at 20 °C, for 5 min) 72 h after infection, and the medium was collected. Purification was performed with the Probond® Purification System (Invitrogen) under native binding conditions with wash and elution buffers made according to the manufacturer's instructions. Purity of the protein was checked and the MW of the recombinant protein was determined by running a 10% SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) under reducing conditions. The size of the protein was determined using Precision Plus Protein™ Standards Dual Color (Bio-Rad Laboratories, Inc., Hercules, CA, USA) and MW marker and bands were visualized using the Colloidal Blue Staining Kit™ (Invitrogen).

CA activity/inhibition assay

An Applied Photophysics stopped-flow instrument was used for assaying the CA catalyzed CO_2 hydration activity (*Khalifah, 1971*). The method was exactly as described previously (*Berrino et al., 2017*) except that the inhibitor dilutions were done up to 0.5 nM.

Light scattering experiments

Molecular weight determination of zebrafish CA VI–PTX was performed using a Malvern Zetasizer μV instrument (Malvern Instruments Ltd., Worcestershire, UK) running

static light scattering (SLS) and dynamic light scattering (DLS) methods. Analysis was performed using a liquid chromatography instrument (CBM-20A, Shimadzu Corporation, Kyoto, Japan) equipped with autosampler (SIL-20A), UV–VIS (SPD-20A) and fluorescence detector (RF-20Axs). Data were processed using Lab Solution Version 5.51 (Shimadzu Corporation) and OmniSec 4.7 (Malvern Instruments Ltd., Worcestershire, UK) softwares. A sample of the protein (50 μg) was injected on a Superdex 200 5/150 column (GE Healthcare, Uppsala, Sweden) equilibrated with 50 mM NaH_3PO_4, 500 mM NaCl pH 8 buffer. Runs were performed with flow rate of 0.1 ml/min at 20 °C using a thermostated cabin. The MW of the zebrafish CA VI–PTX was determined either by using a standard curve based on MW standard proteins (SEC analysis; CA 29 kDa, alcohol dehydrogenase 150 kDa, β-amylase 200 kDa, BSA 66 kDa, Sigma-Aldrich, Inc., St. Louis, MO, USA) or by calibrating the light scattering detector using the monomeric peak of BSA and light-scattering intensity (SLS).

Sample preparation for mass spectrometry

Prior to the mass spectrometric measurements, the CA VI–PTX sample was buffer-exchanged to 10 mM ammonium acetate (pH 7.5) buffer using Sephadex G-25 M (PD-10) desalting columns (GE Healthcare, Gillingham, UK). Ten 1 ml fractions were collected, and the fractions containing protein were concentrated using Amicon Ultra (5 kDa cut-off) centrifugal filter devices (Merck Millipore, Darmstadt, Germany). Finally, protein concentrations were determined by UV-absorbance at 280 nm, using a sequence-derived extinction coefficient 99,155 M^{-1} cm^{-1}, calculated by ProtParam at http://web.expasy.org/protparam/ (*Gasteiger et al., 2003*). Intact protein mass analysis was performed by diluting the sample to the desired protein concentration with acetonitrile (MeCN), containing 1% of acetic acid (HOAc). Alternatively, CA VI–PTX was digested with trypsin. Briefly, an aliquot of the CA VI–PTX sample was mixed with a sequencing grade modified trypsin (Promega, Madison, WI, USA) (3 mg/ml in water) to obtain a 1:20 (w/w) protease-to-protein ratio. The digest sample was incubated at 37 °C for 1 h, and subsequently diluted to approximately 10 μM with MeCN containing 1% HOAc.

Mass measurements and data analysis

All experiments were performed on a 12-T Bruker Solarix-XR FT-ICR mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany), equipped with an Apollo-II electrospray ionization (ESI) source and a dynamically harmonized ParaCell ICR-cell. All protein/peptide samples were directly infused into the ESI source at a flow rate of 1.5 μL/min. The ESI-generated ions were externally accumulated in the hexapole collision cell for 1 s, and transferred to the ICR cell for trapping, excitation and detection. For each mass spectrum, a total of 300 time-domain transients (1 MWord each) were co-added, and zero-filled once to obtain final 2 MWord broadband data. For collision-induced dissociation tandem mass spectrometry (CID-MS/MS) experiments, the precursor ions of interest were mass-selected in a quadrupole and fragmented in the collision cell by increasing the collision voltage to the appropriate value. The mass spectra were externally calibrated with ESI-L Low concentration tuning mix (Part no G1969-85000; Agilent Technologies,

Santa Clara, CA, USA). The instrument was operated and the data were acquired by using Bruker ftmsControl 2.0 software. The mass spectra were subsequently transferred to Bruker DataAnalysis 4.4 software for further processing. Spectral de-isotoping and charge-state deconvolution (to obtain monoisotopic peptide masses) was accomplished with a Bruker SNAP2 peak-picking module. The obtained mass lists were then uploaded to GPMAW 10.0 software (Lighthouse Data, Odense, Denmark) for tryptic peptide identification. Only specific tryptic peptides were searched with a maximum mass error of 5 ppm. The glycosylation sites in CA VI–PTX were identified by incorporating typical high-mannose/complex glycans with a various number of residues into the four putative N-glycosylation sites and searched against the obtained mass lists.

Homology modeling of zebrafish CA VI

We built a 3D model of zebrafish CA VI starting from PDB 3FE4, human CA VI (*Pilka et al., 2012*); and 4AVS, human SAP component (*Kolstoe et al., 2014*), as templates for the CA domain and PTX domain, respectively. The CA template includes residues 32–280 of human CA VI, missing 14 residues in the N-terminus of the mature protein, and 28 residues in the C-terminus. Briefly, the predicted amphipathic helix (APH) region of human CA VI (287–303) and four additional residues (283–286) were modeled as an alpha helix, and the helix was subsequently docked to the C-terminal face of 3FE4. This extended model of human CA VI was used as a template in homology modeling the CA domain plus APH of zebrafish CA VI. The model of the PTX domain of zebrafish CA VI was docked to the model of CA+APH domains. Finally, a pentameric model of CA VI–PTX was created by superimposing five copies of the monomer model over each monomer in the pentameric SAP structure (PDB 4AVS).

The C-terminal alpha helix was generated automatically ab initio for the region predicted to form an APH when the full sequence of human CA VI was given as a modeling target to I-TASSER 4.0 (*Roy, Kucukural & Zhang, 2010*). The helix was separated from the model and subsequently docked to 3FE4, using the HADDOCK 2.1 server (http://haddock.science.uu.nl/) (*de Vries et al., 2007*). A list of potential interface residues in 3FE4 were predicted through CPORT (http://milou.science.uu.nl/services/CPORT/) (*de Vries & Bonvin, 2011*), and only those lying on the C-terminal face of 3FE4 were set as interacting residues in docking. For the APH, the residues located on the hydrophobic side were chosen as interacting residues. The resulting model from this docking step was used as a template to model the full CA domain of zebrafish CA VI at the MODELLER server (*Webb & Sali, 2016*) using UCSF Chimera v. 1.10 (*Pettersen et al., 2004*) as the interface program. The PTX domain of zebrafish CA VI (residues 317–530) was also modeled by MODELLER, with 4AVS as template. The two partial models of zebrafish CA VI were joined by docking with HADDOCK, again predicting interacting residues with CPORT. Most of the predicted interacting residues were located near the C-terminus of the CA domain and near the N-terminus of the PTX domain, and the residues in these regions were chosen as active residues in docking. The structural superimpositions of the PTX domains of the CA VI–PTX monomer on the SAP pentamer (4AVS) were carried out with the MatchMaker tool in UCSF Chimera to generate the final pentamer model.

Zebrafish maintenance and ethical permissions

Wild-type zebrafish of the AB strain were maintained at 28.5 °C under standard conditions (*Westerfield, 2007*). We express the embryonic ages in hours post-fertilization (hpf) and days post-fertilization (dpf). Embryos/larvae were collected from the breeder tanks with a sieve and rinsed with embryonic medium (Sarsted, Nümbrecht, Germany) into Petri dishes. Embryos/larvae were kept in Petri dishes in embryonic medium supplemented with 1-phenyl-2-thiourea (Sigma-Aldrich) at 28.5 °C until they were used in experiments. The maximum number of larvae on a 9 cm diameter Petri dish was 50. Embryonic medium contained 5 mM NaCl, 0.17 mM KCl, 0.33 mM $CaCl_2$, 0.33 mM $MgSO_4$ and 10–5% methylene blue (Sigma-Aldrich). Zebrafish housing and care in the Zebrafish facility of the University of Tampere have been approved by the National Animal Experiment Board of Finland, administered through the Provincial Government of Western Finland, Province Social and Health Department Tampere Regional Service Unit (permit # LSLH-2007-7254/Ym-23). Using five-day old zebrafish as a model organism requires no specific ethical permission, neither does studying tissues collected from euthanized adult fish.

Morpholino injections of zebrafish embryos

Knockdown of *ca6* was carried out using two different antisense morpholino oligonucleotides (MOs) (GeneTools LLC, Philomath, OR, USA): one translation-blocking (MO1 5′-CTGCCTGTGCTCTGAACTGTTTCTC-3′) and the other splicing-blocking, to target intron–exon boundary before exon 9 (MO2 5′-GCTTGCCTTGAGAAGGAAA GATCAT). The random control (RC) MOs (5′-CCTCTTACCTCAGTTACAATTTATA-3′) were used as control MOs. The supplied MOs were re-suspended in sterile water at 1 mM stock concentration. Immediately prior to injection, *ca6*-MOs were diluted to the intended concentration of 125 μM. In order to monitor injection efficiency, 0.2% dextran rhodamine B and 0.1% phenol red (final concentrations; Sigma, Poole, UK) were included in the solution, and the final KCl concentration was adjusted to 1 M. About 1 nl of antisense MO solution was injected into the yolk of approximately 500 one- to two-cell stage embryos, without randomization. The MO-injected embryos were screened for the presence of fluorescence after 24 h to select the true *ca6* morphants using Lumar V1.1 fluorescence stereomicroscope (Carl Zeiss MicroImaging GmbH, Göttingen, Germany) and AxioVision software version 4.9. The non-fluorescent embryos were eliminated.

Microscopy and live image analysis of zebrafish phenotypes

Gross phenotypic appearance was analyzed by light-field microscopy. For each experiment, typically 10–20 *ca6*-MO-injected larvae were screened with a similar number of matched controls. Larvae were first euthanized using 0.05% tricaine (Sigma-Aldrich) in embryo medium and embedded in 17% high MW methyl cellulose in 15 × 30 mm transparent polypropylene Petri dish for taking images of the developing embryos/larvae from 1 to 5 dpf using Zeiss Stereo Microscope (Carl Zeiss MicroImaging GmbH, Göttingen, Germany) with NeoLumar S 1.5× Objective (Carl Zeiss MicroImaging GmbH,

Göttingen, Germany). The images were analyzed with AxioVision software version 4.9. and scale bars were inserted. Images were cropped and assembled into composite images.

Isolation of total RNA and synthesis of cDNA

Total RNA was isolated at different stages of development, from 0 to 168 hpf whole embryos/larvae, and from different organs of the adult zebrafish. Total RNA was isolated from 30 μg samples using the RNeasy® Mini kit (Qiagen, Hilden, Germany) by following the manufacturer's instructions. The concentration and purity of total RNA were determined using a Nanodrop UV/VIS Spectrophotometer at 260 and 280 nm. Reverse transcriptase PCR was performed using 0.1–5 μg of total RNA to synthesize the first strand cDNA using First Strand cDNA Synthesis kit (High-Capacity cDNA Reverse Transcription Kits; Applied Biosystems, Foster City, CA, USA) with random primers and M-MuLV reverse transcriptase according to the protocol recommended by the manufacturer.

Quantitative real-time PCR

Quantitative real-time PCR (qRT-PCR) primers were designed based on the complete cDNA sequence taken from Ensembl (ENSDART00000057097), using the program Primer Express® Software v2.0 (Applied Biosystems) (forward primer 5′-CAAACATTTAT TTGCCAGCACTCC-3′ and reverse primer 5′-TATGTCCAATAATCTCCATCTACTCC-3′). qRT-PCR was performed using the SYBR Green PCR Master Mix Kit in an ABI PRISM 7000 Detection System™ according to the manufacturer's instructions (Applied Biosystems). The PCR conditions consisted of an initial denaturation step at 95 °C for 10 min followed by 40 cycles at 95 °C for 15 s (denaturation) and 60 °C for 1 min (elongation). The data were analyzed using the ABI PRISM 7000 SDS™ software (Applied Biosystems). Every PCR was performed in a total reaction volume of 15 μl containing 2 μl of first strand cDNA (20 ng cDNA), 1 × Power SYBR green PCR Master Mix™ (Applied Biosystems, Foster City, CA, USA), and 0.5 μM of each primer. We performed these experiments in duplicate and with sample duplicates. The results of *ca6* gene expression were normalized using zebrafish housekeeping gene *gapdh* as internal control. The final results are given as relative expression values, calculated according to the Pfaffl's equation (*Pfaffl, 2001*).

Preparation of zebrafish tissues

The adult zebrafish were euthanized by keeping them in 1% tricaine on ice for more than 10 min followed by decapitation. Different organs were harvested under the microscope and immediately transferred them to 1.5 ml microcentrifuge tube containing RNAlater® (Ambion, Austin, TX, USA) and were stored at −20 °C until further analysis. Simultaneously tissues for immunohistochemical analysis were harvested and immediately fixed with 4% PFA for 24 h at 4 °C. Tissues were transferred to 20% sucrose in PBS and stored at 4 °C until embedding them in Tissue-Tek® O.C.T.™ Compound (Sakura Finetek Europe B.V., Alphen aan den Rijn, The Netherlands). Embedded tissue samples were stored at −20 °C until further analysis.

Antibody testing

Antibody against zebrafish CA VI–PTX was manufactured by Innovagen AB (Innovagen AB, Lund, Sweden) according to their standard immunization schedule, with boosters at 14, 28, 49, and 70 days. Pre-immune serum and three samples of polyclonal antiserum were tested using dot blotting. Bio-Dot® Microfiltration Apparatus (BioRad Laboratories, Inc., Hercules, CA, USA) was used to attach 500 ng of produced and purified native zebrafish CA VI–PTX protein to PROTRAN® nitrocellulose (NC) transfer membrane (Schleicher & Schuell GmbH, Dassel, Germany) according to manufacturer's instructions. Prior to staining, non-specific binding of the primary antibody was prevented using diluted colostrum (1:10 in Tris-Buffered Saline with Tween 20 [TBST]) as a blocking agent for 30 min. Pre-immune serum, bleed 1 (day 41), bleed 2 (day 62), and bleed 3 (day 83) of polyclonal rabbit anti-zebrafish CAVI-PTX (Innovagen AB, Lund, Sweden), diluted 1:100 in TBST, were added to NC strips which were incubated at room temperature for 1 h. Donkey anti-rabbit IgG, horseradish peroxidase linked whole antibody (Amersham Biosciences, GE Healthcare Life Sciences, Little Chalfont, UK) diluted 1:25,000 in TBST was used as secondary antibody. Washing steps were carried out using TBST. Staining was carried out using ImmPACTTM DAB Peroxidase Substrate Kit (Vector Laboratories, Inc., Burlingame, CA, USA). The testing showed that bleed 1 and bleed 2 have a strong reactivity against zebrafish CA VI–PTX (Figs. S3B and S3C). Antiserum of bleed 2 was used in further experiments.

Immunohistochemistry of zebrafish tissues

The Tissue-Tek® O.C.T.™ Compound-embedded samples were cut into 10 μm sections using cryotome and prior to staining, the sections were attached to the glass slide by incubating at 37 °C overnight. Staining procedure of tissue samples was carried out as described above. Alexa Fluor® goat anti-rabbit IgG 1:1,000 (Life Technologies, Carlsbad, CA, USA) was used as a secondary antibody, and sections were mounted with Vectashield Hard Set Mounting Medium with nuclear dye DAPI (4′,6-diamidino-2-phenylindole, Vector Laboratories Inc., Burlingame, CA, USA). The sections were photographed using Zeiss LSM780 Laser Scanning Confocal Microscope with Zeiss Cell Observer.Z1 microscope, Plan-Apochromat 40×/1.4 (oil) objective, with pulsed diode laser 405 nm and multiline Argon laser: 488 nm, and Quasar spectral GaAsP PMT array detector (Carl Zeiss Microscopy GmbH, Göttingen, Germany). Images were analysed with Zeiss ZEN2Lite.

Behavioral analysis of 4 and 5 dpf *ca6* knockdown zebrafish larvae

Larvae were tested for behavioral consequences due to *ca6* knockdown by measuring swimming pattern at 4 and 5 dpf. The *ca6* knockdown larvae and two controls, namely uninjected wild-type and RC MO-injected, were raised in embryo medium. Larvae (approximately 10/flask) were placed in a 23 × 43 × 45 mm TC Flask T25 (Sarstedt AG & Co., Nümbrecht, Germany) containing 40 ml embryo medium at 3 dpf and allowed to acclimate to the flask for 24 h at 28.5 °C standard conditions. At 4 and 5 dpf their swimming patterns were observed by a 1 min video recording, with a printed 1 × 1 cm

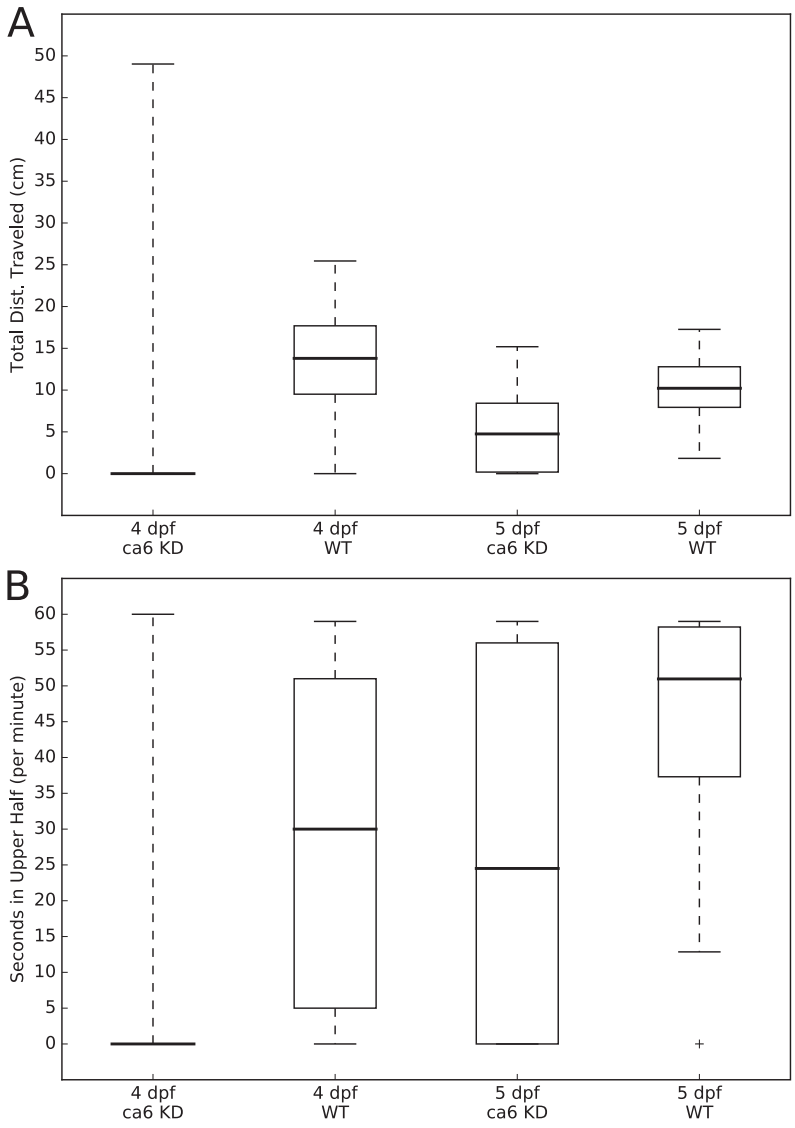

Figure 3 Zebrafish wild-type and *ca6* knockdown movement analysis. Boxplots from video analysis of 1 min of swimming of zebrafish larvae. KD, knockdown; WT, wild-type. Same data used for both (A and B). Statistics of both analyses are given in Table 3. (A) Total distances traveled. (B) Time spent in the upper half of the tank (seconds, out of 60 s). Full-size DOI: 10.7717/peerj.4128/fig-3

grid behind the flask. In total, the movement patterns of 284 zebrafish were recorded and measured: 41 of 4 dpf WT, 130 of 4 dpf KD, 32 of 5 dpf WT, and 81 of 5 dpf KD. Sample sizes of at least 30 per group were chosen a priori because normality of distributions could not be assumed.

The movements of all of the larvae were analyzed using the MtrackJ plugin (*Meijering, Dzyubachyk & Smal, 2012*) within the ImageJ program (*Schneider, Rasband & Eliceiri, 2012*). Tracking and recording of fish movements and analysis of movement data were assigned to two separate researchers to avoid biasing the analysis. Distances traveled (cm per 1 min, for Fig. 3A) and time spent in the upper half of the tank (seconds, out of 60 s, for Fig. 3B) were calculated for each fish, compiled by group, and presented as boxplots using the Matplotlib (*Hunter, 2007*) Python library. Statistical testing of similarity between each group, using the Kolmogorov–Smirnov two sample test, was performed using the Stats module of the SciPy Python library (*van der Walt, Colbert & Varoquaux, 2011*). The two-sample Kolmogorov–Smirnov test was chosen because it makes no assumption about the distribution of data.

RESULTS

Non-mammalian CA VI contains an additional PTX domain

We retrieved 78 CA VI protein sequences from 75 non-mammalian species in NCBI GenPept, all of which have the C-terminal PTX domain. The PTX domain in CA VI is less conserved than the CA domain. The multiple sequence alignment of the 78 CA VI sequences (Fig. S1) shows that there are 83 perfectly conserved amino acids within the catalytic domain (within MSA columns 30–288), whereas only 19 amino acids in the PTX domain are perfectly conserved (within MSA columns 355–566). The region between the CA and PTX domains consists of a moderately conserved and gapless region (MSA columns 300–320) flanked by two highly variable regions of flexible length (MSA columns 292–297 and 326–344). The presence of CA and PTX domains in non-mammalian CA VI sequences has also been documented in the Pfam database since many years (*Finn et al., 2016*), for example in http://pfam.xfam.org/protein/E9QB97_DANRE.

Figure 1 presents the phylogenetic tree of CAs VI, IX, XII, and XIV, clearly showing that the longer, non-mammalian isoforms (with a PTX domain) are orthologs of mammalian CA VI. The pairwise arrangement of VI/IX vs. XII/XIV is the same as in previous phylogenetic work (*Hewett-Emmett, 2000*), suggesting that these four CA isozymes descend from one common ancestor. Figure 2 shows a phylogenetic tree of all human PTXs and selected CA-linked PTX domains, which indicates that the novel PTX domains would be most closely related to the short PTXs, CRP and APCS or SAP.

Platypus (*Ornithorhynchus anatinus*) is probably an exception in the pattern of mammals not having a PTX domain associated with CA VI

A genomic fragment not assigned to any chromosome (Contig22468 in assembly WUGSC 5.0.1/ornAna1) contains an exon which codes for a PTX domain unlike any that we find in other mammalian species, and most similar to CA VI-linked PTX domains in non-mammalian species. The phylogenetic tree in Fig. 2 demonstrates that this platypus PTX sequence is orthologous with the PTX sequences associated with CA VI in non-mammalian species. What is more, a BLASTN search of Contig22468 against the platypus genome showed that it partially matches a region in chromosome 5 right after

the *CA6* locus. More specifically, the first 703 bp of Contig22468 match the last 703 bases (99.86% identity, a single mismatch) of Contig3933.5.

The adjacent location of Contig3933.5 to Contig3933.4, the fragment containing the exons coding for the *CA6* ortholog (ENSOANG00000013215), would put the exon coding for the PTX domain in the correct location and orientation to be part of the platypus *CA6* gene if Contig22468 were placed in this position. Therefore, we tentatively label this PTX domain as "CA-linked" and suggest that Contig22468 would be more correctly mapped starting from Chr5:18954728 in platypus genome assembly OANA5. With this evidence, we also suggest that CA VI in platypus contains a PTX domain, and consequently, that the loss of PTX domain occurred after the separation of monotreme and therian lineages in mammals.

One further phylogenetic tree was made based on CA domain sequences, showing that phylogeny of CA VI follows the expected vertebrate phylogeny, with platypus placed outside of marsupials and placental mammals (Fig. S2).

Exon lengths suggest that the region after the CA domain in CA VI descends from the transmembrane helix of the ancestral form

Mammalian CA VI proteins contain an additional C-terminal region of at least 25 residues, which is dissimilar to anything in other vertebrate CA isoforms and of unknown structure. Non-mammalian CA VI contains a sequence homologous to this extension as a spacer region between the CA and PTX domains. In order to investigate the most likely origin of the spacer region, we compared the exon lengths in *CA6* and the most closely related CA genes (*CA9*, *CA12*, and *CA14*) and short PTXs. The length of the exon coding for the spacer between CA and PTX domains in zebrafish *ca6* is 84 bp, and the coding sequence of the homologous exon in human *CA6* is 83 bp. The exons coding for the region containing the transmembrane (TM) helices (penultimate exons) in *CA9*, *CA12*, and *CA14* are 82, 85, and 85 bp in length, respectively. Assuming a novel juxtaposition of exons between genes coding for the ancestral TM form of *CA6* and a short PTX, the final exon of *CA6* and the first exon of the PTX gene are less likely to have been retained. Because they contain non-coding UTR sequences and lack splice donor and acceptor sites, they would be unlikely to be spliced correctly as continuous, protein-coding sequence. Taken together, this suggests that only the exon coding for the cytoplasmic domain of ancestral CA VI was lost and replaced by the single exon coding for the PTX domain. This also implies that the last exon in mammalian *CA6* and the penultimate exon of non-mammalian *CA6*, predicted to code for an APH (see below), and the penultimate exons of *CA9*, *CA12*, and *CA14*, coding for the TM helix, are highly likely to share a common ancestry.

The region after the CA domain is predicted to contain an APH

The pattern of hydrophobic residues repeating approximately every fourth residue is obvious in the alignment of the region following the CA domain (final domain in mammalian CA VI, or the segment between CA and PTX domains in non-mammalian CA VI), as seen in Fig. 4C and in the larger alignment of Fig. S1. The helical wheel

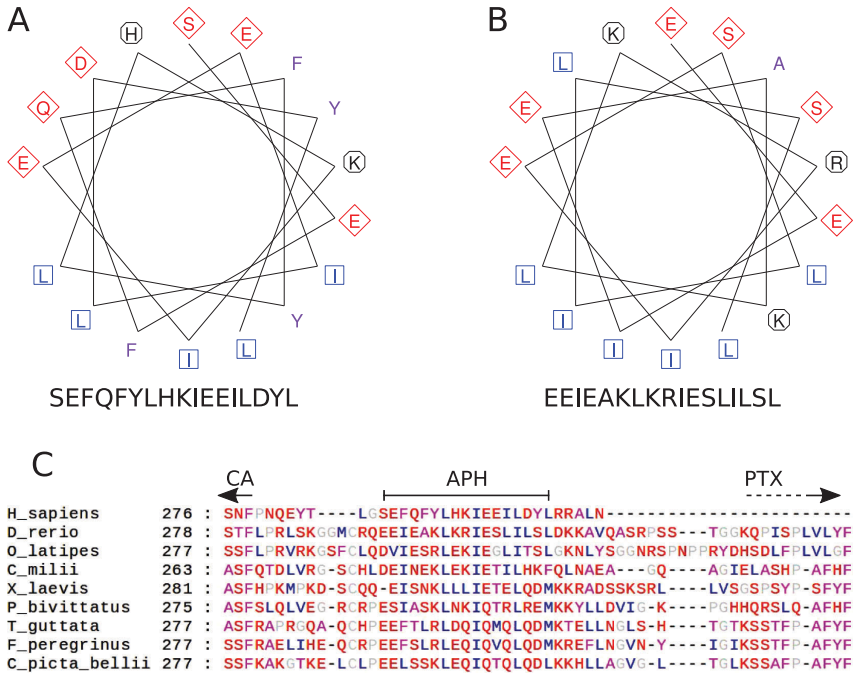

Figure 4 Amphipathic helix analysis in CA VI. (A) Helical wheel diagram of human CA VI (287–303). (B) Helical wheel diagram of zebrafish CA VI (293–310). Multiple sequence alignment of the spacer region of CA VI from indicated species. (C) CA indicates the end of the catalytic CA domain; APH is the suggested amphipathic helix, which is analyzed in (A) and (B); and PTX indicates the approximate start of the pentraxin domain (not applicable to *Homo sapiens* CA VI).

Full-size DOI: 10.7717/peerj.4128/fig-4

visualizations of Figs. 4A and 4B indicate that when folded as an alpha helix, this region of human and zebrafish CA VI, respectively, would be an APH, with one side lined with mainly hydrophobic residues (in blue and lilac). Furthermore, this region (292–312) in zebrafish CA VI is also predicted to have a high potential to form a coiled-coil structure by the COILS algorithm (*Lupas, Van Dyke & Stock, 1991*) in InterProScan at http://www.ebi.ac.uk/interpro/sequence-search (*Jones et al., 2014*). The APH region is a unique feature of CA VI, present in both non-mammalian and mammalian sequences.

Duplication of an adjacent glucose transporter gene is associated with the loss of PTX from CA VI

The genes next to *CA6* provide a clue for a possible cause of losing the PTX-encoding exon in mammalian *CA6*. We have observed 17 non-mammalian genomes with a chromosomal arrangement of *CA6*, then one glucose transporter gene (*SLC2A5/SLC2A7*), followed by the gene *GPR157*, whereas most mammalian genomes present the gene order *CA6*, *SLC2A7*, *SLC2A5*, and *GPR157*. The reconstructed syntenic block for therian mammals in the region after *CA6* in Genomicus (http://www.genomicus.biologie.ens.fr/genomicus-86.01)

(*Muffato et al., 2010*) also shows the duplicated glucose transporter, whereas those for ancestral tetrapods and bony fish lineages only have a single *SLC2A5/SLC2A7* ortholog. We were not able to find any single genome containing a PTX-coding exon with *CA6* and both *SLC2A5* and *SLC2A7*. Hence, the available genomic evidence suggests that the loss of the PTX-domain-coding exon and the duplication of the adjacent glucose transporter gene may have occurred simultaneously, close to the divergence time of the mammalian lineage. The rearrangements during the gene duplication would also provide a plausible mechanism for the exon loss.

Sequencing of zebrafish ca6 cDNA confirms a 530-residue product

We produced a PCR-amplified cDNA of zebrafish *ca6* for recombinant protein production. The resulting sequence had five synonymous substitutions compared to Ensembl ENSDART00000132733 (Fig. S4) and three unresolved bases leading to one unknown amino acid residue. Except for the unknown residue, the translation is identical to the predicted 530-residue protein (Ensembl ENSDARP00000119189 or UniProt E9QB97, Fig. S5). The cDNA sequence has been submitted to ENA database (http://www.ebi.ac.uk/ena) as LT724251 and its translation to UniProt as A0A1R4AHH7. The other predicted Ensembl transcript (ENSDART00000079007) codes for a protein of 538 residues, in which an additional 24 bp exon creates an insertion before the PTX domain.

Sequence alignment predicts three disulfides in zebrafish CA VI

Cysteine pairs 44/226 in (CA domain, MSA columns 51/234 in Fig. S1), 352/408, and 487/518 (PTX domain, columns 390/453 and columns 532/564, respectively in Fig. S1) are expected to form disulfides by sequence conservation in the multiple sequence alignment. All three disulfides are also structurally verified. The one in CA domain is seen in all structures of extracellular CAs, e.g., human CA VI in PDB 3FE4 (*Pilka et al., 2012*), and the disulfide 352/408 in the PTX domain is homologous to the one in short PTXs, e.g., human CRP in PDB 3PVN (*Guillon et al., 2014*). The third disulfide, 487/518, is also supported by proximity in our molecular model (Fig. 5A). There is one further unpaired Cys290, in the region between the CA and PTX domains (and missing from the model), which is also conserved in 76 of 78 non-mammalian sequences (Fig. S1).

3D model of zebrafish CA VI–PTX is compatible with predicted APH and disulfides

We made a homology-based model of the CA and PTX domains and combined it with an alpha helical model of the predicted APH region, using protein–protein docking to create the nearly full model. Figure 5A shows the model, in two orientations, with the CA domain at the top and the PTX domain at the bottom. The APH (pink) fills a non-polar cavity on the surface of the CA domain. The precise orientation of the PTX domain is impossible to predict with certainty, but the current model shows it leaning against the CA domain and APH. The most highly variable regions, for which no template was

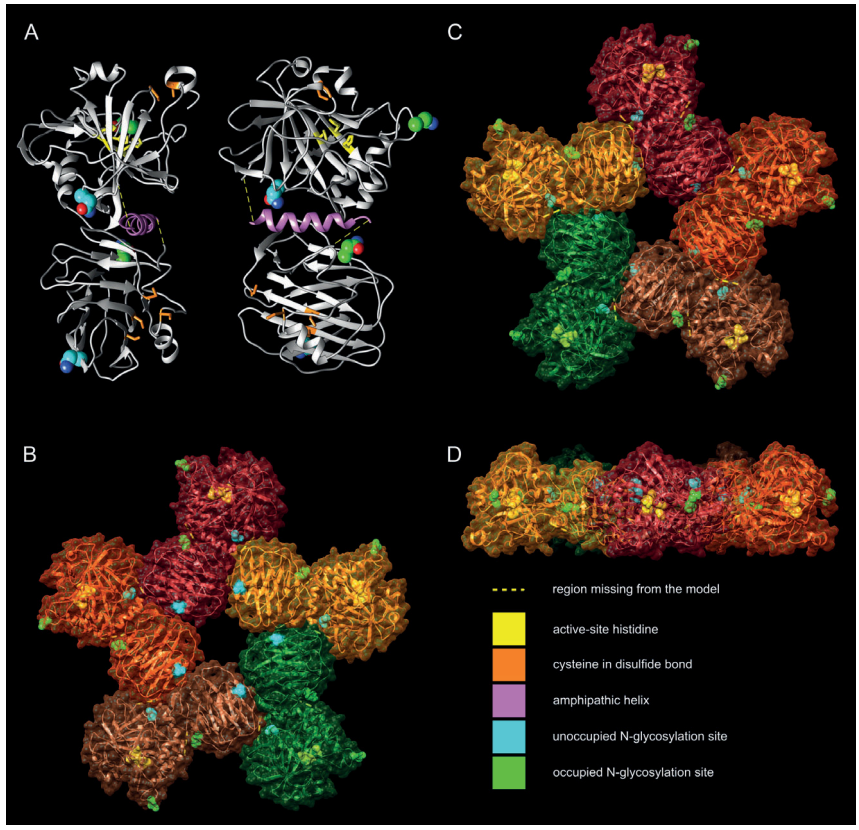

Figure 5 Molecular models of zebrafish CA VI–PTX. (A) One protomer shown in two orientations, CA domain at the top, PTX at the bottom. Potential glycosylation site Asn residues are shown as spheres, active-site histidines and assumed disulfide cysteines as sticks. (B) Front view of the pentamer model. In (B–D), individual protomers are shown in different colors. Asn residues in glycosylation sites and active-site histidines are shown in spheres, cysteines not highlighted. (C) Back view of the pentamer model. (D) Side view of the pentamer model, seen from the top of (C), with back view downwards.
Full-size DOI: 10.7717/peerj.4128/fig-5

available, were not modeled (residues 281–292 and 311–317), indicated by yellow dotted lines (Fig. 5A). In addition, the N-terminus of the model is incomplete, missing residues 20–31, which are not visible in the template 3FE4.

The zinc-binding histidines in the active site of the CA domain are shown as yellow sticks in Fig. 5A (zinc not shown), with the active-site cavity opening upwards. Disulfide-forming cysteines are presented as orange stick models. The disulfide in the CA domain and the one in the beta sheet of the PTX domain (lowest in Fig. 4A) are also present in the templates. The information of the predicted third disulfide on the surface of the PTX domain was not used when building the model, but the cysteines ended in close proximity so that the disulfide could be constructed by minor refinement of the model. This

disulfide would lock the C-terminus of the PTX domain on the surface of the domain. The presumably unpaired Cys290 is part of an unmodeled region.

Based on the pentamerization tendency of mammalian PTX domains, we constructed an additional pentameric model of zebrafish CA VI (Figs. 5B–5D) by superimposing the PTX domains of five copies of the monomer model on the pentameric structure of SAP (PDB 4AVS). Individual monomers are presented in different surface colors. There are no serious steric clashes in the model, and the domain axes align to make a flat pentamer complex (Fig. 5D), even if no pentamer constraints were applied for the monomer model. Furthermore, adjacent monomers form an additional protein–protein interface between the sides of their PTX and CA domains. The general shape of the modeled pentamer is a flat, roughly planar five-pointed star, thickness 4–5 nm and an approximate diameter 15 nm. The active site of CA faces outward in the pentamer so that the zinc-binding histidines (yellow spheres in panels B–D) are exposed in the active-site cavity, as seen in the center of panel D.

The four potentially N-glycosylated Asn residues (in the motif Asn-X-Ser/Thr) are all on the surface of the monomer, shown as spheres in Fig. 4A. In contrast, the pentamer model only shows three of them on the surface of the pentamer. Asn210, shown in cyan, is buried between the monomers, conforming well with the observed non-glycosylated status for this Asn residue. The coloring of the potential glycosylation sites in Figs. 4A–4D reflects their observed glycosylation status (presented below under mass spectrometry).

Recombinant CA VI–PTX shows a high catalytic activity

Zebrafish CA VI–PTX was produced in insect cells with high yield. The purified protein showed a single band close to the expected size in SDS-PAGE (Fig. 6, measured MW 58.6 kDa, theoretical 58.107 kDa without glycans, signal peptide excluded). Carbonate dehydratase activity was analyzed kinetically in the presence or absence of acetazolamide. The kinetic parameters of CA VI–PTX (k_{cat} and k_{cat}/K_m) were then compared with those of the thoroughly investigated CAs, namely the cytosolic and ubiquitous human isozymes α-CA I (hCA I) and II (hCA II). The CA VI–PTX possesses considerable carbonate dehydratase activity as shown in Table 2. A k_{cat} of 8.9×10^5 s^{-1} and a k_{cat}/K_m of 1.3×10^8 $M^{-1} \times s^{-1}$ show that the enzymatic activity of CA VI–PTX is almost in the same range with the very highly active human CA II. Data also show that CA VI–PTX was efficiently inhibited, with an inhibition constant of 5 nM, by the clinically-used sulfonamide, acetazolamide (5-acetamido-1,3,4-thiadiazole-2-sulfonamide).

Light scattering analysis by LC–SLS–DLS confirms multimeric structure

The molecular size of native recombinantly produced zebrafish CA VI–PTX was estimated by SLS and DLS analysis after liquid chromatography. Gel filtration analysis indicated main peak eluting at 1.52 ml retention volume according to A280 (Fig. 7, black curve). This was associated with SLS intensity peak with identical shape. Analysis of the scattering

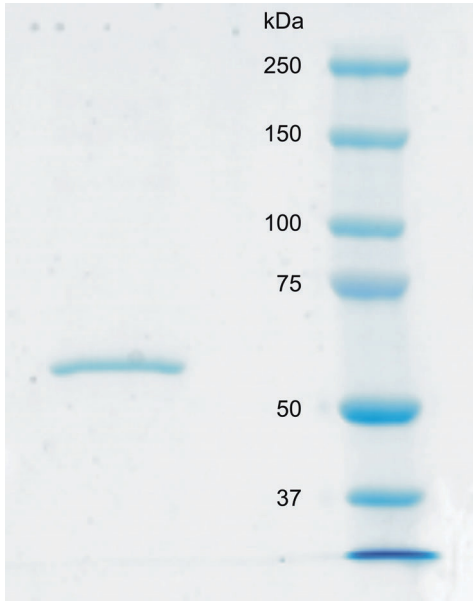

Figure 6 SDS-PAGE of recombinantly produced zebrafish CA VI–PTX. Left: purified recombinant zebrafish CA VI–PTX, molecular mass calculated from mobility 58.6 kDa. Right: molecular weight standards. Full-size DOI: 10.7717/peerj.4128/fig-6

Table 2 Kinetic parameters for CO_2 hydration reaction catalyzed by selected α-CA isozymes.

Enzyme	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (M^{-1} s^{-1})	K_I(AAZ) (nM)
hCA I[a]	2.0×10^5	4.0	5.0×10^7	250
hCA II[a]	1.4×10^6	9.3	1.5×10^8	12
Pentraxin-CA VI[b]	8.9×10^5	6.5	1.3×10^8	5

Notes:
AAZ, acetazolamide, 5-acetamido-1,3,4-thiadiazole-2-sulfonamide.
[a] Human recombinant isozymes, stopped flow CO_2 hydratase assay method (pH 7.5) (*Nishimori et al., 2007*).
[b] Zebrafish recombinant enzyme, stopped flow CO_2 hydratase assay method (pH 7.5), this work.

intensity (SLS) results in a MW estimate of 280 ± 11 kDa for the peak, and the estimate was homogeneous throughout the elution peak (Fig. 7, near-horizontal line across the peak in dark gray). In addition, DLS data was collected for the eluted peak indicating particle size of 7.69 ± 0.29 nm, as Rh (hydrodynamic radius), which is consistent with the determined MW. The MW estimate based on the retention volume in gel filtration is slightly smaller (214 ± 10 kDa), possibly due to off-globular shape of the molecule. The small peak eluting before the main peak (~1.1 ml retention volume) indicated the presence of aggregated protein, resulting in high scattering intensity. According to A280, this is less than 5% of the protein sample. Altogether, the light scattering analysis combined with gel filtration indicates oligomeric assembly for the protein, a pentameric form being the most probable oligomeric state.

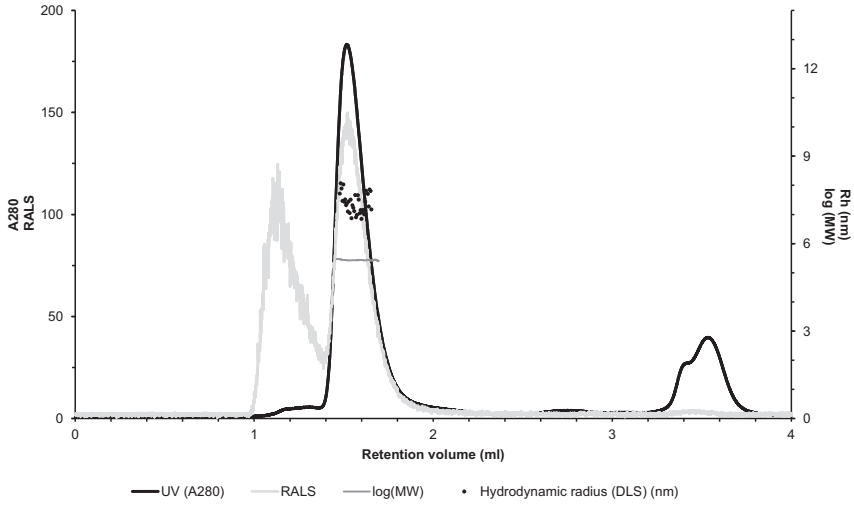

Figure 7 Assessment of the oligomeric size of zebrafish CA VI. Gel permeation chromatography was used to study the characteristics of recombinantly produced zebrafish CA VI. The left *Y*-axis shows the UV absorption intensity (280 nm wavelength) and light scattering (LS) intensity. UV intensity was used for the determination of the protein concentration. Molecular weight (MW) was calculated using LS intensity and shown on the right *Y*-axis. Hydrodynamic radius (R_h) was calculated from the dynamic light scattering signal, and is also shown on the right *Y*-axis. In addition, the oligomeric size of zebrafish CA VI was evaluated based on the penetration time using molecular weight marker proteins as a standard. Full-size DOI: 10.7717/peerj.4128/fig-7

Mass spectrometry confirms post-translational modifications

All attempts to characterize the intact CA VI–PTX with ESI FT-ICR mass spectrometry failed, despite the extensive sample desalting/purification prior to the measurements. This may be due to a slight protein precipitation observed during the sample preparation. Therefore, in-solution trypsin digestion was selected as the main route for structural characterization of CA VI–PTX. The digestion was performed in non-reducing conditions to preserve disulfide bonds in the structure. The digestion resulted in 97% sequence coverage with 64 specific tryptic peptides identified (Figs. 8 and 9, and fuller details in Fig. S6 and Table S1).

The peptide map in Fig. 9 shows that the tryptic peptides were found within both protein domains, although somewhat larger peptides (up to ~14 kDa) were found within the PTX domain. Out of all identified peptides, 12 contained disulfide bonds (either intra- or interpeptide). These peptides confirmed the putative disulfide bonds, Cys 44/226 in the CA domain, and Cys 352/408 and Cys 487/518 in the PTX domain. Cys290 in the spacer region is most likely free but the corresponding tryptic peptide (LSKGGMCR) was not observed to confirm this. These disulfide bonds are fully consistent with the 3D structural model of CA VI–PTX.

Carbonic anhydrase VI–pentraxin contains four putative N-glycosylation sites (Asn210, Asn258, Asn339 and Asn394), having a canonical NxS/T consensus sequence

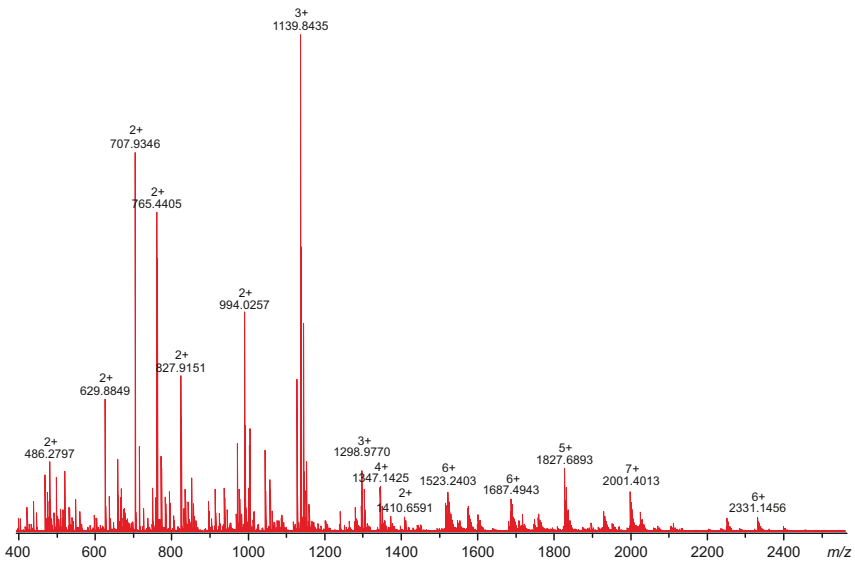

Figure 8 High-resolution mass spectrum of the tryptic digest of CAVI-PTX. The mass spectrum was measured by direct infusion on a 12-T Fourier transform ion cyclotron resonance instrument using positive-ion electrospray ionization. Monoisotopic *m*/*z* values and charge states obtained through peak deconvolution are indicated for the most abundant peaks. Full-size DOI: 10.7717/peerj.4128/fig-8

(marked in Fig. 9). Among the identified tryptic peptides, 12 glycopeptides were found. On the basis of these peptides, CA VI–PTX carries two glycans, a core-fucosylated oligomannose type glycan $GlcNAc_2(Fuc)Man_3$ at Asn258 and an oligomannose type glycan $GlcNAc_2Man_3$ at Asn339, located in the CA domain and PTX domain, respectively. These glycosylation sites and glycan structures were further verified by CID-MS/MS experiments of the representing glycopeptides [248–266] (3416.5084 Da) and [331–347] (2819.3059 Da) (Fig. 10). As no other glycan variants were observed among the peptides, it seems that the glycosylation in CA VI–PTX (produced in insect cells) is rather homogenous. These results are consistent with accessibility of the sites predicted by our model. Interestingly, the glycosylation site at Asn258 is conserved in 77 out of 78 non-mammalian CA VI sequences in the sequence alignment Fig. S1 (columns 266–268). The tryptic peptide [191–216] (2988.4744 Da) was only observed in a free form, indicating that Asn210 is non-glycosylated in the CA domain. Similarly, the peptides spanning the Asn394 residue were all observed without any glycans attached (Fig. S6), suggesting that this site is non-glycosylated in the PTX domain.

Immunohistochemistry shows cell surface localization of CA VI–PTX in various tissues

Recombinant zebrafish CA VI–PTX protein was used to raise a rabbit polyclonal antiserum, which worked well in immunofluorescence studies. Figure 11 shows positive

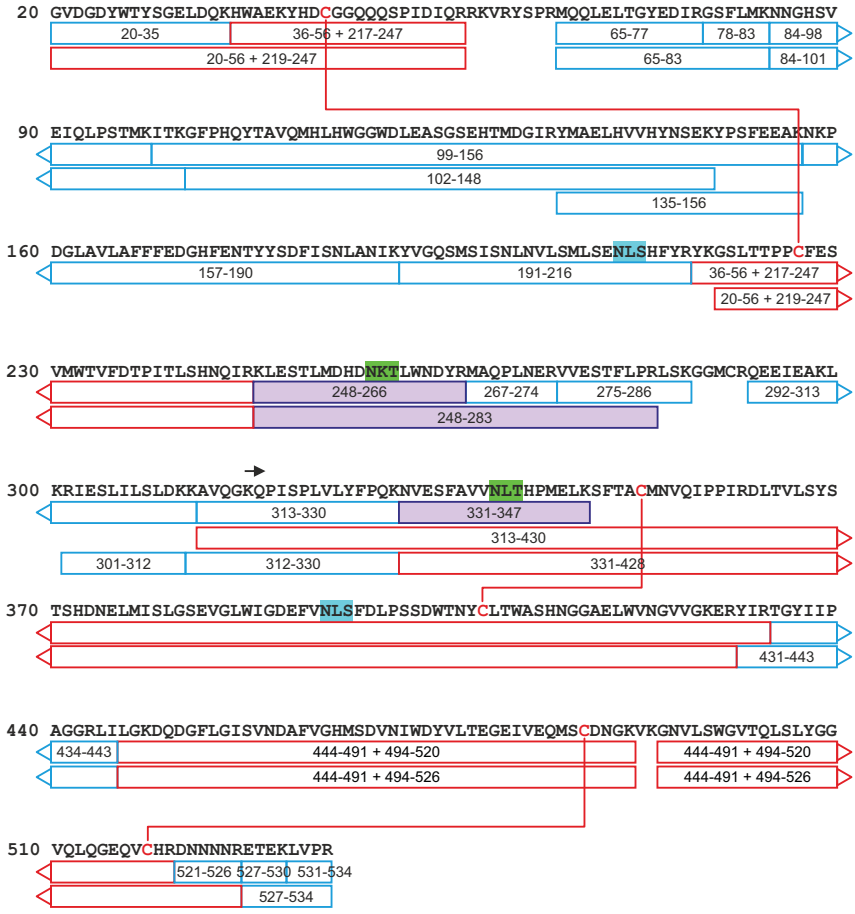

Figure 9 A tryptic peptide map of selected peptides of zebrafish CA VI–PTX. The identified tryptic peptides are indicated with blue boxes showing the start and the end residues. The confirmed disulfide bonds are indicated with red lines with the corresponding peptides indicated with red boxes. The four potential N-glycosylation sites are indicated with a distinct background color (blue: unoccupied N-glycosylation site; and green: occupied N-glycosylation site). The three observed glycopeptides are marked with purple boxes. The start of the PTX domain (KQP . . .) has been indicated with a black arrow. This figure shows a minimum amount of peptides for maximal coverage, whereas all identified peptides are shown in Fig. S6. Full-size DOI: 10.7717/peerj.4128/fig-9

staining in the skin, heart, gills, and swim bladder. The strongest signal is seen on cell surfaces, while the intracellular staining was detectable but weaker.

To get further insights into *ca6* expression in zebrafish, we also studied the expression pattern in different tissues of adult zebrafish by qRT-PCR. As shown in Table 3, relative expression of *ca6* mRNA was found to be prominent in the fins/tail, and brain. Low levels of expression were observed in the gills, kidney, teeth, skin, and spleen. A very faint signal was detected in the swim bladder, intestine, pancreas, liver, eggs, and heart.

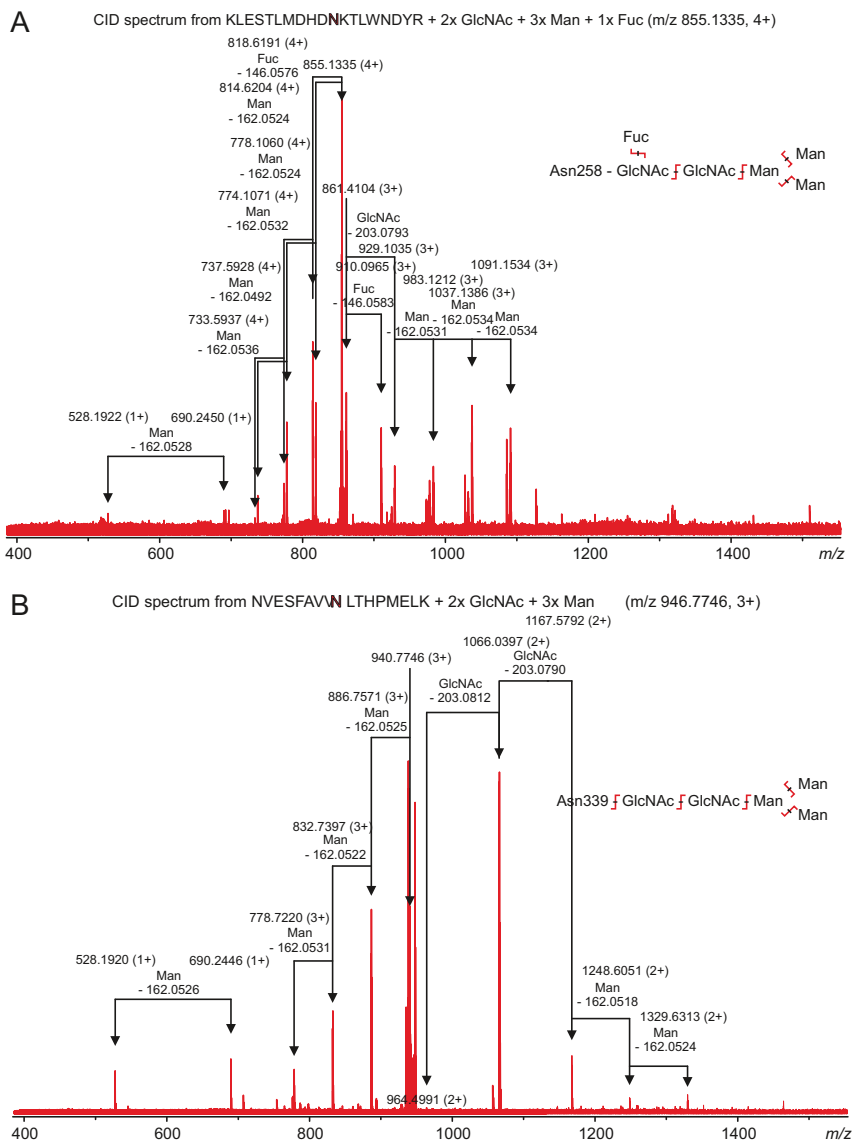

Figure 10 Characterization of zebrafish CA VI–PTX glycopeptides by tandem mass spectrometry. The precursor ions of the two observed glycopeptides, with monoisotopic masses of 3416.5084 Da and 2819.3059 Da (residues 248–266 and 331–347, respectively), were mass-selected in a quadrupole for collision-induced dissociation tandem mass spectrometry. The fragmentation patterns are consistent with the presence of the standard N-glycosylation core pentasaccharide with fucosylation in the innermost *N*-acetylglucosamine residue in the glycopeptide 248–266 (A) and a similar non-fucosylated pentasaccharide in the glycopeptide 331–347 (B). Full-size DOI: 10.7717/peerj.4128/fig-10

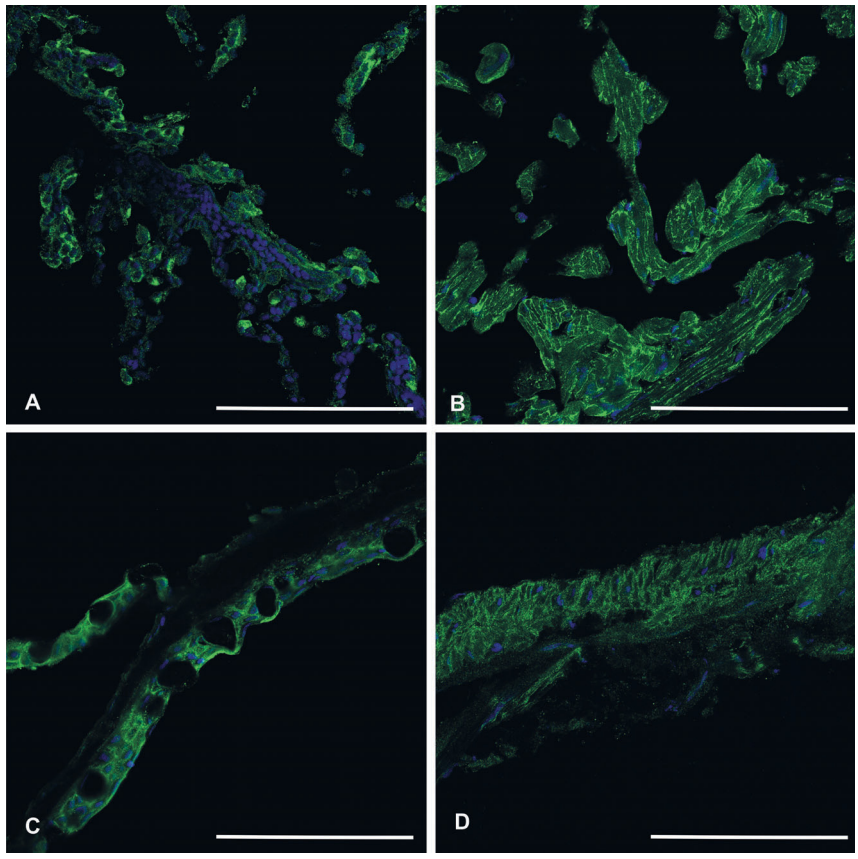

Figure 11 Immunohistochemistry of CA VI–PTX in adult zebrafish tissues. Tissue sections stained by anti-zebrafish CA VI–PTX (green) and nuclear staining by DAPI (blue). Gills (A), heart (B), skin (C), swim bladder (D). The strongest signal (A–D) is present on the cell surfaces, even though the cell interior gives some background staining. Scale bars 100 μm. Full-size DOI: 10.7717/peerj.4128/fig-11

Zebrafish cannot swim properly in the *ca6* knockdown model

Gene-specific antisense MOs have been widely used to inhibit gene expression in zebrafish larvae (*Eisen & Smith, 2008*). We designed two different MOs, one for translational blocking of *ca6* mRNA and the other for blocking intron splicing before exon 9. Both MOs were used to repeat all knockdown experiments with highly similar results, suggesting equally efficient knockdown in both kinds of *ca6* morphants. We did not see any morphological differences between uninjected and RC MO-injected embryos/larvae over the period of five days of development. The *ca6* morphant zebrafish embryos between 1 and 3 dpf were also devoid of any notable morphological changes, but interestingly, at the end of 4 dpf we consistently observed an underdeveloped or deflated swim bladder in *ca6* morphant larvae (Fig. 12).

Table 3 **Relative expression ratios of *ca6* mRNA in adult zebrafish tissues.**

Tissue	Relative expression
Fin/tail	214.82
Teeth	6.81
Spleen	2.37
Kidney	19.44
Brain	293.46
Swim bladder	0.61
Heart	0.01
Gills	68.70
Skin	4.32
Intestine	0.41
Pancreas	0.27
Eggs	0.01
Liver	0.02

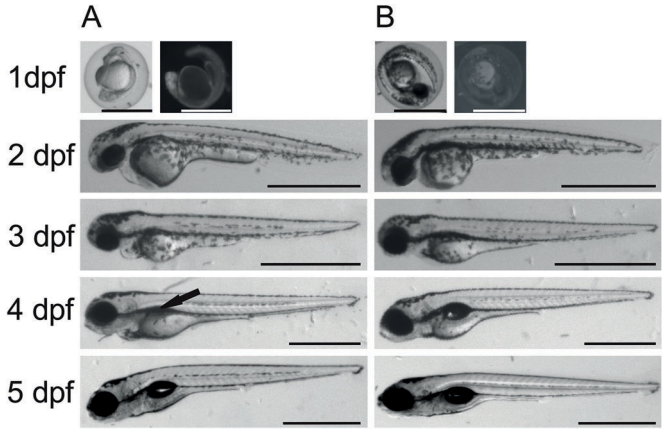

Figure 12 **Comparison between morpholino-injected and wild-type zebrafish larvae.** The morphant larvae (A) showed consistently a deflated swim bladder at 4 dpf (arrow), which returned to normal morphology at 5 dpf. Wild-type larvae of the same ages are shown for comparison (B). Scale bars 1 mm.
Full-size DOI: 10.7717/peerj.4128/fig-12

The quantitative expression analysis of *ca6* mRNA was done in wild-type and the *ca6* morphant zebrafish at different stages of development. As seen in Fig. 13, the mRNA expression in wild-type embryos was highest at 24 hpf, with slightly lower values later. The levels of *ca6* mRNA were consistently higher in the morphant embryos compared to the wild-type, possibly because of compensatory upregulation of the gene caused by the absence of CA VI protein. The peak expression of *ca6* was at 48 hpf in the *ca6* morphant embryos.

In order to measure swimming activity of morphant vs. wild-type, we calculated total distances traveled for individual larvae, and they are presented as boxplots in Fig. 3.

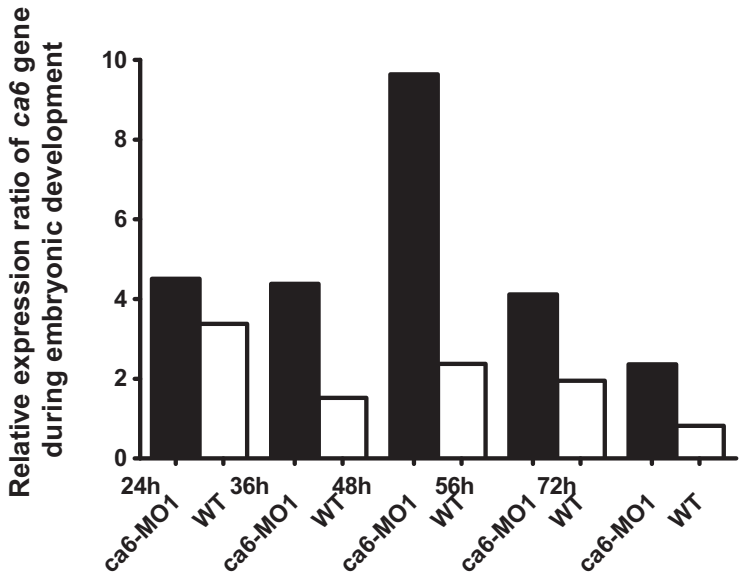

Figure 13 Developmental expression pattern of *ca6* in 1–5 dpf larvae. The expression levels of the *ca6* gene was studied using qRT-PCR from the total mRNA isolated from 1 to 5 dpf of *ca6* morphant and wild-type larvae. The results of *ca6* gene expression were normalized using *gapdh* as internal control.
Full-size DOI: 10.7717/peerj.4128/fig-13

Table 4 Statistics of swimming pattern analysis of *ca6* morphant and wild-type zebrafish.

A[a]	Median	Mean	SD	Range	*p*-Value
Day 4					
KD	0.00	1.87	6.59	0.00–49.03	4.28×10^{-19} [b]
WT	13.80	13.12	6.01	0.00–25.45	1.90×10^{-3} [c]
Day 5					
KD	4.75	4.89	4.26	0.00–5.19	1.16×10^{-7} [b]
WT	10.22	10.38	3.18	1.83–17.27	
B[d]	**Median**	**Mean**	**SD**	**Range**	***p*-Value**
Day 4					
KD	0	9.13	20.46	0.00–60.00	8.68×10^{-11} [b]
WT	31	29.59	22.44	0.00–60.00	4.98×10^{-3} [c]
Day 5					
KD	24.51	26.96	26.23	0.00–60.00	2.98×10^{-4} [b]
WT	51.47	45.29	16.42	0.00–60.00	

Notes:
KD, knockdown; WT, wild-type.
[a] Swimming distances (cm).
[b] KD compared to WT.
[c] Day 4 WT compared to day 5 WT.
[d] Time spent in upper half of the flask.

Two-sample Kolmogorov–Smirnov statistical analyses were performed between relevant group pairs to determine if they could have been drawn from the same distribution. Day 4 knockdown larvae swam less (median 0.00 cm) than day 4 wild-type larvae (median 13.80 cm, p-value 4.28×10^{-19}), and similarly day 5 knockdown larvae swam less (median 4.75 cm) than day 5 wild-type larvae (median 10.22 cm, p-value 1.16×10^{7}). Full details of the swimming data are shown in Table 4. Taken together with the clearly observed swim bladder deficiency in 4 dpf larvae (Fig. 12) and the presence of CA VI in adult zebrafish swim bladder, we suggest that CA VI is required either for swim bladder development or swim bladder function. When CA VI expression is mainly restored in 5 dpf larvae, the swimming pattern also returns to almost normal.

DISCUSSION

This study consists of the characterization of a novel type of a CA, CA VI containing a PTX domain, by means of sequence analyses, phylogenetics, molecular modeling, experiments on a recombinantly produced protein, knockdown of the *ca6* gene in zebrafish embryos, and expression studies by immunohistochemistry and qRT-PCR. The bioinformatic and experimental analyses build a coherent picture of the structure of this novel domain combination, and the evolutionary analysis shows a history of domain gains and losses. Based on our previous work and the findings in this study, we propose that CA VI–PTX in zebrafish is needed for filling the swim bladder, and possibly in a novel type of membrane anchoring and immune function.

The PTX domain found associated with non-mammalian CA VI is a novel member of the PTX family. We have shown it to be most closely related with the short PTXs, CRP and SAP (Fig. 2). The association of a CA domain with a PTX domain is new in both the PTX and CA families. SAP and CRP are more closely similar to each other than either is to the CA-associated PTX domain. This could indicate that the CA-associated PTX domain had diverged from a common ancestor before the duplication that created SAP and CRP, but we cannot take this for granted, because adaptation to create a viable domain interface may have accelerated the rate of change in the CA-associated PTX domain.

The phylogenetic tree in Fig. 1 shows that the TM CAs IX, XII, and XIV and secretory CA VI share a common ancestor. We propose that the quartet has arisen in the two whole-genome duplications in early vertebrates. Figure 14 presents a plausible sequence of events that could have led to present-day domain structures in CA VI. Briefly, we assume that the exon coding for the cytoplasmic domain in ancestral CA VI was replaced by an exon coding for a PTX domain (probably by a duplication or a move of an exon coding for a short PTX in early vertebrates), and the TM helix transformed into an APH (Figs. 4 and 5). Later, presumably in the therian mammal lineage, the PTX domain was lost, leaving the APH in the C-terminus of CA VI. These hypotheses are supported by the following observations: (1) Comparison of exon lengths suggests the TM-helix-coding exon as the most likely ancestor of the exon coding the spacer region after the CA domain in CA VI; (2) the losses of the CP domain in early CA VI and of PTX domain in mammalian lineage are more parsimonious assumptions than their acquisition in

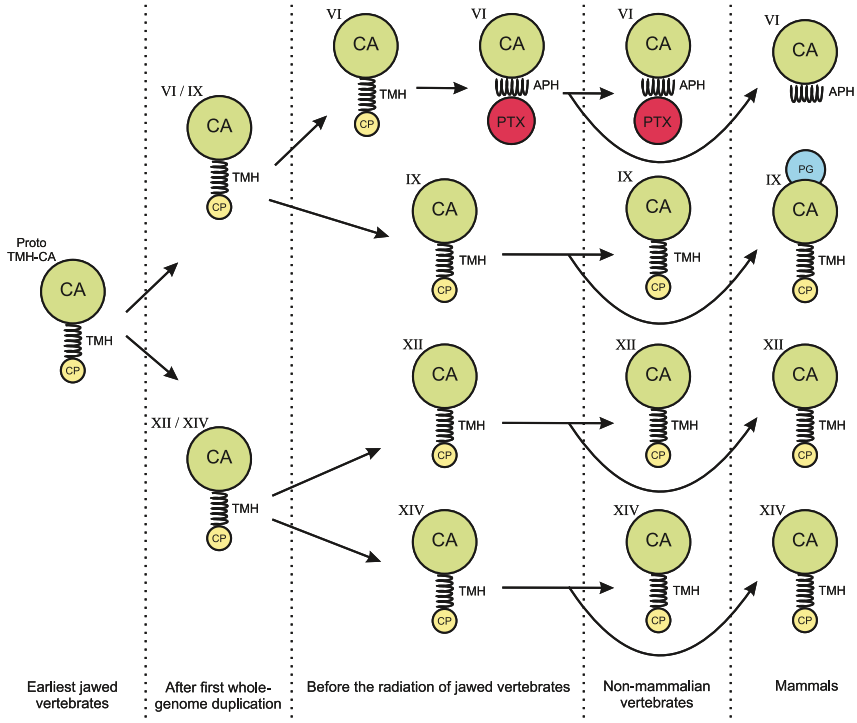

Figure 14 Hypothesis of evolution of the domain composition in CA VI and the transmembrane CA isoforms. CA, catalytic CA domain; TMH, transmembrane helix; APH, amphipathic helix; PTX, pentraxin domain; PG, proteoglycan domain. Image credit: Original digital art by Jukka Lehtiniemi.
Full-size DOI: 10.7717/peerj.4128/fig-14

multiple lineages; (3) the duplication of the glucose transporter genes *SLC2A5* and *SLC2A7*, as seen in therian mammals, is evidence of rearrangements in the region adjacent to the PTX-domain-coding exon of the *CA6* locus, which we assume to have led to the loss of the PTX domain in mammalian CA VI; and (4) the PTX domain is consistently present in non-mammalian CA VI and missing from mammalian CA VI (most likely excepting platypus).

Considering the monomer MW of 58.1 kDa (plus glycosylation), the LC–SLS–DLS results clearly confirm that zebrafish CA VI is an oligomer. The MW estimated by LC–SLS (280 ± 11 kDa) is slightly less than MW calculated from sequence (290.5 kDa for pentamer, plus glycosylation). Based on the gel filtration retention volume and protein standards, the MW is estimated to be slightly smaller (214 ± 10 kDa), but this result may be affected by column interactions and deviation from the globular shape. Furthermore, the hydrodynamic radius calculated from light scattering (7.69 ± 0.29 nm; diameter 15.38 ± 0.58 nm) suggests a particle size in the range of 364–434 kDa for globular particle. In this context, it has to be noted that diffusion of the particle is highly dependent on the molecular shape and DLS-based estimate may also be slightly affected by irregular shape.

Taken together, the light scattering results are more compatible with a pentamer than tetramer or hexamer models. The 3D model of CA VI–PTX as a pentamer (Figs. 5B–5D) predicts a shape of a flat, roughly planar five-pointed star, thickness 4–5 nm and approximate diameter 15 nm, i.e., clearly off-globular, which would explain the minor conflicts between observations. What is more, the pentamer model is also supported by known pentamerization of related PTXs (CRP and SAP). However, we need to stress that the relative orientations of the CA and PTX domains in our models are only tentative.

Mass spectrometry confirms that the N-terminus of the mature CA VI–PTX coincides with the predicted signal peptide cleavage site between residues 19 and 20. Glycopeptides with typical N-linked glycans are observed associated with Asn258 and Asn339, whereas the peptides containing Asn210 or Asn394 are only seen in non-glycosylated form (Fig. 9). Consistent with these observations of N-glycosylation, our 3D model of pentameric CA VI–PTX (Figs. 5B–5D) shows that Asn258 and Asn339 are well exposed, whereas Asn210 is fully buried in the protomer/domain interface, and Asn394 would be somewhat hindered at the protomer interface.

We discovered a minor but surprising outcome in the knockdown model regarding the poor floating ability, most likely caused by a deflated swim bladder, both of which we observed consistently in 4 dpf knockdown larvae. The statistically significant lower swimming distances and stationary positioning at the bottom of 4 dpf knockdown larvae, vs. those of 4 dpf wild-type larvae or 5 dpf knockdown larvae, imply that the knockdown larvae gain normal swimming function as the knockdown action of the injected MOs is relieved (Fig. 12). CA VI–PTX function within the swim bladder is further supported by immunohistochemistry and qRT-PCR, showing expression of both CA VI–PTX protein and mRNA in the swim bladder specimens. However, at the current point we cannot distinguish whether the swim bladder dysfunction observed in 4 dpf larvae is due to delayed development or the need of CA VI–PTX in swim bladder inflation.

C-reactive protein and serum amyloid P are known to bind carbohydrates, i.e., they are lectins (*Hind et al., 1984*; *Kottgen et al., 1992*). The calcium-binding residues in the sugar binding site are partially conserved between these two lectins and the CA-associated PTX domain. In addition, PTXs are a subfamily of the Concanavalin A-like lectin/glucanase family, which contains numerous other lectins (leguminous plant lectins, animal galectins, etc.) and other proteins interacting with carbohydrates (http://www.ebi.ac.uk/interpro/entry/IPR013320). In our immunohistochemistry results the CA VI–PTX protein shows mostly a strong cell-surface staining pattern (Fig. 11), even if the protein is predicted to be a secreted, soluble protein. We assume that the PTX domain in CA VI would also be a lectin and anchor the protein on the cell surface via sugars in glycoconjugates. Binding to plasma membrane glycoconjugates would also explain why the loss of the TMH was tolerated, i.e., TM helix anchoring was replaced by lectin anchoring. If sugar binding by CA VI–PTX can be proved experimentally, non-mammalian CA VI would represent the first case of an enzyme which is attached on the cell surface by lectin binding.

Lectins and other PRMs are an important part of the innate immune system in fishes, which is more diverse than that of mammals (*Vasta et al., 2011*; *Sunyer, Zarkadis & Lambris, 1998*). Although teleost fish lack lymph nodes and bone marrow, the anterior part of the fish kidney is considered a functional ortholog of mammalian bone marrow. Thus, it represents the main hematopoietic lymphoid tissue of teleosts, and is thought to be an immunologically responsive organ (*Zapata & Amemiya, 2000*). The role of maintenance of mucosal homeostasis is served in teleosts by the gut, skin, and gills, which all contain mucosa-associated lymphoid tissue (*Salinas, Zhang & Sunyer, 2011*). These are among the tissues where zebrafish CA VI–PTX has its highest expression, and therefore we assume that this protein is a part of the innate immune system.

Interestingly, we have shown that mouse CA VI is also highly expressed in the gut, specifically in the immunologically active Peyer's patches (*Pan et al., 2011*). In another study, we demonstrated that there is a likely role for *Car6* in immune stimulated lung tissues (*Patrikainen et al., 2016*) and murine *Car6* is likely involved in mucosa maintenance in both airways and gut (*Leinonen et al., 2004*; *Parkkila et al., 1997*). We formed a preliminary hypothesis that mouse CA VI is involved in immunological functions, which has been confirmed recently (*Xu et al., 2017*), by showing that CA VI isoform B promotes interleukin-12 expression. However, a gene regulatory function is unlikely for zebrafish CA VI, with the estimated diameter of 15 nm for the pentamer making it too large to enter the nuclear pores. The locations of high *ca6*/*CA6* expression in fish and in mammals are similar in that they allow delivery of CA VI on the physical barrier against external environment (gut, skin, and gills in zebrafish; skin, saliva, milk, and lungs in human/mouse), consistent with a function associated with primary immune defense. Summing up, we suggest that both mammalian and fish CA VI are components of the innate immune system, with or without a PTX domain.

Given the dynamic nature of genomes, with transposition and translocation events constantly shuffling exons, it is hard to see the choreography of domain moves in CA VI during vertebrate evolution as anything more than chance events. However, in order to remain stably in a genome, the changes must be at least tolerated, or possibly provide some advantage to their carrier. We see the addition of the PTX domain in early jawed vertebrates as a tolerated change, in which membrane attachment through a TM helix was replaced by lectin anchoring. As we have suggested, the new domain context may have led to the CA domain of CA VI adopting functionality within the innate immune system. Then later, when the PTX domain was lost, presumably through the local segmental duplication leading to a duo of glucose transporters (*SLC2A5* and *SLC2A7*), the addition of another glucose transporter may have been more of an advantage than the loss of pentamerization and membrane anchoring in CA VI, and thus this chromosomal arrangement became fixed in early therian mammals. The loss of the PTX domain may also have opened the way for using the APH in forming dimers. We have a preliminary result of human CA VI being a mixture of monomer and dimer forms in solution (A. Yrjänäinen, 2017, unpublished data), in which we speculate dimerization to be mediated by the amphipathic helices being able to join in a coiled-coil fashion when unhindered by a further C-terminal domain.

This study has given us many ideas for future research. We plan to take a closer look at the complex evolution of non-mammalian PTXs, which might shed more light on the origin of the CA VI-linked PTX domain and on structure-related constraints on its surface. We have also started work on comparisons of per-residue conservation patterns of the CA domain in mammalian vs. non-mammalian CA VI. Testing the sugar-binding ability of CA VI–PTX will be the obvious way to explore the lectin hypothesis.

ACKNOWLEDGEMENTS

We thank Aulikki Lehmus and Marianne Kuuslahti for the skillful technical assistance with most experiments; Leena Mäkinen, and Hannaleena Piippo, for their technical assistance with zebrafish experiments, and Jukka Lehtiniemi for the artwork of Fig. 14. Thanks are due to Alma Yrjänäinen and Linda Urbański for the help with immunohistochemistry experiments and collecting tissues. We thank Mataleena Parikka for the help with adult zebrafish tissue collection. The authors thank Ritva Romppanen for preparing samples for mass spectrometry analysis. We acknowledge Biocenter Finland for infrastructure support in light scattering experiments. Core facilities at BioMediTech and Faculty of Medicine and Life Sciences, University of Tampere, were essential in microscopy (Tampere Imaging Facility), zebrafish experiments (Zebrafish Laboratory), and in DNA sequencing (Sequencing Facility).

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by grants from the Jane & Aatos Erkko Foundation (Seppo Parkkila), the Sigrid Jusélius Foundation (Seppo Parkkila), the Tampere Tuberculosis Foundation (Seppo Parkkila), the Finnish Cultural Foundation (Harlan R. Barker), and the Academy of Finland (Seppo Parkkila, Vesa P. Hytönen). There was no additional external funding received for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Jane & Aatos Erkko Foundation.
Sigrid Jusélius Foundation.
Tampere Tuberculosis Foundation.
Finnish Cultural Foundation.
Academy of Finland.

Competing Interests

The authors declare that they have no competing interests. Seppo Parkkila and Vesa Hytönen were part-time employees of Fimlab Laboratories Ltd. (laboratory center owned by the local university hospital district) at the time of performing the first experiments of this study. Ashok Aspatwar was an employee of Fimlab Laboratories Ltd.

for some periods during this study. Csaba Ortutay is an employee of Hiducator Ltd., Prajwol Manandhar is an employee of The Center for Molecular Dynamics Nepal, and Mika Hilvo is an employee of Zora Biosciences Ltd.

Author Contributions

- Maarit S. Patrikainen conceived and designed the experiments, performed the experiments, analyzed the data, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Martti E.E. Tolvanen conceived and designed the experiments, performed the experiments, analyzed the data, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Ashok Aspatwar conceived and designed the experiments, performed the experiments, analyzed the data, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Harlan R. Barker conceived and designed the experiments, performed the experiments, analyzed the data, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Csaba Ortutay conceived and designed the experiments, performed the experiments, analyzed the data, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Janne Jänis conceived and designed the experiments, performed the experiments, analyzed the data, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Mikko Laitaoja conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, reviewed drafts of the paper.
- Vesa P. Hytönen conceived and designed the experiments, performed the experiments, analyzed the data, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Latifeh Azizi performed the experiments, analyzed the data, prepared figures and/or tables, reviewed drafts of the paper.
- Prajwol Manandhar performed the experiments, analyzed the data, prepared figures and/or tables, reviewed drafts of the paper.
- Edit Jáger performed the experiments, analyzed the data, reviewed drafts of the paper.
- Daniela Vullo performed the experiments, analyzed the data, prepared figures and/or tables, reviewed drafts of the paper.
- Sampo Kukkurainen performed the experiments, reviewed drafts of the paper, original discovery of CA VI–PTX.
- Mika Hilvo analyzed the data, reviewed drafts of the paper, original idea that the PTX might actually pentamerize non-mammalian CA VI.
- Claudiu T. Supuran conceived and designed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, reviewed drafts of the paper.
- Seppo Parkkila conceived and designed the experiments, contributed reagents/materials/analysis tools, wrote the paper, reviewed drafts of the paper.

Animal Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

Zebrafish housing and care in the Zebrafish facility of the University of Tampere have been approved by the National Animal Experiment Board of Finland, administered through the Provincial Government of Western Finland, Province Social and Health Department Tampere Regional Service Unit (permit # LSLH-2007-7254/Ym-23). Using five day old zebrafish as a model organism requires no specific ethical permission, likewise studying tissues after euthanizing adult fish.

DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

The zebrafish ca6 cDNA sequence has been deposited to the ENA database (http://www.ebi.ac.uk/ena) and assigned the identifier LT724251. The translated protein sequence is available in UniProt (http://www.uniprot.org) as A0A1R4AHH7.

Data Availability

The following information was supplied regarding data availability:

Multiple sequence alignments for phylogenetic trees are provided as files Data S1–S6, and are referenced in the appropriate locations in the text. A multi-page protein sequence alignment is provided as Fig. S1.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.4128#supplemental-information.

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* **215(3)**:403–410 DOI 10.1006/jmbi.1990.9999.

Berrino E, Bua S, Mori M, Botta M, Murthy VS, Vijayakumar V, Tamboli Y, Bartolucci G, Mugelli A, Cerbai E, Supuran CT, Carta F. 2017. Novel sulfamide-containing compounds as selective carbonic anhydrase I inhibitors. *Molecules* **22(7)**:1049 DOI 10.3390/molecules22071049.

Bottazzi B, Inforzato A, Messa M, Barbagallo M, Magrini E, Garlanda C, Mantovani A. 2016. The pentraxins PTX3 and SAP in innate immunity, regulation of inflammation and tissue remodelling. *Journal of Hepatology* **64(6)**:1416–1427 DOI 10.1016/j.jhep.2016.02.029.

de Vries SJ, Bonvin AM. 2011. CPORT: a consensus interface predictor and its performance in prediction-driven docking with HADDOCK. *PLOS ONE* **6(3)**:e17695 DOI 10.1371/journal.pone.0017695.

de Vries SJ, van Dijk AD, Krzeminski M, van Dijk M, Thureau A, Hsu V, Wassenaar T, Bonvin AM. 2007. HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. *Proteins: Structure, Function, and Bioinformatics* **69(4)**:726–733 DOI 10.1002/prot.21723.

Eisen JS, Smith JC. 2008. Controlling morpholino experiments: don't stop making antisense. *Development* **135(10)**:1735–1743 DOI 10.1242/dev.001115.

Fernley RT, Wright RD, Coghlan JP. 1979. A novel carbonic anhydrase from the ovine parotid gland. *FEBS Letters* **105(2)**:299–302 DOI 10.1016/0014-5793(79)80634-1.

Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research* **44(D1)**:D279–D285 DOI 10.1093/nar/gkv1344.

Flicek P, Amode MR, Barrell D, Beal K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fairley S, Fitzgerald S, Gil L, Gordon L, Hendrix M, Hourlier T, Johnson N, Kahari AK, Keefe D, Keenan S, Kinsella R, Komorowska M, Koscielny G, Kulesha E, Larsson P, Longden I, McLaren W, Muffato M, Overduin B, Pignatelli M, Pritchard B, Riat HS, Ritchie GR, Ruffier M, Schuster M, Sobral D, Tang YA, Taylor K, Trevanion S, Vandrovcova J, White S, Wilson M, Wilder SP, Aken BL, Birney E, Cunningham F, Dunham I, Durbin R, Fernandez-Suarez XM, Harrow J, Herrero J, Hubbard TJ, Parker A, Proctor G, Spudich G, Vogel J, Yates A, Zadissa A, Searle SM. 2012. Ensembl 2012. *Nucleic Acids Research* **40(D1)**:D84–D90 DOI 10.1093/nar/gkr991.

Garlanda C, Bottazzi B, Bastone A, Mantovani A. 2005. Pentraxins at the crossroads between innate immunity, inflammation, matrix deposition, and female fertility. *Annual Review of Immunology* **23(1)**:337–366 DOI 10.1146/annurev.immunol.23.021704.115756.

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research* **31(13)**:3784–3788 DOI 10.1093/nar/gkg563.

Guillon C, Bigouagou UM, Folio C, Jeannin P, Delneste Y, Gouet P. 2014. A staggered decameric assembly of human C-reactive protein stabilized by zinc ions revealed by X-ray crystallography. *Protein and Peptide Letters* **22(3)**:248–255 DOI 10.2174/0929866522666141231111226.

Henkin RI, Lippoldt RE, Bilstad J, Edelhoch H. 1975. A zinc protein isolated from human parotid saliva. *Proceedings of the National Academy of Sciences of the United States of America* **72(2)**:488–492 DOI 10.1073/pnas.72.2.488.

Hewett-Emmett D. 2000. Evolution and distribution of the carbonic anhydrase gene families. *EXS* **90**:29–76.

Hind CR, Collins PM, Renn D, Cook RB, Caspi D, Baltz ML, Pepys MB. 1984. Binding specificity of serum amyloid P component for the pyruvate acetal of galactose. *Journal of Experimental Medicine* **159(4)**:1058–1069 DOI 10.1084/jem.159.4.1058.

Hunter JD. 2007. Matplotlib: a 2D graphics environment. *Computing in Science & Engineering* **9(3)**:90–95 DOI 10.1109/mcse.2007.55.

Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30(9)**:1236–1240 DOI 10.1093/bioinformatics/btu031.

Karhumaa P, Leinonen J, Parkkila S, Kaunisto K, Tapanainen J, Rajaniemi H. 2001. The identification of secreted carbonic anhydrase VI as a constitutive glycoprotein of human and rat milk. *Proceedings of the National Academy of Sciences of the United States of America* **98(20)**:11604–11608 DOI 10.1073/pnas.121172598.

Khalifah RG. 1971. The carbon dioxide hydration activity of carbonic anhydrase. I. Stop-flow kinetic studies on the native human isoenzymes B and C. *Journal of Biological Chemistry* **246**:2561–2573.

Kolstoe SE, Jenvey MC, Purvis A, Light ME, Thompson D, Hughes P, Pepys MB, Wood SP. 2014. Interaction of serum amyloid P component with hexanoyl bis(D-proline) (CPHPC).

Acta Crystallographica Section D: Biological Crystallography **70(8)**:2232–2240 DOI 10.1107/s1399004714013455.

Kottgen E, Hell B, Kage A, Tauber R. 1992. Lectin specificity and binding characteristics of human C-reactive protein. *Journal of Immunology* **149**:445–453.

Leinonen J, Parkkila S, Kaunisto K, Koivunen P, Rajaniemi H. 2001. Secretion of carbonic anhydrase isoenzyme VI (CA VI) from human and rat lingual serous von Ebner's glands. *Journal of Histochemistry & Cytochemistry* **49(5)**:657–662 DOI 10.1177/002215540104900513.

Leinonen JS, Saari KA, Seppanen JM, Myllyla HM, Rajaniemi HJ. 2004. Immunohistochemical demonstration of carbonic anhydrase isoenzyme VI (CA VI) expression in rat lower airways and lung. *Journal of Histochemistry & Cytochemistry* **52(8)**:1107–1112 DOI 10.1369/jhc.4a6282.2004.

Lupas A, Van Dyke M, Stock J. 1991. Predicting coiled coils from protein sequences. *Science* **252(5009)**:1162–1164 DOI 10.1126/science.252.5009.1162.

Medzhitov R. 2007. Recognition of microorganisms and activation of the immune response. *Nature* **449(7164)**:819–826 DOI 10.1038/nature06246.

Meijering E, Dzyubachyk O, Smal I. 2012. Methods for cell and particle tracking. *Methods in Enzymology* **504**:183–200 DOI 10.1016/b978-0-12-391857-4.00009-4.

Melis M, Atzori E, Cabras S, Zonza A, Calo C, Muroni P, Nieddu M, Padiglia A, Sogos V, Tepper BJ, Barbarossa IT. 2013. The gustin (CA6) gene polymorphism, rs2274333 (A/G), as a mechanistic link between PROP tasting and fungiform taste papilla density and maintenance. *PLOS ONE* **8(9)**:e74151 DOI 10.1371/journal.pone.0074151.

Muffato M, Louis A, Poisnel CE, Crollius HR. 2010. Genomicus: a database and a browser to study gene synteny in modern and ancestral genomes. *Bioinformatics* **26(8)**:1119–1121 DOI 10.1093/bioinformatics/btq079.

NCBI Resource Coordinators. 2016. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research* **44(D1)**:D7–D19 DOI 10.1093/nar/gkv1290.

Nishimori I, Minakuchi T, Onishi S, Vullo D, Cecchi A, Scozzafava A, Supuran CT. 2007. Carbonic anhydrase inhibitors: Cloning, characterization, and inhibition studies of the cytosolic isozyme III with sulfonamides. *Bioorganic & Medicinal Chemistry* **15(23)**:7229–7236 DOI 10.1016/j.bmc.2007.08.037.

Pan PW, Kayra K, Leinonen J, Nissinen M, Parkkila S, Rajaniemi H. 2011. Gene expression profiling in the submandibular gland, stomach, and duodenum of CAVI-deficient mice. *Transgenic Research* **20(3)**:675–698 DOI 10.1007/s11248-010-9441-2.

Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20(2)**:289–290 DOI 10.1093/bioinformatics/btg412.

Parkkila S, Kaunisto K, Rajaniemi L, Kumpulainen T, Jokinen K, Rajaniemi H. 1990. Immunohistochemical localization of carbonic anhydrase isoenzymes VI, II, and I in human parotid and submandibular glands. *Journal of Histochemistry & Cytochemistry* **38(7)**:941–947 DOI 10.1177/38.7.2113069.

Parkkila S, Parkkila A, Lehtola J, Reinila A, Sodervik H, Rannisto M, Rajaniemi H. 1997. Salivary carbonic anhydrase protects gastroesophageal mucosa from acid injury. *Digestive Diseases and Sciences* **42**:1013–1019.

Parkkila S, Parkkila AK, Vierjoki T, Stahlberg T, Rajaniemi H. 1993. Competitive time-resolved immunofluorometric assay for quantifying carbonic anhydrase VI in saliva. *Clinical Chemistry* **39**:2154–2157.

Patrikainen MS, Pan P, Barker HR, Parkkila S. 2016. Altered gene expression in the lower respiratory tract of *Car6*$^{-/-}$ mice. *Transgenic Research* **25(5)**:649–664 DOI 10.1007/s11248-016-9961-5.

Patrikainen M, Pan P, Kulesskaya N, Voikar V, Parkkila S. 2014. The role of carbonic anhydrase VI in bitter taste perception: evidence from the *Car6*$^{-/-}$ mouse model. *Journal of Biomedical Science* **21**:82 DOI 10.1186/s12929-014-0082-2.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25(13)**:1605–1612 DOI 10.1002/jcc.20084.

Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research* **29(9)**:e45 DOI 10.1093/nar/29.9.e45.

Pilka ES, Kochan G, Oppermann U, Yue WW. 2012. Crystal structure of the secretory isozyme of mammalian carbonic anhydrases CA VI: implications for biological assembly and inhibitor development. *Biochemical and Biophysical Research Communications* **419(3)**:485–489 DOI 10.1016/j.bbrc.2012.02.038.

Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European molecular biology open software suite. *Trends in Genetics* **16(6)**:276–277 DOI 10.1016/s0168-9525(00)02024-2.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61(3)**:539–542 DOI 10.1093/sysbio/sys029.

Roy A, Kucukural A, Zhang Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols* **5(4)**:725–738 DOI 10.1038/nprot.2010.5.

Salinas I, Zhang YA, Sunyer JO. 2011. Mucosal immunoglobulins and B cells of teleost fish. *Developmental and Comparative Immunology* **35(12)**:1346–1365 DOI 10.1016/j.dci.2011.11.009.

Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9(7)**:671–675 DOI 10.1038/nmeth.2089.

Shatzman AR, Henkin RI. 1981. Gustin concentration changes relative to salivary zinc and taste in humans. *Proceedings of the National Academy of Sciences of the United States of America* **78(6)**:3867–3871 DOI 10.1073/pnas.78.6.3867.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7(1)**:539 DOI 10.1038/msb.2011.75.

Sunyer JO, Zarkadis IK, Lambris JD. 1998. Complement diversity: a mechanism for generating immune diversity? *Immunology Today* **19(11)**:519–523 DOI 10.1016/s0167-5699(98)01341-3.

Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research* **34**:W609–W612 DOI 10.1093/nar/gkl315.

Thatcher BJ, Doherty AE, Orvisky E, Martin BM, Henkin RI. 1998. Gustin from human parotid saliva is carbonic anhydrase VI. *Biochemical and Biophysical Research Communications* **250(3)**:635–641 DOI 10.1006/bbrc.1998.9356.

van der Walt S, Colbert SC, Varoquaux G. 2011. The NumPy array: a structure for efficient numerical computation. *Computing in Science & Engineering* **13(2)**:22–30 DOI 10.1109/mcse.2011.37.

Vasta GR, Nita-Lazar M, Giomarelli B, Ahmed H, Du S, Cammarata M, Parrinello N, Bianchet MA, Amzel LM. 2011. Structural and functional diversity of the lectin repertoire in

teleost fish: relevance to innate and adaptive immunity. *Developmental and Comparative Immunology* **35(12)**:1388–1399 DOI 10.1016/j.dci.2011.08.011.

Webb B, Sali A. 2016. Comparative protein structure modeling using MODELLER. *Current Protocols in Protein Science* **86**:2.9.1–2.9.37 DOI 10.1002/cpps.20.

Westerfield M. 2007. *The Zebrafish Book: A Guide for the Laboratory Use of Zebrafish (Danio rerio)*. Eugene: M. Westerfield.

Xu J, Xu X, Wang B, Ma Y, Zhang L, Xu H, Hu Y, Wu J, Cao X. 2017. Nuclear carbonic anhydrase 6B associates with PRMT5 to epigenetically promote IL-12 expression in innate response. *Proceedings of the National Academy of Sciences of the United States of America* **114(32)**:8620–8625 DOI 10.1073/pnas.1700917114.

Zapata A, Amemiya CT. 2000. Phylogeny of lower vertebrates and their immunological structures. *Current Topics in Microbiology and Immunology* **248**:67–107 DOI 10.1007/978-3-642-59674-2_5.