Characterisation of trafficking signals shared by different types of exported proteins in the human malaria parasite *Plasmodium falciparum*

Dissertation

with the aim of achieving a doctoral degree
at the Faculty of Mathematics, Informatics and Natural Sciences
Department of Biology, University of Hamburg

submitted by

Ann-Katrin Ullrich

Hamburg, 2016

Dissertationsgutachter:	Dr. Tobias Spielmann
	Prof. Dr. Thorsten Burmester
Datum der Disputation:	24.06.2016

Declaration on oath

I hereby declare, on oath, that I have written the present dissertation by my own and have not used other than the acknowledged resources and aids.

_____

Ann-Katrin Ullrich

Hamburg, May 9, 2016

Language Certificate

I am a native speaker, have read the present PhD thesis and hereby confirm that it complies with the rules of the English language.

Toronto, April 25, 2016

Tatianna Wong

Summary

Malaria is still one of the most devastating diseases worldwide. Vaccine development of the first promising candidate is currently ongoing but its efficacy has already been shown to fall short of commercially available vaccines applied against other diseases. Beside vaccine development, widespread resistances against different antimalarial drugs challenge the attempts to control this important disease.

The most severe form of malaria is caused by the protozoan parasite *Plasmodium falciparum*. During the erythrocytic stage the intracellular parasite exports a large number of proteins beyond its own parasite specific boundaries into the host red blood cell (RBC). This leads to host cell modifications that include the formation of Maurer's clefts, membranous export structures in the host cell, and knobs, structures on the RBC surface. These structures are necessary for the survival and virulence of the parasite, as they enable the sequestration within the vasculature of the host. This avoids the clearance of the infected red blood cell (iRBC) by the spleen and this is believed to be a major reason for the malaria-associated pathology.

Two different types of proteins are exported – those that harbour a PEXEL motif (*Plasmodium* export element) and PEXEL negative exported proteins (PNEPs). PEXEL proteins are processed in the motif, an event essential for export. Thereafter, the new N-terminus containing a single conserved residue from the PEXEL motif takes over and guides further export. This region is comparable to the N-terminus of PNEPs, which also promotes export. The mature N-termini of PEXEL proteins and the N-termini of PNEPs are exchangeable and hence these two regions can be seen as a common core export domain of all exported proteins. However, due to (i) only a single conserved export promoting residue deriving from the PEXEL motif (here termed the 'primary export signal'), (ii) the lack of a comparable residue in many PNEPs and (iii) a clear indication for additional trafficking regions further downstream (here termed the 'secondary export signal'), the export promoting capacity and a possible consensus has so far remained elusive for this core export region.

In this thesis, sequences promoting export and sequences hampering export were assessed, both in naturally occurring N-termini of exported and non-exported proteins as well as in a neutral N-terminus that was specifically designed for this purpose. This

neutral N-terminus enabled the investigation of the impact on export of specifically inserted amino acids and their particular position within the N-terminus. By means of this neutral N-terminus, proline (P) was found to be a strong blocking signal. It was further revealed that the primary positive signal is much more flexible than expected from the PEXEL consensus it derives from. Beside the typical PEXEL residues glutamic acid (E) and glutamine (Q) – the amino acids annotated as the most frequent as a primary positive signal – for instance threonine (T) and leucine (L) also promoted high levels of export. In contrast aspartic acid (D), which is annotated the third most frequently found amino acid at this position in the PEXEL motif, did not promote export at all. Hence, the consensus of the PEXEL motif may need to be revisited.

The secondary positive signal was found to range widely both in sequence and position. No concrete sequence was uncovered to promote export as a secondary positive signal from a natural protein. However, it was found that the residues acting as the primary export signal in the neutral N-terminus could also serve as a secondary export signal if multiples of these residues were used, albeit always promoting lower efficiency of export than as a primary signal. Furthermore, it was uncovered that contrary to some previous hypotheses, negative charge of the amino acids alone was not decisive for export. As amino acids with a high α-helix- or β-turn-forming propensity were overrepresented in N-termini of exported proteins, and such residues generally promoted better export, specific secondary structures may be beneficial for export. Finally it was found that larger, rather than smaller amino acids better promoted export.

In conclusion this thesis provides data suggesting an increased plasticity of the primary export signal, uncovers the first sequence blocking export and informs on the amino acid residues that may contribute to the secondary positive export signal. The data however also indicate that the secondary export signal may be complex in nature and difficult to define in simple terms. Nevertheless, these results will form the basis for future attempts to predict the capacity of a region to act as the core export domain.

Zusammenfassung

Malaria ist immer noch eine der verheerendsten Krankheiten weltweit. Die Impfstoffentwicklung des ersten erfolgversprechenden Kandidaten läuft zurzeit, allerdings wurde bereits gezeigt, dass die Wirksamkeit des Impfstoffes, im Vergleich zu sonst kommerziell eingesetzten Vakzinen gegen andere Krankheiten, weit zurückbleibt. Neben der schlechten Impfstofflage hemmen weit verbreitete Resistenten des Erregers gegen die unterschiedlichen Antimalaria-Wirkstoffe die Kontrolle dieser bedeutenden Krankheit.

Die schwerste Form der Malaria wird durch den Protozoen *Plasmodium falciparum* hervorgerufen. Während der erythrozytären Phase exportiert der intrazelluläre Parasit eine erhebliche Anzahl an Proteinen über seine eigenen Parasiten-spezifischen Grenzen hinaus in die Wirtszelle, die rote Blutzelle. Die dadurch hervorgerufenen Modifikationen umfassen die Bildung von Maurer's clefts, membranösen Exportstrukturen in der Wirtszelle, und Knobs, Strukturen an der Oberfläche der roten Blutzelle. Diese Strukturen sind erforderlich für das Überleben und die Virulenz des Parasiten, da sie die Sequestrierung der infizierten roten Blutzellen innerhalb der Blutgefäße des Wirtes ermöglichen und so die Beseitigung der infizierten Erythrozyten durch die Milz verhindern. Die Sequestrierung wiederum wird als wichtiger Grund für Malaria-assoziierte Pathologie angenommen.

Zwei unterschiedliche Typen von Proteinen werden exportiert – solche die ein PEXEL-Motiv (*Plasmodium* export element) aufweisen und PEXEL negative exported proteins (PNEPs). PEXEL-Proteine werden innerhalb des Motivs prozessiert, ein essentielles Ereignis für den Proteinexport. Anschließend übernimmt der neue N-terminus, welcher nur noch eine konservierte Aminosäure des PEXEL-Motivs enthält, den weiteren Export. Der N-terminus der prozessierten PEXEL-Proteine ist vergleichbar mit dem N-terminus von PNEPs, der ebenfalls Proteinexport vermittelt. Diese beiden Regionen sind austauschbar in ihrer Fähigkeit Export zu vermitteln, weshalb diese beiden Bereiche als eine gemeinsame Exportdomäne aller exportierten Proteine angesehen werden können. Bisher ist jedoch nur die konservierte Export-vermittelnde Aminosäure des PEXEL-Motivs in dem Bereich bekannt (welche von nun an als ‚primäres Exportsignal' bezeichnet wird). Solch eine vergleichbare Aminosäure fehlt bei vielen PNEPs. Außerdem gibt es deutliche Hinweise, dass sowohl der nach der Prozessierung

entstandene PEXEL N-terminus wie auch die PNEP N-termini, Sequenz-abwärts weitere Exportsignale enthalten (von nun an als ‚sekundäre Exportsignale' bezeichnet). Aufgrund dieser Umstände war bisher sowohl die Export-vermittelnde Eigenschaft als auch ein möglicher Konsensus der gemeinsamen Exportdomäne schwer zu definieren.

In dieser Arbeit wurden Export-fördernde und -blockierende Sequenzen untersucht, sowohl von natürlich vorkommenden N-termini exportierter und nicht-exportierter Proteine, als auch in einem neutralen N-terminus, welcher eigens für diesen Zweck entwickelt wurde. Dieser neutrale N-terminus ermöglicht es, die Auswirkung eingefügter Aminosäuren (und deren spezifische Position) innerhalb des N-terminus auf den Export zu untersuchen. Mithilfe dieser neutralen Region konnte gezeigt werden, dass Prolin (P) ein starkes Export-blockierendes Signal ist. Ferner wurde deutlich, dass das primäre Exportsignal weitaus flexibler ist als durch den Konsensus des PEXEL-Motivs hätte angenommen werden können. Neben den klassischen PEXEL-Aminosäuren Glutaminsäure (E) und Glutamin (Q) – die Aminosäuren, die am häufigsten als primäre Exportsignale annotiert sind – rufen beispielsweise auch Threonin (T) und Leucin (L) hohe Exportlevels hervor. Asparaginsäure (D) hingegen, annotiert als die dritthäufigste Aminosäure an der entsprechenden Position innerhalb des PEXEL-Motivs, resultierte nicht in Export. Somit ist vermutlich eine neue Analyse des PEXEL-Konsensus' notwendig.
Das sekundäre Exportsignal war sehr divers, sowohl die Sequenz als auch die Position betreffend. Deswegen konnte keine Sequenz eines natürlichen Proteins ermittelt werden, die Export als ein sekundäres Exportsignal hervorruft. Allerdings konnte aufgedeckt werden, dass Aminosäuren, die als primäres Exportsignal fungieren können, ebenfalls als sekundäres Exportsignal wirken können, wenn mehrere dieser Aminosäuren eingesetzt werden. Die Exportrate war bei diesen Proteinen allerdings immer niedriger als beim Einsatz der jeweiligen Aminosäure als primäres Exportsignal. Ferner konnte gezeigt werden, dass – konträr zu einigen früheren Hypothesen – die negative Ladung einer Aminosäure nicht alleine ausschlaggebend für Proteinexport ist. Da Aminosäuren mit hohen α-helix- oder β-turn formenden Eigenschaften in N-termini exportierter Proteine überrepräsentiert sind (und solche Aminosäuren meist bessere Exportraten hervorgerufen haben), sind bestimmte Sekundärstrukturen eventuell förderlich für Proteinexport. Des Weiteren schienen größere, eher als kleinere, Aminosäuren besseren Export zu vermitteln. Zusammenfassend stellt diese Thesis

Daten bereit, die die Plastizität des primären Exportsignals deutlich erweitern. Außerdem deckt sie die erste Export-blockierende Sequenz auf und zeigt die Aminosäuren auf, die eventuell innerhalb des sekundären Exportsignals mitwirken, wobei dieses Signal in natürlichen Proteinen sehr komplex und schwer definierbar zu sein scheint. Dennoch bilden diese Daten die Voraussetzung für zukünftige Vorhersagen der Exportkapazität der gemeinsamen Exportdomäne.

Table of contents

List of figures

Abbreviations

%	Per cent
α	Alpha/anti
A	Adenine
A, Ala	Alanine
aa	Amino acid(s)
ACT	Artemisinin based combination therapy
ad	Add to
ATG	Methionine, start codon
ATP	Adenosine triphosphate
β	Beta
BLAST	"Basic Local Alignment Tool"
BNI	Bernhard Nocht Institute
bp	Base pair (s)
°C	Degrees Celsius
C	Cytosine
C, Cys	Cysteine
C-	Carboxy
cDNA	Complementary DNA
cm	Centimetre
CO_2	Carbon dioxide
crt	Chloroquine Resistance Transporter
D	Dimension
D, Asp	Aspartic acid
Da	Dalton
DAPI	4´6-Diamino-2-phenylindol
DBL	"Duffy binding-like"
DHFR	Dihydrofolate reductase
dH_2O	Distilled water
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribose nucleoside triphosphate
DDT	dichlorodiphenyltrichloroethane
E, Glu	Glutamic acid
ECL	Enhanced chemiluminescence
E. coli	*Escherichia coli*
e.g.	*Exempli gratia* (for example)
EGTA	Ethylene Glycol Tetraacetic Acid
ER	Endoplasmic Reticulum
et al.	*Et alii* (and others)
ETRAMP	Early transcribed membrane protein
EXP	Exported protein
F, Phe	Phenylalanine
F	Farad
Fig.	Figure
fw	Forward
g	Gram; Gravitational force
GBP	Glycophorin-binding protein

GFP	Green Fluorescent Protein
G	Guanine
G, Gly	Glycine
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
gDNA	Genomic DNA
GRASP	Golgi Reassembly Stacking Protein
h	Hour
H, His	Histidine
hDHFR	Humane Dihydrofolat Reductase
H_2O	Water
HRP	Horseradish peroxidase
HSP	Heat shock protein
I, Ile	Isoleucine
i.e.	Id est (that is to say)
in vitro	Outside of the living
in vivo	Within the living
in silico	In silicon (performed on computer/via computer simulation)
k	Kilo
K, Lys	Lysine
KAHRP	Knob-associated histidine-rich protein
kb	Kilo base
l	Litre
L, Leu	Leucine
LB	Lysogeny Broth
µ	Micro
µm	Microliter
µm	Micrometre
M	Mole
M	Mega
M, Met	Methionine
m	Milli
MAHRP	Membrane Associated Histidine
MCs	Maurer's clefts
MCS	Multiple cloning site
min	Minute
ml	Millilitre
mDHFR	Mouse dihydrofolat reductase
mTRAP	Merozoite TRAP
n	Nano
N, Asn	Asparagine
N-	Amino-
N_2	Nitrogen gas
NCBI	National Center for Biotechnology Information
NEB	New England Biolabs
O_2	Oxygen
OD	Optical density
ON	Over night
P.	*Plasmodium*
P, Pro	Proline

PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEXEL	*Plasmodium* Export Element
Pf	*Plasmodium falciparum*
PfEMP1	*Plasmodium falciparum* Erythrocyte Membrane protein-1
P. falciparum	*Plasmodium falciparum*
P. knowlesi	*Plasmodium knowlesi*
PlasmoDB	Plasmo database (database for *Plasmodia*)
P. malariae	*Plasmodium malariae*
P. ovale	*Plasmodium ovale*
P. vivax	*Plasmodium vivax*
pH	Potentia Hydrogenii (potential of hydrogen)
PTEX	*Plasmodium* translocon for exported proteins
PNEP	PEXEL negative exported protein
PV	Parasitophorous vacuole
PVM	Parasitophorous vacuole membrane
Q, Gln	Glutamine
rev, rv	Reverse
R, Arg	Arginine
RBC	Red blood cell
REX	Ring Exported Protein
RNA	Ribonucleic acid
rpm	Rounds per minute
RPMI	Roswell Park Memorial Institute
RT	Room temperature
s, sec	Second
S, Ser	Serine
SBP	Skeleton Binding Protein
SERA	Serine repeat antigen
SP	Signal peptide
sp.	*Species*, singular
spp.	*Species*, plural
ssp.	*Subspecies*
T	Thymine
T, Thr	Threonine
Tg, *T. gondii*	*Toxoplasma gondii*
TM	Transmembrane domain
TRAP	Thrombospondin Related Adhesion Protein
TRX2	Thioredoxin 2
TVN	Tubovesicular network
U	Units
UTR	Untranslated region
UV	Ultraviolet
V, Val	Valine
V	Volt
Vol.	Volume
VTS	Vacuolar Targeting Signal
v/v	Volume/volume
W, Trp	Tryptophan

WHO	World Health Organization
WT	Wild type
w/v	Weight/Volume
Y, Trp	Tyrosine

1 Introduction

1.1 Malaria

Malaria is, besides AIDS and tuberculosis, considered to be one of the most important infectious diseases worldwide. The name malaria derives from the Latin words 'mala aria', which means bad air. This designation was based on the belief that the disease is caused by bad air, rising from swamps (Capanna, 2006). Alphonse Laveran firstly described the parasite in 1880. Seventeen years later Ronald Ross discovered that the pathogen was transmitted via mosquitoes. Since then, immense progress has been made in our understanding of the parasite and in the control of this disease, but nevertheless it remains a major problem in many areas of the world.

1.1.1 Epidemiology

Around 1900 malaria was spread over the entire world, except for Canada, Greenland, Argentina, Chile and Australia (Hay *et al*., 2004). In the 1960s the eradication of malaria succeeded in Europe through the draining of swamps and the systematic use of insecticides (de Zulueta, 1998) but the disease is still widespread in many areas of the world. Almost half of the world's population still lives in regions at risk of malaria infection (Figure 1.1). In 2015 around 214 million malaria cases were reported which led to 438.000 deaths. Around 90% of the cases occurred in sub-Saharan Africa (WHO, 2015). Nearly 70% of the deaths caused by malaria are children under five years of age (WHO, 2015). Of the remaining deaths, many are pregnant women and immunosuppressed people (Gallup & Sachs, 2001; Greenwod *et al*., 2008; Wellems *et al*., 2009; WHO, 2015). Nevertheless, since 2000, the infection rates have been reduced up to 60% due to better prevention and control measurements (WHO, 2016; see section 1.1.3).

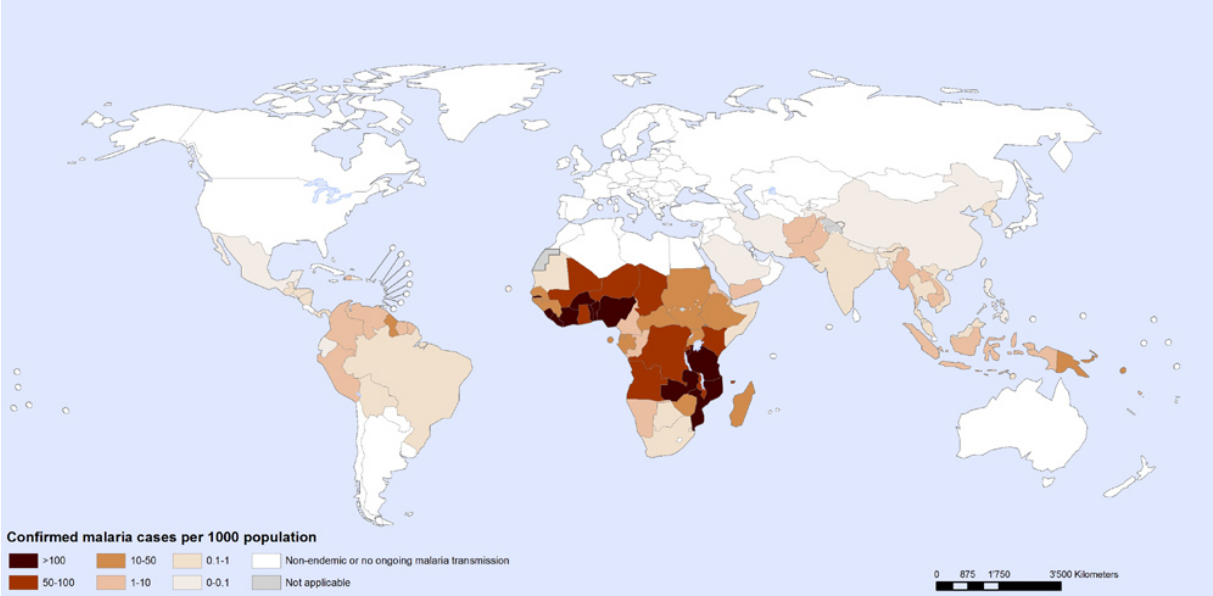


Figure 1.1 – Countries with ongoing transmission of malaria in 2013 (WHO malaria report, 2014). Malaria occurs mainly in Africa, other regions are Central and South America as well as Asia and the West Pacific Ocean. The map shows regions with a high ongoing malaria transmission rate (dark brown) to no transmission rate at all (white).

1.1.1.1 The pathogen *Plasmodium*

Malaria is caused by vector-borne pathogens of the genus *Plasmodium*. This protozoan parasite belongs to the phylum of the Apicomplexa. Other important members of the Apicomplexa are *Toxoplasma*, *Babesia* and *Cryptosporidium* (Escalante & Ayala, 1995; Wellems *et al*., 2009). This phylum is characterised by an apical complex (see section 1.2.2). Around 200 different *Plasmodium* species are known (Dhangadamajhi *et al*., 2010). Five of these species can infect humans – *Plasmodium falciparum, Plasmodium vivax, Plasmodium ovale, Plasmodium malariae and Plasmodium knowlesi*. *P. falciparum* and *P. vivax* together are responsible for around 95% of all reported malaria cases (Garcia, 2010). *P. vivax* is widely distributed in Asia, the West Pacific Ocean and in Central and South America (Mendis *et al*., 2001). *P. knowlesi* is found in Southeast Asia (Singh & Daneshvar, 2013). In comparison to the other human-infecting species, the infection with *P. knowlesi* is rare. The other three human infecting parasite species *P. falciparum*, *P. malariae* and *P. ovale* can be found in the tropical regions (Snow *et al*., 2005; Collins & Jeffery, 2005; Collins & Jeffery, 2007).

1.1.1.2 Malaria transmission

The transmission *of Plasmodium* spp. occurs via female *Anopheles* mosquitoes during a

blood meal. As the male *Anopheles* mosquitoes feed on nectar and other sugar sources only, they do not transmit malaria parasites. Females instead do need sugar as an energy source but also proteins of blood for the development of their eggs (Tahir *et al*., 2003). *Anopheles* mosquitoes develop in stagnant water from an egg to a larva and a pupa, which takes eleven days to four weeks – depending on the species and the ambient temperature. The final adult phase in nature lasts for one week for the males and two weeks for the females (CDC, 2012; WHO, 2014). *Anopheles* mosquitoes can be distinguished from other genera by their characteristic posture, as their bodies are mostly positioned in a 30 to 45 degree angle to the ground (CDC, 2015). Between 30 and 40 *Anopheles* species are known to transmit *Plasmodium* parasites to humans. *Anopheles gambiae* is the most important of them in the African region (CDC, 2015). *P. knowlesi* – in contrast to all the other human infecting *Plasmodium* species – can only be transmitted by *Anopheles leucosphyrus* (Cox-Singh & Singh, 2008).

1.1.2 Clinic

After the bite of an *Anopheles* mosquito and the transmission of parasites to a human, first an asymptomatic phase takes place. During this phase the parasites replicate within the liver of their host (see section 1.2.1). Subsequently the erythrocytic phase takes place, where the parasites develop within red blood cells (see section 1.2.2; CDC, 2013). These blood stage parasites cause the symptoms of malaria. The first symptoms can occur a few days up to one year after infection, depending on the *Plasmodium* species. They can vary from headaches, fever and chills to convulsions and vomiting and range from a mild to a heavy form. Severe symptoms of malaria are splenomegaly, anaemia, cerebral malaria, prostration and acidosis-driven respiratory distress. Adults can also suffer from multiple organ failure (Greendwood *et al*., 2008; Wellems *et al*., 2009; WHO, 2014). The severity of the disease depends on different factors, for example the parasite species and the age, immunity status and genetic background of the host (Greenwood *et al*., 2008; Wellems *et al*., 2009; WHO, 2014).

1.1.2.1 The different types of Malaria

The five human infecting species of *Plasmodium* are clinically distinguishable due to their blood stage cycle time and their clinical picture. *Plasmodium vivax*, *ovale* and

malariae show a synchronous development cycle in the human blood. The newly formed parasites are released synchronously which leads to fever attacks of the patient. *P. vivax* and *P. ovale* preferentially infect young red blood cells, the so called 'reticulocytes' (McKenzie *et al*., 2001; McKenzie *et al*., 2002). The infection of reticulocytes is limited due to their low availability what leads to a comparably low parasitaemia – the percentage of infected compared to not infected red blood cells. These two species cause the Malaria tertiana, which received its name because of the continuous fever attacks reoccurring in a three day rhythm coinciding with the 48 hour blood stage cycle of these species (McKenzie *et al*., 2002; Collins & Jeffery, 2005; Collins & Jeffery, 2007). *P. vivax* and *P. ovale* infections often cause mild disease, although *P. vivax* was also reported to cause severe disease (Mueller *et al*., 2007; Naing *et al*., 2014). The ability of *P. vivax* and *P. ovale* to form hypnozoites is special for these species. Hypnozoites are parasite forms that persist in a dormant mode in liver cells until they become activated again and cause a relapse of the disease (Shute, 1946; Markus, 1976; Cogswell, 1992).

Plasmodium malariae mainly infects senescent red blood cells (McKenzie *et al*., 2001; McQueen & McKenzie, 2004). This species causes the Malaria quartana. Its lifecycle takes 72 hours, therefore two fever free days follow each fever attack. The course of disease is usually mild. This species can also cause relapses but this is due to persistent erythrocytic forms not hypnozoites. These relapses can occur up to 50 years after infection (Collins & Jeffery, 2007).

Plasmodium knowlesi is actually a macaque-infecting parasite and therefore does not belong to the classical human infecting group of *Plasmodium* species. It was firstly diagnosed 1965 to infect humans (Chin *et al*., 1965). Its blood stage cycle takes only 24 hours and is synchronous which leads to daily fever attacks. As it causes severe forms of malaria, its diagnostic discrimination, especially from *P. vivax* – which is closely related to *P. knowlesi* – is very important. As a human to human transmission is very rare, the infection of *P. knowlesi* is categorised as a zoonosis (Cox-Singh & Singh, 2008).

The most important human infecting species is *Plasmodium falciparum*. It causes the most severe form of malaria, the Malaria tropica or falciparum malaria. Its asynchronous blood stage cycle, which takes 48 hours, usually does not cause regular fever attacks like the other human infecting species but constant fever (Arnot & Gull, 1998; Gerald *et al*.,

2011). As *P. falciparum* parasites infect red blood cells of every age, the parasitaemia can get very high. Therefore anaemia is one of the major symptoms of Malaria tropica. Further, infected red blood cells (iRBC) adhere to the endothelium of blood vessels of major organs (see section 1.3.1) and can therefore escape the clearance by the spleen. The sequestration of iRBCs within the blood vessels of the brain can lead to blocked capillaries and inflammation in the brain, causing cerebral malaria. This is a major complication caused by *P. falciparum* and can induce coma, brain injuries and, in progress, death (Trampuz *et al*., 2003; Josling & Llinás, 2015).

1.1.2.2 Diagnosis

A fast and effective diagnosis of malaria is important as this reduces complications as well as the mortality rate. But due to the different possible clinical presentations of the disease and differing treatment indications, it is critical to determine the infecting *Plasmodium* species, a task that can be difficult (Tangpukdee *et al*., 2009). In addition, an infection with more than one *Plasmodium* species is frequent and further hampers the correct diagnosis. The discrimination of *P. vivax* and *P. ovale* from other *Plasmodium* species is possible based on so called 'Schüffner's dots' present in the infected red blood cells of Giemsa stained blood smears. These dots are characteristic for these two species (Udagama *et al*., 1988; Coatney *et al*, 1971). *P. falciparum* also shows dotty structures when investigated by microscopy. In this case these structures are called 'Maurer's clefts' (see section 1.3.1.1). Until now no systemic comparison to Schüffner's dots was conducted (Mundwiler-Pachlatko & Beck, 2013). Differentiation of all *Plasmodium* species is possible by microscopy of Giemsa stained thin blood smears which is the gold standard in Malaria diagnosis. In terms of *P. falciparum* only so called 'ring stage' parasites can be detected in blood smears (see section 1.3.1). Different *Plasmodium* species can also be discriminated from one another by PCR analysis (Johnston *et al*., 2006). The detection of *Plasmodium* specific antigens is another tool for the diagnosis of a malaria infection.

1.1.3 Vector control, drugs and malaria vaccines

1.1.3.1 Vector control

As malaria occurs predominately in low income regions that are often difficult to access (WHO, 2016), cheap and easily available control strategies are needed. In the middle of the 20th century Dichlordiphenyltrichlorethane (DDT) was used to reduce mosquito populations to combat malaria. This insecticide was – in combination with the drying of swamps – very efficient in reducing malaria infections (Beard, 2006). After its commercial availability in 1945 (WHO, 1979), DDT was used extensively worldwide until the mid 1960s. By then the eradication of malaria from the United States and Europe through the use of the insecticide succeeded (Attaran & Maharaj, 2000). However, due to arising health issues of exposed people and resistances of mosquitoes, the use of DDT was lowered in the 1970th (Beard, 2006). Today it is applied only in small doses in Africa (WHO, 2006). The most successful vector control strategy to date is the use of insecticide-treated mosquito bed nets (ITNs) that prevent the transmission of the parasite from the mosquito to the human host (Binka & Akweongo, 2006). An additional protection through combining indoor residual spraying (IRS) and ITNs may be beneficial but was not finally proven so far (WHO, 2014).

1.1.3.2 Drugs

Malaria patients can be treated with different kinds of drugs, depending on the parasite species and the severity of the disease. Quinine was the first drug for malaria treatment. It was discovered in the 17th century (Achan *et al*., 2011) and it is said that British colonialists drunk quinine-carrying Tonic Water to avoid malaria infections. They mixed it with gin to bypass the bitter taste and thereby invented the Gin Tonic (Meyer *et al*., 2004). In 1820 two French chemists managed to isolate quinine from the cinchona bark (Achan *et al*., 2011). From then on the substance was used frequently to treat malaria. In 1934 the synthesis of chloroquine was started (Coatney, 1963), which is closely related to quinine. It belonged to a new class of antimalarials, the four-amino quinolones (Bray *et al*., 1996), which was invented by German scientists (Krafts *et al*., 2012). Firstly categorised as too toxic, the drug was ignored for a decade until further inventions showed the high impact as an antimalarial drug (Krafts *et al*., 2012). Mefloquine, which is structurally related to quinine, was introduced in the early seventies (Croft, 2007). It

was principally used as a prophylactic because of its long half life (Schlagenhauf *et al*., 2010). In 1972 artemisinin was isolated from *Artemisia annua* (sweet wormwood) by Chinese scientists (Tu, 2011). It was already known from Chinese writings over 2000 years ago that this plant was used to cure patients with characteristic fever attacks (Cox, 2002; Dondorp *et al*., 2009). After a long process of investigating, artemisinin finally found commercial use. It is a fast killing agent with a very short plasma half life and by that reduces the risk for the development of resistant parasites (Vries & Dien, 1996). Artemisinin shows the highest recorded parasite killing effect of all antimalarial drugs applied until today (White, 2004).

1.1.3.3 Drug resistance and the onset of ACT

Constant invention of new drugs was necessary because resistances of the parasites occurred to all of the so far introduced drugs. The first resistances to chloroquine, which was extensively used, were documented around 1960 (Wellems & Plowe, 2001).

The treatment of malaria with artemisinin as the first line drug started in the 1990s. Already in the late nineties an artemisinin based combination therapy (ACT) was introduced and it was hoped that this would prevent the development of resistances in the parasite. For this combination therapy, artemisinin (e.g. the derivate Artesunate) was given in combination with a partner drug, for instance mefloquine. Mefloquine persists much longer than artemisinin at active concentrations in the human body due to its half life of two to four weeks. Thus this second drug eliminates any parasites remaining after the artemisinin levels drop to sublethal levels, decreasing the likelihood of emerging resistant parasites (Eastman & Fidock, 2009; Frey *et al*., 2010).

A second advantage of an artemisinin based therapy is the clearance of sexual stages of the pathogen, in addition to the elimination of the asexual stages (Eastman & Fidock, 2009; Frey *et al*., 2010). As the WHO recommends ACT, most countries switched to this as their standard treatment (WHO, 2006). Nevertheless, first resistances against artemisinin have recently been confirmed in five countries – Cambodia, the Lao People's Democratic Republic, Myanmar, Thailand and Vietnam – all in the Greater Mekong subregion (WHO, 2015; Winzeler & Manary, 2014).

1.1.3.4 Vaccines

A key control strategy against infectious diseases is the use of vaccines. However, despite great efforts to develop a malaria vaccine, no vaccine is at present commercially available and clinical trials have so far led to mostly sobering results. The reasons for this are the high complexity of the disease and the fact that many *P. falciparum* proteins, suitable as vaccine candidates, are highly polymorphic (Sutherland, 2007).

Various vaccine candidates from many different phases of the parasite life cycle have already been investigated during the last decades. The vaccine candidate RTS,S/AS01 (RTS,S) is now the first agent that made it to clinical trial IV. It was developed in the late 1980s by researchers of GlaxoSmithKline in cooperation with the PATH malaria vaccine initiative and is the first licensed vaccine against a parasitic disease ever. It acts by inducing humoral and cellular immunity that prevents the infection of liver cells by the parasite (Foquet *et al*., 2014). The vaccine, tested in children that are the most affected group, showed an efficacy of only 27-39%, depending on the age of the children (WHO, 2016). Nevertheless this vaccine is already a milestone and the WHO recommended a first pilot implementation (WHO, 2016). RTS,S is also considered the initial point for the development of potential further vaccine candidates that may show a better efficacy.

1.2 Cell biology of malaria

1.2.1 Life cycle of *Plasmodium falciparum*

The lifecycle of the five human infecting *Plasmodium* species is complex but comparable (Figure 1.2). It differs only in the duration of the different phases. The first part of the lifecycle is the sexual phase, the sporogony, which takes place in the mosquito. With a bite of an *Anopheles* mosquito so called 'gametocytes' are taken up together with the blood of the human. Gametocytes are the sexual stages of *Plasmodium* spp. With the blood meal they enter the midgut of the mosquito and further develop to gametes, leading to the formation of female (macrogamete) and male (microgamete) stages that upon fertilisation form a motile parasite stage termed 'ookinete'. The ookinete penetrates the gut wall and under the basal lamina of the midgut epithelium develops to an oocyst. This oocyst then ruptures and releases up to 1.000 so called 'sporozoites' which migrate to the mosquito's hemolymph to invade the salivary gland epithelium. With another bite of the mosquito several hundred sporozoites are transmitted with the

saliva into the dermis of the human host. In the dermis sporozoites find blood vessels and enter the blood stream to reach the liver. Through invasion of liver cells, the hepatocytes, the liver phase begins. This so called 'exoerythrocytic schizogony' is the first of two asexual multiplications in the human host. Within the liver cell the parasite develops to a liver schizont, reproduces asexually and forms up to 30.000 daughter cells termed ‘merozoites’. Under destruction of the host hepatocyte, these merozoites are released into the sinusoid lumen in parasite-filled vesicles called merosomes. By that the parasite bypasses the recognition through the immune system and ensures the release of merozoites into the bloodstream (Sturm *et al*., 2006). The merosomes then rupture to release merozoites that now invade red blood cells (RBC) to initiate the second asexual reproduction phase termed the ‘erythrocytic schizogony’. Within the RBC the parasites develop in several developmental stages (see section 1.2.2) to merozoites again, which are released and invade new RBCs, leading to a continuous multiplication of the parasite in the blood. A few merozoites, upon entry into new RBCs, develop to gametocytes instead of developing into asexual stages. Through a mosquito bite these gametocytes can be taken up and the lifecycle of the parasite starts again (Greenwood *et al*., 2008).

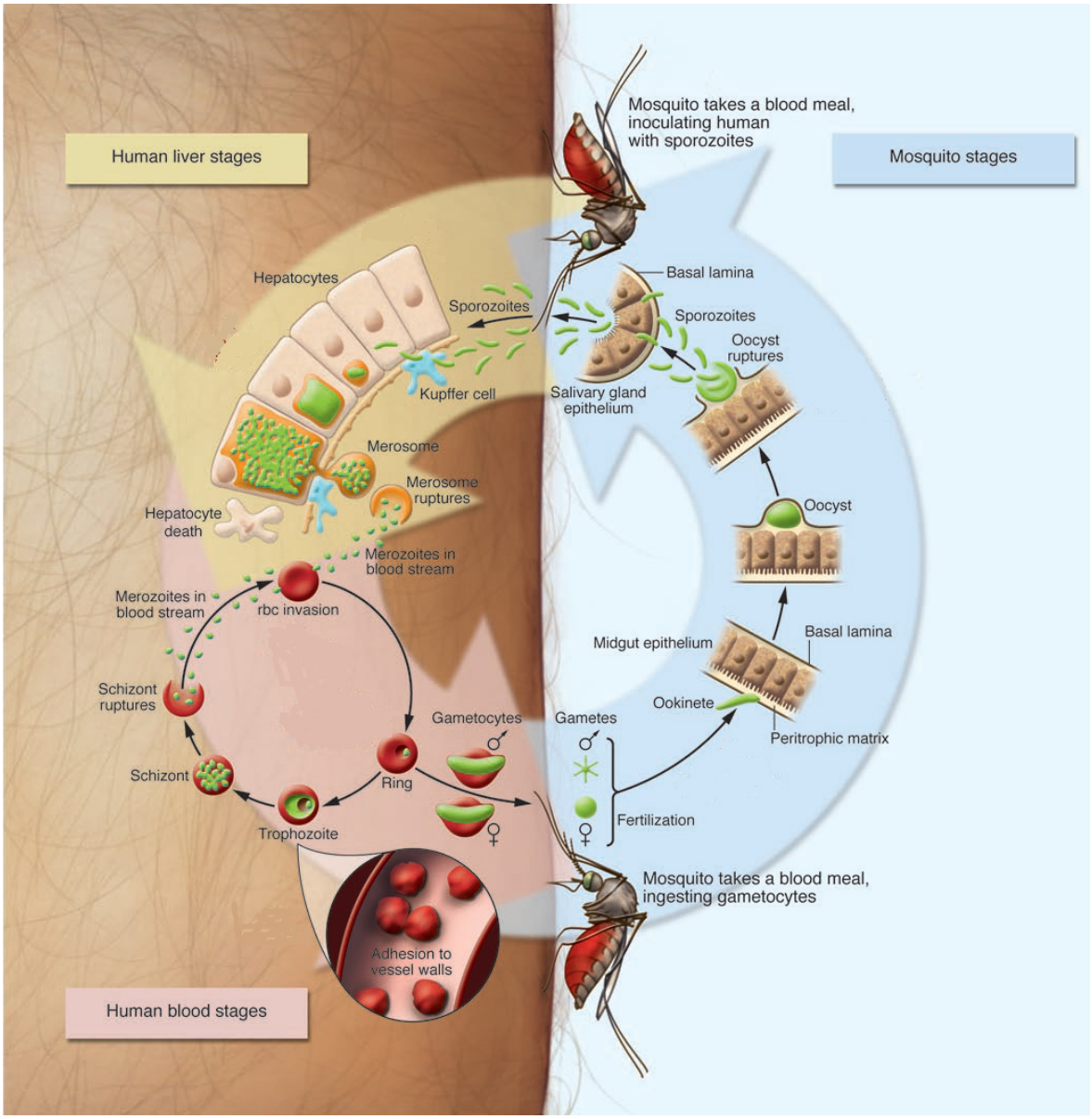


Figure 1.2 – The lifecycle of *Plasmodium*. Gametocytes (parasites are shown in green) are taken up by an *Anopheles* mosquito and develop within the mosquito to sporozoites. Through another mosquito bite these sporozoites are transferred to the human host, develop within the liver to merozoites and are released into the blood stream where they invade RBCs. Here new merozoites are formed and released. A few of the merozoites develop into gametocyctes and can be ingested by the mosquito through another blood meal. Modified from Greenwood *et al.*, 2008.

1.2.2 The erythrocytic lifecycle

The human blood stage (Figure 1.3) is the life cycle stage of malaria parasites that causes the symptoms of malaria in the host. This phase starts with the first contact of a RBC with a merozoite. The merozoite binds to the RBC plasma membrane and re-orientates, so that the apical end is in contact with the erythrocyte. The invasive forms of the *Apicomplexa* possess secretory organelles. These organelles are situated at the apical end of the merozoites, forming the apical complex that is important for the invasion

process (Wright & Rayner, 2014). These organelles comprise the rhoptries, micronemes and dense granules. In addition an apical polar ring acts as a microtubule organising centre (Sam-Yellowe, 1996). During invasion, the rhoptries and the micronemes secrete proteins to form a tight junction complex at the boundary point to the RBC (Hanssen *et al*., 2013; Riglar *et al*., 2011; Singh *et al*., 2010). The tight junction is passed around the surface of the invading merozoite, appearing like a belt around the parasite. This inclusion is driven by an actin-myosin motor, anchored to the inner membrane complex (IMC) of the merozoite. In addition to its role in invasion, the IMC is also responsible for the typical ovoid shape of the merozoite (Baum *et al*., 2006; Jones *et al*., 2006). Finally the moving junction closes behind the merozoite. With this the merozoite is enveloped by the host cell membrane and forms a new compartment – the parasitophorous vacuole (PV) with its parasitophorous vacuole membrane (PVM; Wright & Rayner, 2014). The whole process – from merozoite release until completed invasion – takes less than two minutes (Glushakova *et al*., 2005; Gilson & Crabb, 2009).

After the merozoite entered the RBC the erythrocytic schizogony starts. The freshly invaded parasite loses its apical apparatus and transforms from an egg-shaped into a round-shaped appearance. In Giemsa stained thin blood smears the parasite appears as a (signet) ring when investigated by light microscopy. This stage is therefore called the 'ring stage' of the parasite. It takes approximately 24 hours. After that the parasite develops to a so called 'trophozoite'. This stage is mainly characterised by rapid growth of the parasite and takes 12 to 14 hours. Finally the parasite develops into a so called 'schizont' (Bannister *et al*., 2000). During this stage the parasite multiplies its DNA, undergoes segmentation and forms up to 32 daughter merozoites (Maier *et al*., 2009). This phase takes 8 to 10 hours. After rupture of the schizont, the newly formed merozoites are released and can invade new RBC (Cowman & Crabb, 2006). This completes one round of the continuous asexual blood cycle which lasts 48 hours in *P. falciparum* (Bannister *et al*., 2000; Maier *et al*., 2009).

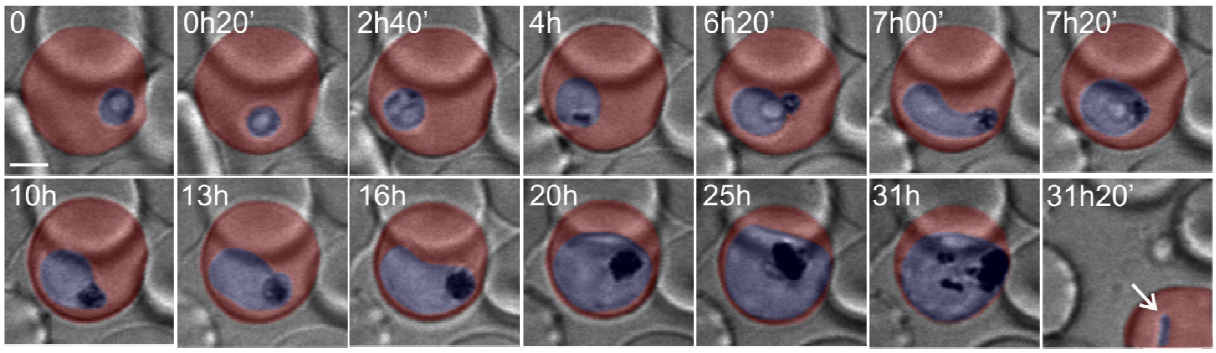


Figure 1.3 – Selected points in time during the erythrocytic cycle of *P. falciparum*. Different stages of the same developing parasite collected by time lapse microscopy are shown. After invasion of the RBC (red) the merozoite (blue) develops to a ring stage parasite. Point 0 already shows a late ring stage parasite. It further develops to a trophozoite (2h40' – 16h) and matures to a schizont (20h – 31h). With completion of the cycle the schizont ruptures and the newly formed merozoites invade new RBCs (31h20', white arrow shows a newly invaded RBC by a merozoite). Modified from Grüring *et al.*, 2011.

1.3 Protein export

The major function of RBCs is to carry haemoglobin and therefore oxygen through the body of the human. The erythrocyte is an ideal niche for *Plasmodium* parasites as it lacks organelles such as the nucleus and further lacks expression of the major histocompatibility complex. This provides some protection of the infected cell from immune recognition. As a disadvantage, the RBC also lacks a protein trafficking system which has to be installed *de novo* by the parasite.

1.3.1 Host cell modifications

For installing a protein trafficking system, about 10% of all *Plasmodium falciparum* proteins are exported (Spielmann & Gilberger, 2015). Hence the parasite extensively modifies its host red blood cell during the erythrocytic stage (Deitsch & Wellems, 1996; Kyes *et al.*, 2001; Rowe & Kyes, 2004). Establishing such an exomembrane system in the host cell cytosol is a prerequisite for the trafficking of proteins to the host cell surface but also potentially for the uptake of nutrients (Lauer *et al.*, 1997; Hanssen *et al.*, 2010). The introduced host cell changes can range from newly synthesised channels for the uptake and disposal of proteins and nutrients (Saliba *et al.*, 1998; Desai *et al.*, 2000; Staines *et al.*, 2004) over the change in ion channel properties (Decherf *et al.*, 2004; Staines *et al.*, 2007) and membrane rigidity of the iRBC (Glenister *et al.*, 2002) to the formation of knobs on the host cell surface (Trager *et al.*, 1966; Leech *et al.*, 1984). Among other things the formation of knobs is a crucial factor in parasite virulence.

These structures are formed by different exported proteins like KAHRP (knob associated histidine-rich protein) which self-associate on the cytoplasmic side of the host cell and lead to cup-shape structures bulging out the host cell membrane. The cytoplasmic domain of the major virulence factor PfEMP1 (*P. falciparum* erythrocyte membrane protein 1) is anchored in the knobs via an interaction with KAHRP. The amino-terminal domain of PfEMP1 is exposed at the outer membrane of the host cell and can interact with host cell receptors (Oh *et al*., 2000; Maier *et al*., 2009; Rowe *et al*., 2009) leading to cytoadherence to blood vessels and sequestration of infected erythrocytes in the vasculature of the major organs of the host (Miller *et al*., 2002; Rowe *et al*., 2009). In addition this can lead to the association with uninfected RBCs, which is called 'rosetting' (Miller *et al*., 1994). The degree of sequestration depends on the expression of parasite proteins on the host cell surface, which is directly connected to parasites survival and virulence (Crabb *et al*., 1997; Horrocks *et al*., 2005; Rug *et al*., 2006).

The production and export of proteins starts immediately after invasion of the parasite into the RBC and peaks in ring stage parasites (Marti *et al*., 2004). These modifications culminate in the presentation of PfEMP1 on the surface of the erythrocyte at the transition of the parasite to the trophozoite stage and the ensuing sequestration is the reason why only ring stage parasites can be detected in the blood stream of infected patients.

1.3.1.1 Maurer's clefts

A key host cell modification belonging to the exomembrane system are the Maurer's clefts (MCs). These parasite-induced vesicular structures are believed to be involved in protein trafficking (Wickham *et al*., 2001; Mundwiler-Pachlatko & Beck, 2013; Spielmann & Gilberger, 2015). Many of the exported proteins that are trafficked from the parasite into the host cell localise permanently or transiently to Maurer's clefts. Host cell surface proteins such as PfEMP1 are detected at the Maurer's clefts and knock outs of some resident Maurer's clefts proteins prevent PfEMP1 from reaching the host cell surface (Cooke *et al*., 2006; Maier *et al*., 2008; Spycher *et al*., 2008). Despite the absence of PfEMP1, virulence factor trafficking also seems to depend on similar proteins and similar structures in rodent malaria parasites (deNiz *et al*., 2016). It is to date unclear how Maurer's clefts are formed but it has been proposed that they are formed out of the

PVM (Aikawa *et al*., 1975; Atkinson *et al*., 1987; Kara *et al*., 1988). Maurer's clefts are generated rapidly after invasion. Once generated, no new clefts are formed (Grüring *et al*., 2011). Maurer's clefts can be attached to one another or to the PVM but are most frequently attached to the RBC membrane (Hanssen *et al*., 2008; Wickert *et al*., 2003; Wickert *et al*., 2004). This interaction occurs via structures termed 'tethers' (Pachlatko *et al*., 2010) or by remodelled host cell derived actin cables (Cyrklaff *et al*., 2012). Individual Maurer's clefts appear as discrete stained puncta in Giemsa stained thin blood smears (Atkinson & Aikawa, 1990; Henrich *et al*., 2009).

1.3.1.2 Other host cell modifications

Another structure described in infected RBCs that may be involved in protein export are the so called 'J-dots'. These mobile structures are distinguishable from Maurer's cleft by their differing constituents (Külzer *et al*., 2010; Külzer *et al*., 2012). They may be involved in the trafficking of exported proteins, including PfEMP1, within the host cell cytoplasm for instance to the Maurer's clefts (Kulzer *et al*., 2010; Petersen *et al*., 2016).
A further host cell modification described in *P. falciparum* infected RBCs is the tubovesicular network (TVN). It consists of an interconnected network of tubovesicular membranes that is continuous with the PVM from where it reaches into the cytoplasm of the host cell. This network was proposed to be involved in gaining access to and then taking up nutrients (Lauer *et al*., 1997; Tamez *et al*., 2008). The mechanism of nutrient uptake via the TVN remains unclear to date.

1.3.2 The mechanism of protein export

The establishment of host cell modifications and the required protein trafficking system within the host cell is realised through the export of parasite proteins into the host cell. Proteins that are exported from the parasite into the RBC or onto the RBC surface have to cross the parasite plasma membrane (PPM), the parasitophorous vacuole (PV), the parasitophorous vacuole membrane (PVM) or have to reach the RBC membrane (Figure 1.4; Marti & Spielmann, 2013).

Exported proteins harbour at least one hydrophobic region (Sargeant *et al*., 2006; Heiber *et al*., 2013) that permits the entry into the secretory pathway (Lingelbach &

Przyborski, 2006; Chang *et al*., 2008). Soluble exported proteins harbour a hydrophobic signal peptide (SP), which is processed in the ER and thereby prepares the protein for further export (Boddey *et al*., 2010; Russo *et al*., 2010). Transmembrane (TM) proteins show hydrophobic transmembrane domains for entry in the secretory pathway (Templeton, 2009). All proteins to be exported are transported by the ER and the Golgi apparatus via the vesicular transport pathway to the PPM (Figure 1.4; Lingelbach & Przyborski, 2006). Transmembrane proteins are inserted integral into the vesicular membrane upon ER entry and as a result are anchored in the lipid layers of membranes. Within these vesicles soluble proteins are transported as cargo. After transport to the periphery of the parasite, the vesicles fuse with the PPM and soluble proteins are directly released into the PV, whereas the membrane embedded proteins are extracted out of the PPM into the PV via a so far unknown translocon in the PPM (Figure 1.4; Grüring *et al*., 2012; Marti & Spielmann, 2013). This is achieved by unfolding of the transmembrane proteins, what leads to detachment and release of the protein into the PV (Grüring *et al*., 2012; Mesen-Ramirez *et al*., 2016). In some cases transmembrane proteins can also be translocated across or post translationally re-inserted into membranes (Deponte *et al*., 2012). In the PV the proteins are transported via PTEX (see section 1.3.3) – a translocon, which is situated at the inner face of the PVM – to the host cell cytosol (Figure 1.4; de Koning-Ward *et al*., 2009; Beck *et al*., 2014; Elsworth *et al*., 2014; Mesen-Ramirez *et al*., 2016). The exact mechanism is still unknown.

Exported TM proteins are found in different vesicular structures in the host cell or on the host cell surface, embedded in the RBC plasma membrane (Cooke *et al*., 2004; Spielmann & Gilberger, 2010; Deponte *et al*., 2012; Prajapati *et al*., 2014). Soluble exported proteins are found in the host cell cytosol or become attached to structures in the host cell via protein-protein interactions (Blobel *et al*., 1975 I & II; Deponte *et al*., 2012).

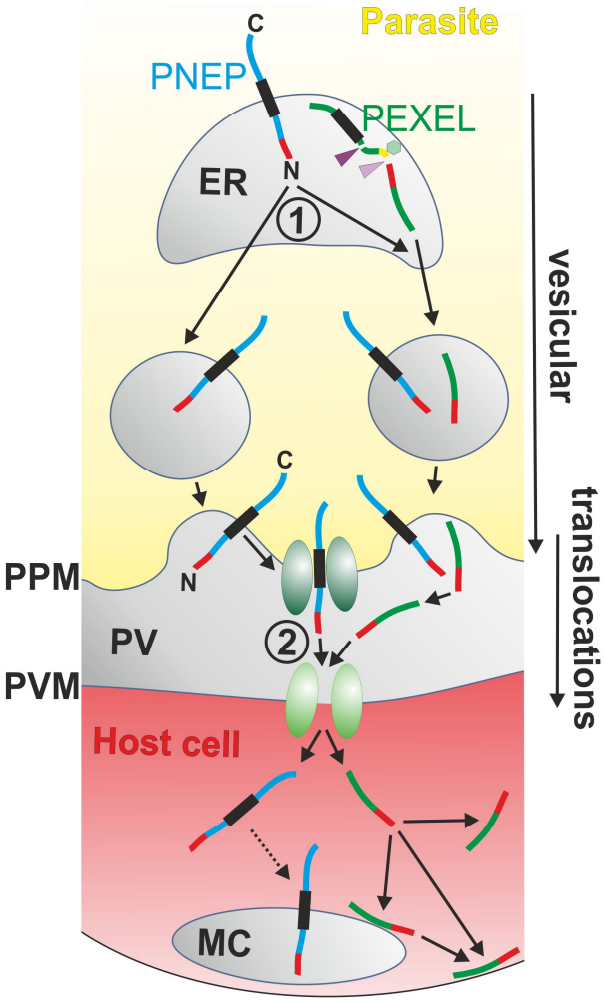


Figure 1.4 – Trafficking pathway of exported proteins from the parasite to the host cell (cytoplasm). Exported proteins (PEXEL proteins and PNEPs, see section 1.3.4) are embedded in vesicles or inserted into the membrane of vesicles. After fusion with the PPM or transport via a first translocon (dark green ellipses) the proteins are translocated via PTEX (light green ellipses; see section 1.3.3) into the host cell cytosol, where they occur as soluble or transmembrane proteins. ER: endoplasmic reticulum; PPM: parasite plasma membrane; PV: parasitophorous vacuole; PVM: parasitophorous vacuole membrane; MC: Maurer's clefts; black bars: hydrophobic regions. 1. & 2. are possible intersections within the export pathway of PEXEL proteins and PNEPs. Grüring *et al.*, 2012.

1.3.3 The translocon PTEX

To reach the host cell, all exported proteins have to pass a translocon at the PVM. A translocon is a protein complex that mediates the translocation of proteins across membranes (Johnson & van Waes, 1999). In *Plasmodium* parasites, the *Plasmodium* translocon of exported proteins (PTEX) is putatively responsible for the export of proteins from the parasite across the PVM into the host cell (Figure 1.5; de Koning-Ward *et al.*, 2009). The PTEX complex is released very early in the cycle into the PV by the merozoite apical organelles and is then associated with the inner face of the PVM. The complex consists of five proteins –heat shock protein 101 (HSP101), exported protein 2

(EXP2), thioredoxin 2 (Trx2), PTEX88 and PTEX150. HSP101 is a member of the AAA+ ATPase domain HSP100/ClpA/B chaperone family. As an ATPase it may be the energy source for the translocation process. The putative ring-shaped hexamer is a chaperone and therefore probably binds the protein and unfolds it (de Koning-Ward *et al*., 2009). Unfolding of the protein is important for the export process (Gehde *et al*., 2009; Grüring *et al*., 2012; Heiber *et al*., 2013). In agreement with the supposed crucial role of HSP101 for translocation, conditional inactivation showed that this molecule is essential for protein export (Beck at al., 2014; Elsworth *et al*., 2014). The unfolded protein is then passed to EXP2, which is a membrane-associated protein and the putative pore forming unit of the translocon within the PVM (Johnson *et al*., 1994; de Koning-Ward *et al*., 2009; Gold *et al*., 2015; Mesen-Ramirez *et al*., 2016). This complex member is believed to act as an oligomer of at least eight subunits through which the protein is believed to be passed to reach the host cell. Substrates arrested during translocation were found to be in a complex with EXP2, indicating that this protein indeed plays a role in the translocation process (Mesen-Ramirez *et al*., 2016) but details on its role in PTEX are still elusive. Trx2, possibly involved in protein oxidation and reduction, is probably located in close proximity to PTEX150, PTEX88 and HSP101 and might pass the folded protein to HSP101 (de Koning-Ward *et al*., 2009; Goldberg & Cowman, 2010). Trx2 and PTEX88 are considered to be accessory factors (de Koning-Ward *et al*., 2009; Matz *et al*., 2013; Matz *et al*., 2015). PTEX150 is essential for export (Elsworth *et al*., 2014) but its function within the complex is to date unknown (de Koning-Ward *et al*., 2009).

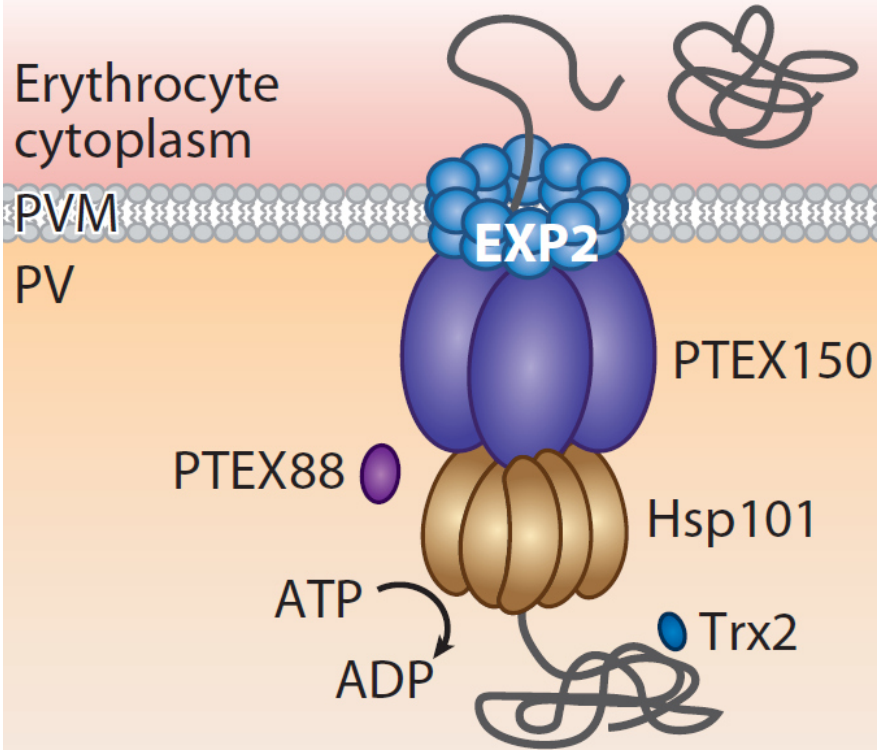


Figure 1.5 – A schematic of PTEX. Supposed arrangement of PTEX components. The protein is passed from the PV via the PTEX and its components Trx2 (blue), PTEX88 (purple), PTEX150 (violet) and the chaperone HSP101 (brown) to the putative pore-forming unit EXP2, which is located within the PVM, to reach the host cell. Modified from Boddey & Cowman, 2013.

1.3.4 Types of exported proteins – PEXEL proteins and PNEPs

Exported proteins can be divided into two groups (Figure 1.6A). The first group shows a five residue export motif, which is termed the *Plasmodium* export element (PEXEL) or vacuolar transport signal (VTS) (Marti *et al.*, 2004; Hiller *et al.*, 2004). The PEXEL motif has the consensus RxLxE/Q/D (Boddey *et al.*, 2013) and mediates the export of soluble and transmembrane proteins into the host cell (Spillman *et al.*, 2015). The arginine (R) and the leucine (L) of the motif are conserved. The fifth PEXEL position is also conserved but more diverse and most prominently holds a glutamic acid (E), an aspartic acid (D) or a glutamine (Q) (Marti *et al.*, 2004; Hiller *et al.*, 2004; Przyborski *et al.*, 2005; Sargeant *et al.*, 2006; van Ooij *et al.*, 2008; Boddey *et al.*, 2009; Boddey *et al.*, 2013; Schulze *et al.*, 2015), although no systematic testing of these residues has been carried out. An often N-terminally recessed signal peptide is located ~20 amino acids upstream of the PEXEL motif (Deponte *et al.*, 2012).

In the ER, the PEXEL motif is cleaved by the aspartic protease plasmepsin V after the leucine, between amino acid position 3 and 4 of the motif. This cleavage is thought to be essential for protein export (Chang *et al*., 2008; Boddey *et al*., 2009; Boddey *et al*., 2010; Russo *et al*., 2010). PEXEL position 1 and 3 are decisive for plasmepsin V cleavage and PEXEL position 5 – which is the second position after plasmepsin V cleavage (P2) – is important for the export of the protein (Boddey *et al*., 2009). Thus the processing uncovers the export sequence for the transport to the erythrocyte (Boddey *et al*., 2009; Grüring *et al*., 2012; Tarr *et al*., 2013). The new mature N-terminus usually starts with a non-charged amino acid (PEXEL position 4), which is followed, according to the PEXEL consensus, by a negatively charged glutamic (E) or aspartic acid (D) or by glutamine (Q) which is polar (Figure 1.6A). This mature protein is then exported. The location of the motif is conserved in PEXEL proteins (Sargeant *et al*., 2006) to guarantee the co-translational processing by plasmepsin V (Boddey *et al*., Feb 2016). The amino acid sequence downstream of the PEXEL motif can hold additional export promoting information (see section 1.3.5). Hence both regions – the beginning of the mature N-terminus (xE/Q/D, henceforth referred to as the 1° export signal) and the downstream sequence (henceforth referred to as the 2° export signal) – can influence protein export (Figure 1.6; Grüring *et al*., 2012; Tarr *et al*., 2013). However, in contrast to the 1° signal, the 2° signal is to date almost entirely undefined in position and sequence.

The second group of exported proteins does not contain a PEXEL motif and is therefore called PEXEL negative exported proteins (PNEPs; Figure 1.6A; Spielmann & Gilberger, 2010). The so far known PNEPs are diverse in structure (Blisnick *et al*., 2000; Spycher *et al*., 2003; Hawthorne *et al*. 2004; Spielmann *et al*., 2006; Pachlatko *et al*., 2010; Külzer *et al*., 2012; Heiber *et al*., 2013) and up to now no consensus motif promoting export was found. Several PNEPs do not contain a signal peptide but an internal hydrophobic region that serves as a transmembrane domain (TM) (Spielmann & Gilberger, 2010). Others harbour more than one transmembrane domain or a classical N-terminal signal peptide (SP) with or without a transmembrane domain further downstream (Külzer *et al*., 2012; Heiber *et al*., 2013). PNEPs use at least in part the same export pathway as PEXEL proteins (Figure 1.4; Grüring *et al*., 2012; Beck *et al*., 2014; Elsworth *et al*., 2014). For instance PTEX is required for the export of both types of proteins to the host cell (Beck *et al*., 2014; Elsworth *et al*., 2014). Most PNEPs known so far localise to Maurer's clefts or in the case of Maurer's clefts associated histidine rich protein 2 (MAHRP2) to the

Maurer's clefts associated tethers (Pachlatko *et al.*, 2010). Some PNEPs that are soluble in the host cell have also been described (Külzer *et al.*, 2012; Heiber *et al.*, 2013).

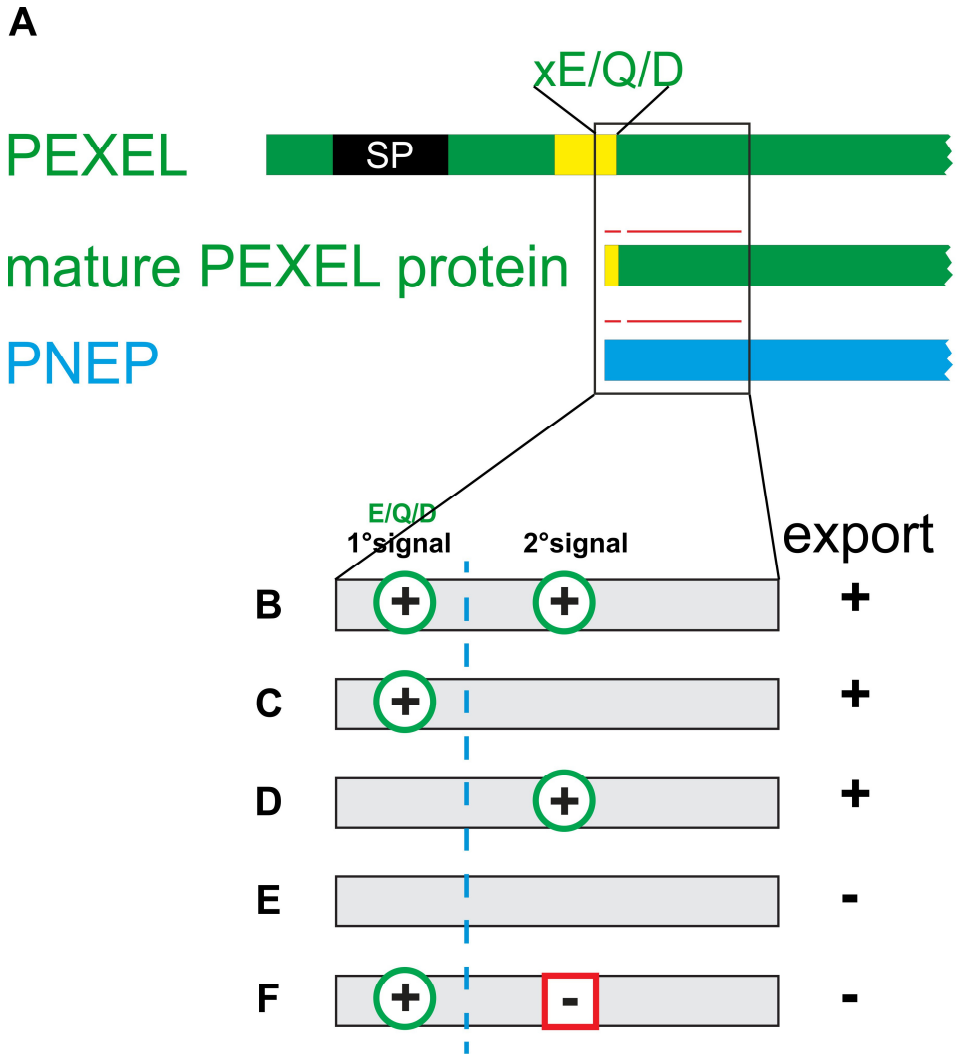


Figure 1.6 – A schematic of PEXEL proteins and PNEPs. (A) PEXEL proteins are processed within the PEXEL, which leaves the xE/Q/D as the new N-terminus (mature PEXEL protein). PNEPs do not exhibit such a motif but harbour an N-terminal region required for export that is exchangeable with the mature PEXEL N-terminus. This region might therefore represent a common core export domain (box). **(B-F)**: Based on experimental data, this core export domain can harbour different trafficking signals: 1° and 2° positive signals (green circled plus) or negative signals (red boxed minus sign). Grey bars represent neutral sequences. Positive signals alone lead to the export of the protein **(B, C, D)**, blocking sequences prevent export **(F)**. Neutral N-termini do not mediate export **(E)**. Combined from Grüring *et al.*, 2012, and Marti & Spielmann, 2013.

The first 20 amino acids of PNEPs lacking a classical N-terminal signal peptide are sufficient to promote export of a reporter carrying also a PNEP TM (Grüring *et al.*, 2012; Heiber *et al.*, 2013). Up to now, around two dozen PNEPS are known in *P. falciparum*. Members of this group are for example SBP1, MAHRP1 and MAHRP2 as well as REX1 and REX2 (Blisnick *et al.*, 2000; Spycher *et al.*, 2003; Hawthorne *et al.*, 2004; Spielmann *et al.*, 2006; Pachlatko *et al.*, 2010). However, it is likely that there are many more

PNEPs, as a consensus sequence to identify them is so far lacking (Heiber *et al*., 2013). For other *Plasmodium* spp. even more PNEPs may exist, as these species seem to contain a lower number of PEXEL proteins (Sargeant *et al*., 2006).

The major *P. falciparum* virulence factor PfEMP1 (see section 1.3.1), a large multi-variant molecule, is encoded by a family of approximately 60 different *var* genes present in the parasite genome (Flick & Chen, 2004). These genes are expressed in a mutually exclusive manner, ensuring that only one variant is expressed at a given time. Switching of the variant is the reason for antigenic variation of the iRBC (Flick & Chen, 2004). PfEMP1 is a special case in terms of the classification into PEXEL proteins and PNEPs. It was originally categorised as a protein containing a PEXEL-like motif (Marti *et al*., 2004; Sargeant *et al*., 2006) or a VTS (Hiller *et al*., 2004) with the sequence KxLxD. However, this PEXEL-like sequence is not cleaved by plasmepsin V (Boddey *et al*., 2013) and as a consequence, PfEMP1 was later classified as a PNEP (Boddey *et al*., 2013). Using the example of PfEMP1, the importance of the discovery of a possible common core export domain – shared by PEXEL proteins and PNEPS – becomes obvious.

1.3.5 Positive and negative signals

Both groups of exported proteins – PEXEL proteins and PNEPs – hold the export information in their N-termini. In the case of PEXEL proteins, the mature N-terminus, after most of the PEXEL was removed, is essential for export. This mature N-terminus is sufficient to drive export of a reporter containing a PNEP TM. This N-terminal export region is exchangeable between PEXEL proteins and PNEPs (Grüring *et al*., 2012), indicating that processed PEXEL proteins and PNEP N-termini are comparable and that both groups share a common core export domain (Figure 1.6; Marti & Spielmann, 2013; Spielmann & Gilberger, 2015). Based on previous results, this core domain is likely to be located within the first 20 amino acids of the N-terminus (Figure 1.6A, box; Haase *et al*., 2009; Grüring *et al*., 2012; Heiber *et al*., 2013). However, the exact sequence determinants or the actual core domain are so far unknown, although some parameters were already tested: The remaining part of the PEXEL motif (1° positive signal) after plasmepsin V cleavage, an E, Q or D in position 2, is essential for export (Figure 1.6B, C; Boddey *et al*., 2009). However, mutation of this residue resulted in protein export of all of the so far investigated proteins when only the mature N-termini were examined

(Grüring *et al*., 2012; Tarr *et al*., 2013). That indicates that possible export signals further downstream (2° positive signal) can substitute for the primary positive signal (Figure 1.6D). These 2° positive signals probably also consist of negatively charged amino acids (Haase *et al*., 2009; Saridaki *et al*., 2009; Bhattacharjee *et al*., 2009) and are therefore maybe comparable to the primary positive signal. This implicates a certain resemblance of PNEP and mature PEXEL N-termini (Spielmann & Gilberger, 2010). It was further supposed that an overall negative net charge in PNEP N-termini is important for protein export (Spycher *et al*., 2006; Saridaki *et al*., 2009). However, scrambling of PNEP N-termini resulted in reduction or even abolishment of protein export, although the overall net charge of the N-terminus did not change (Grüring *et al*., 2012). Hence the net charge of the N-terminus seems not to be decisive for protein export.

Negative signals are those which block the export of a protein even though there is a classical primary positive signal located in the N-terminus – for instance, there is a perfect PEXEL left over such as E, Q or D at position 2 of the N-terminus but the protein is not exported (Figure 1.6F). Two such N-termini were discovered in proteins of the PV (SERA7) and the PVM (ETRAMP5; Grüring *et al*., 2012), indicating that the N-termini of these proteins evolved in such way that additional sequences prevent their export to maintain their correct localisation. N-termini can also be neutral, thus harbouring no export or blocking region. Proteins exhibiting such an N-terminus are also not exported (Figure 1.6E).

Taken together these data indicate that the presence or absence of positive and negative signals within the N-terminus decides about the export profile of the protein (Marti & Spielmann, 2013) but no definitive and systematic data on this topic exist.

1.4 Aim of this study

Plasmodium falciparum parasites remodel their host red blood cell by exporting a large number of proteins. This is essential for its survival and virulence. Different groups of exported proteins exist but they may share a common export domain and follow a partially shared export pathway. However, it is exactly this shared export domain that is only poorly characterised. This thesis has the aim to define the sequence requirements of this shared export domain. This will provide crucial information for the prediction of PNEPs to obtain the complete *Plasmodium* exportome and will form the basis to understand the common nature of the export steps shared between PNEPs and PEXEL proteins.

2 Material and methods

2.1 Material

2.1.1 Technical devices

Device	Specifications	Brand / Distributor
Agarose gel chamber	Sub Cell GT basic	Bio-Rad, München
Analytical Balance	870	Kern
Blot device Gel holder cassettes Foam pads Electrode assembly Cooling unit	Mini Protean Tetra Cell System	Bio-Rad, München
Centrifuge	Megafuge 1.0R J2- HS Ultracentifuge Rotor JA-12 Avanti J-26S XP Rotor JA-14	Heraues, Hannover Beckman Coulter, Krefeld Beckman Coulter, Krefeld
Table centrifuge	Eppendorf 5415 D	Eppendorf, Hamburg
Casting gel stuff Casting stand Casting plates Casting frames 12 –wells combs	Mini Protean	Bio-Rad, München
Developer	Curix 60	AGFA-Gevaert, Mortsel/Belgium
Developer cassette	Cronex Quanta III	Dupont, Neu Isenburg
Electrophoresis chamber	Mini Protean 67s	Bio-Rad, München
Electroporator	Gene Pulser X- Cell	Bio-Rad, München
Electroporator	Nucleofector II AAD-1001N	Amaxa Biosystems, Germany
Ice machine	EF 156 easy fit	Scotsmann, Vernon Hills/USA
Bacterial incubator	Thermo function line	Heraeus, Hannover
P. falciparum cell culture incubator	Heratherm IGS400	Thermo Scientific, Langenselbold
Shaking incubator	Max Q4000	Barnstead, Iowa/ USA
Light Microscope	Axio Lab A1	Zeiss, Jena
Fluorescence Microscope	Axioscope 1	Zeiss, Jena

Device	Specifications	Brand / Distributor
Microscope digital camera	Orca C4742-95	Hamamatsu Phototonics K.K., Japan
Microwave	Micro 750W	Whirlpool, China
Laboratory scale	Atilon	Acculab Sartorius, Göttingen
PCR Mastercycler	epgradient	Eppendorf, Hamburg
Photometer	BioPhotometer plus	Eppendorf, Hamburg
pH-meter	SevenEasy	Mettler-Toledo, Gießen
Pipettes	1-10/200/1000 µl	Gilson, Middleton, USA
Pipettor	Pipetboy acu	IBS,
Power supply	EV31 Power Source 300 V	Consort, Belgium VWR, Taiwan
Roller mixer	STR6	Stuart
Sterile laminar flow bench	Steril Gard III Advance	Baker, Stanford USA
Thermoblock	Thermomixer compact	Eppendorf, Hamburg
Ultrapure water purification system	Milli Q	Millipore
UV transiluminator	PHEROlum289	Biotec Fischer,Reiskirchen
Vacuum pump	BVC Control	Vacuubrand, Deutschland
Vortexer	Genie 2	Scientific Industries, USA
Waterbath	1083	GFL, Burgwedel

2.1.2 Chemicals

Reagent	Brand / Distributor
Acetic acid	Roth, Karlsruhe
Acrylamide/Bisacrylamide solution (40%)	Roth, Karlsruhe
Agar LB (Lennox)	Roth, Karlsruhe
Agarose	Invitrogen, USA
AlbumaxII	Gibco, Life Technologies, USA
Albumin bovine Fraction V (BSA)	Biomol, Hamburg
Ammonium persulfate (APS)	Applichem, Darmstadt
Ampicillin	Roche, Mannheim
Bacto™ yeast extract	Roth, Karlsruhe
Bacto™ Peptone	Roth, Karlsruhe
Blasticidin S	Invitrogen, USA
Bromophenol blue	Roth, Karlsruhe

Reagent	Brand / Distributor
Coomassie Brilliant Blue G-250	Merck, Darmstadt
Calciumchlorid ($CaCl_2$)	Sigma, Steinheim
Desoxynucleotides (dNTPs)	Thermo Scientific, Lithuania
Developer solution G150 (Western blot)	Agfa, Leverkusen
4',6-diamidino-2-phenylindole (DAPI)	Roche, Mannheim
Dimethyl sulfoxide (DMSO)	Sigma, USA
Dipotassium phosphate	Merck, Darmstadt
Disodium phosphate	Roth, Karlsruhe
Dithiobis(succinimidylpropionate) (DSP)	Thermo Scientific, USA
1,4,-dithiothreitol (DTT)	Biomol, Hamburg
Dulbecco's Phosphate Buffered Saline (DPBS)	PAN, Biotech, Aidenbach
Ethanol	Roth, Karlsruhe
Ethidium bromide (EtBr)	Sigma, Steinheim
Ethylenediaminetetraacetic acid (EDTA)	Biomol, Hamburg
Ethylene glycol tetraacetic acid (EGTA)	Merck, Darmstadt
Fixation solution G334 (Western blot)	Agfa, Leverkusen
Gentamycin	Ratiopharm, Ulm
Giemsa's azure, eosin, methylene blue solution	Merck, Darmstadt
D-Glucose	Merck, Darmstadt
Glycerol	Merck, Darmstadt
Glycine	Biomol, Hamburg
4-(2-Hydoxyethyl)-1-piperazineethanesulfonicacid (HEPES)	Roche, Mannheim
Hydrochloric acid (HCl)	Merck, Darmstadt
Hypoxanthin	Sigma, Steinheim
Isopropanol	Roth, Karlsruhe
Isopropyl-β-D-thiogalactopyranosid (IPTG)	Roth, Karlsruhe
Lysozyme	Fluka Analytical
Magnesium chloride ($MgCl_2$)	Merck, Darmstadt
Manganese(II) chloride ($MnCl_2$)	Merck, Darmstadt
β-Mercaptoethanol	Merck, Darmstadt
Methanol	Roth, Karlsruhe
3-(N-morpholino)propansulfonic acid (MOPS)	Sigma, Steinheim
Milk powder	Roth, Karlsruhe
Percoll	GE Healthcare, Sweden
Phenylmethylsulfonylfluorid (PMSF)	Sigma, Steinheim
Potassium chloride	Merck, Darmstadt

Reagent	Brand / Distributor
Potassium dihydrogen phosphate	Merck, Darmstadt
Protease inhibitor cocktail ("Complete Mini")	Roche, Mannheim
Rubidium chloride	Sigma, Steinheim
RPMI (Roswell Park Memorial Institute)-Medium	Applichem, Darmstadt
Saponin	Sigma, Steinheim
Sodium acetate	Merck, Darmstadt
Sodium chloride	Gerbu, Gaiberg
Sodium bicarbonate	Sigma, Steinheim
Sodium dodecyl sulfate (SDS)	Applichem, Darmstadt
Sodium dihydrogen phosphate	Roth, Karlsruhe
Sodium hydroxide	Merck, Darmstadt
Sorbitol	Sigma, Steinheim
Tetanolysin	Sigma, Steinheim
N, N, N, N-Tetramethylethylenediamin (TEMED)	Merck, Darmstadt
Tris base	Roth, Karlsruhe
Tris-EDTA (TE)	Invitrogen, Karlsruhe
Trichloroacetic acid	Roth, Karlsruhe
Triton X-100	Biomol, Hamburg
Water for molecular biology (Ampuwa)	Fresenius Kabi, Bad Homburg
WR99210 (WR)	Jacobus Pharmaceuticals, Washington (USA)
Yeast extract	Becton Dickinson, Heidelberg

2.1.3 Lab ware and disposables

Labware and disposables	Specifications	Manufacturer
Chromatography paper		Whatman
Conical falcon tubes	15ml, 50ml	Sarstedt, Nümbrecht
Cryotubes	1.6 ml	Sarstedt, Nümbrecht
Culture bottles	50 ml	Sarstedt, Nümbrecht
Disposable pipette tips	1-10/ 20-200/ 100-1000 µl	Sarstedt, Nümbrecht
Filter, round	150mm	Macherey-Nagel, Düren
Filter tips	1-10/ 20-200/ 100-1000 µl	Sarstedt, Nümbrecht
Glass cover slips	24 x 65 mm Thickness 0.13-0.16	R. Langenbrinck, Emmendingen
Glass slides		Engelbrecht, Edermünde

Labware and disposables	Specifications	Manufacturer
Gloves, latex		Kimtech Science EcoShield™
Gloves, purple nitrile		Kimtech Science
IFA glass slides	10 wells ER-208B-CE24 6.7 mm	Thermo Scientific, USA
Leukosilk tape		BSN medical GmbH
Medical X-Ray screen film blue sensitive	CEA RP NEW	AGFA Health Care NV, Mortsel, Belgium
Nitrocellulose blotting Membrane Protran	Amersham 0.45 µm	GE Healthcare, Deutschland
One way cannula		Braun, Melsungen
One way injection		Braun, Melsungen
Parafilm		Bemis, USA
Pasteur pipette		Brand, Wertheim
PCR Reaction tubes	Multiply-µStrip Pro 8-Strip	Sarstedt, Nümbrecht
Petri dishes	5 ml/ 10 ml 15 x 60 and 14 x 90 mm	Sarstedt, Nümbrecht
Plastic pipettes	5/ 10/ 25 ml	Sarstedt, Nümbrecht
Reaction tubes	0,2ml / 0,5ml / 1.5 ml / 2 ml	Sarstedt, Nümbrecht/ Eppendorf Hamburg
Scalpel		Braun, Tuttlingen
Sterile filter	0.22 µm	Sarstedt, Nümbrecht
Transfection cuvettes	0.2 cm	Bio-Rad, München

2.1.4 Kits

Kit	Manufacturer
NucleoSpin. Plasmid	Macherey-Nagel, Düren
NucleoSpin. Extract II	Macherey-Nagel, Düren
QIAamp DNA Mini Kit	Qiagen, Hilden
QIAGEN Plasmid Midi Kit	Qiagen, Hilden
Western Blot ECL-SuperSignal West Pico	Thermo Scientific, Schwerte
Western Blot ECL-Clarity Detection Kit	Bio-Rad, USA

2.1.5 DNA- and protein ladders

DNA- and protein ladder	Manufacturer
GeneRuler™1000 bp ladder	Thermo Scientific, Schwerte
PageRuler™ prestained protein ladder	Thermo Scientific, Schwerte
PageRuler™ unstained protein ladder	Thermo Scientific, Schwerte

2.1.6 Media, buffers and solutions

2.1.6.1 Media, buffers and solutions for microbiological culture

10x Luria-Bertani (LB) medium stock solution	10% NaCl 5% peptone 10% yeast extract in dH_2O, autoclaved
LB medium working solution	1% (w/v) NaCl 0.5% (w/v) peptone 1%(w/v) yeast extract in dH_2O in dH_2O, autoclaved
LB Agar plate solution	1.5% agar-agar 1x LB medium
Ampicillin stock solution	100mg/ml ampicillin in 70% ethanol
Glycerol freezing solution	50% (v/v) glycerol in 1x LB medium

2.1.6.2 Buffers for competent *E. coli* cells

TFBI buffer	30mM acetic acid 50nM $MnCl_2$ 100mM RbCl 10mM $CaCl_2$ 15% (v/v) glycerol pH 5.8 (with 0.2N Acetic acid) ad 500ml dH_20
TFBII buffer	10mM MOPS 75mM $CaCl_2$

	10mM RbCl
	15% (v/v) glycerol
	pH 7.0 (with NaOH)
	ad 500ml dH_20

2.1.6.3 Buffers and solutions for molecular biology analyses

2.1.6.3.1 Plasmid DNA preparation using STET buffer

STET buffer	80g/l saccharose
	18,612g/l EDTA , pH 8,0
	5g/l Triton X-100
	1,211g/l Tris-Base, pH 8,0
	in dH_2O

2.1.6.3.2 DNA precipitation

Sodium acetate	3M, pH 5.2
Ethanol	100%
Tris-EDTA (TE) buffer	10mM Tris-HCl, pH 8.0
	1mM EDTA

2.1.6.3.3 DNA electrophoresis

50x TAE-Buffer	2M Tris base
	1M pure acetic acid
	50mM EDTA
	pH 8.5
	in dH_2O
6x Loading buffer	40% glycerol (v/v)
	2.5% (w/v) xylene cyanol
	2.5% (w/v) bromophenol blue
	in dH_2O

2.1.6.4 Media and solutions for parasite culture and cell biology experiments

2.1.6.4.1 *P. falciparum in vitro* culture

RPMI complete medium	1,587% (w/v) RMPI 1640 12mM $NaHCO_3$ 6mM D-Glucose 0.5% (v/v) Albumax II 0.2mM hypoxanthine 0.4mM gentamycin pH 7.2 in dH_2O sterile filtered
10% Giemsa solution	10ml Giemsa's azure, eosin, methylene blue solution 90ml dH_2O
Synchronization solution	5% (w/v) D-sorbitol in dH_2O sterile filtered
Transfection buffer (Cytomix)	120mM KCl 150µM $CaCl_2$ 2mM EGTA 5mM $MgCl_2$ 10mM K_2HPO_4 / KH_2PO_4 25mM HEPES pH 7.6 in dH_2O sterile filtered
Amaxa transfection buffer	90mM Na_2HPO_4 5mM KCl

	0.15mM $CaCl_2$ 50mM HEPES pH 7.3 in dH_2O sterile filtered
Malaria freezing solution (MFS)	4.2% D-sorbitol 0.9% NaCl 28% glycerol in dH_2O sterile filtered
Malaria thawing solution (MTS)	3.5% NaCl in dH_2O sterile filtered
WR99210 stock solution	20mM WR99210 in 1ml DMSO sterile filtered
WR99210 working solution	1:1000 dilution of stock solution in RPMI complete medium
Blasticidin S (BSD) working solution	5mg/ml BSD in RPMI complete medium sterile filtered
G418 working solution	50mg/mL in RPMI complete medium sterile filtered
Human red blood cells	Blood bank,

	Universitätsklinikum, sterile, concentrate; bloodgroup 0+, Eppendorf (UKE), Hamburg

2.1.6.4.2 Buffers and solutions for cell biology and biochemical assays

Parasite lysis buffer	4% SDS 0.5% Triton X-100 0.5x PBS in dH_2O
Percoll stock solution	90% (v/v) Percoll 10% (v/v) 10 x PBS
80% Percoll solution	89% (v/v) Percoll stock solution 11% (v/v) RPMI compl. medium 4% (w/v) sorbitol sterile filtered
60% Percoll solution	67% (v/v) Percoll stock solution 33% (v/v) RPMI compl. medium 4% (w/v) sorbitol
40% Percoll solution	44% (v/v) Percoll stock solution 56% (v/v) RPMI compl. medium 4% (w/v) sorbitol
Saponin solution (for selective membrane permeabilisation)	0.03% (w/v) saponin 150mM NaCl 1mM PMSF 2x Protease inhibitor cocktail in dH_2O

Diluting buffer	10mM Tris-HCl, pH 7.5 150mM NaCl 1mM PMSF 2x Protease inhibitor cocktail in dH_2O
DSP (Stock solution)	20mM in DMSO
Quenching buffer	25mM Tris-HCl in 1x PBS

2.1.6.5 Buffers and solutions for protein analyses

2.1.6.5.1 SDS-Page and western blot

10x Running buffer	250mM Tris base 1.92M Glycine 1% (w/v) SDS in dH_2O
Ammonium persulfate (APS)	10% (w/v) in dH_2O
Separating gel buffer	1.5M Tris-HCl, pH 8.8 in dH_2O
Stacking gel buffer	1M Tris-HCl, pH 6.8 in dH_2O
Stacking gel (for two gels, 5%)	0.75ml stacking gel buffer 4.35ml dH_2O 750µl acryl amide (40%) 60µl SDS (10%) 60µl APS (10%) 6µl TEMED

Separating gel (for two gels, 12%)	2.5ml running gel buffer 4.2ml dH_2O 3ml acryl amide (40%) 100µl SDS (10%) 100µl APS (10%) 4µl TEMED
6 x SDS sample buffer	375mM Tris-HCl, pH 6.8 12% (w/v) SDS 60% (v/v) glycerol 0.6M DTT 0.06% (w/v) bromophenol blue in dH_2O
10 x Western transfer buffer	250mM Tris base 1.92M glycerol 0.1% (w/v) SDS in dH_2O
1 x Western transfer buffer	10% 10x Western transfer buffer 20% Methanol in dH_2O
Blocking solution	5% (w/v) milk powder in 1x PBS
Washing buffer	1x PBS

2.1.7 Bacterial and *Plasmodium* strains

E. coli strain XL-10 Gold	Tet^r Δ(*mcrA*)183 Δ(*mcrCB-hsdSMR-mrr*)173 *endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte* [F' *proAB* $lacI^q$ZΔ*M15* Tn*10* (Tet^r) Amy Cam^r]

P. falciparum strain 3D7	NF54 clone, patient isolate (from near Schipol airport, Amsterdam, the Netherlands), likely origin: West Africa

2.1.8 Enzymes

2.1.8.1 Polymerases

FirePol DNA Polymerase [5 U/µl]; Solis Biodyne, Taipei, Taiwan

Phusion High-Fidelity DNA Polymerase [2 U/µl]; NEB, Ipswich, USA

2.1.8.2 Restriction enzymes

Restriction enzyme	Restriction site	Brand/Manufacturer
AvrII	C^CTAGG	NEB, Ipswich / USA
KpnI	GGATC^C	NEB, Ipswich / USA
XhoI	C^TCGAG	NEB, Ipswich / USA
DpnI	GA^TC	NEB, Ipswich / USA

2.1.8.3 Ligases

T4 DNA-Ligase [3 U/µl] NEB, Ipswich, USA

2.1.9 Antibodies

2.1.9.1 Primary antibodies

Antigen	Organism	Dilution		Brand/Manufacturer
		WB	IFA	
Aldolase	Rabbit	1:4000	-	AG Spielmann
GFP	Mouse	1:1000	1:500	Roche, Mannheim
GFP	Rabbit	1:2000	1:500	Thermo Scientific
Triple hemagglutinin (3x HA)	Rat	1:2000	1:500	Roche, Mannheim
myc	Rabbit	1:500	1:200	Cell Signaling, USA

2.1.9.2 Secondary antibodies

Antigen	Conjugate	Organism	Dilution	Application	Brand/Manufacturer
Mouse	HRP	Goat	1:3000	Western blot	Dianova, Hamburg
Rabbit	HRP	Donkey	1:2500	Western blot	Dianova, Hamburg
Rat	HRP	Goat	1:3000	Western blot	Dianova, Hamburg

2.1.9.3 Antibody coupled beads

Antigen	Conjugate	Organism	Application	Brand/Manufacturer
HA	Agarose	Mouse	IP	Thermo Scientific, USA
GFP	Agarose	Camel	IP	Chromotek, München
Streptavidin	Sepharose	-		IBA

2.1.10 Oligonucleotides

Name	Sequence with restriction site (written in lower-case characters)
AAAAAAAAAA_mTRAP_KpnI_fw	CAGCggtaccATGGCAGCTGCTGCCGCAGCAGCAGCTGCCGCTTCTGCATTATATGAACATATGAATAC
AAAAEAAAAA_mTRAP_KpnI_fw	CAGCggtaccATGGCAGCTGCTGCCGAAGCAGCAGCTGCCGCTTCTGCATTATATGAACATATGAATAC
AAAAAAAEAA_mTRAP_KpnI_fw	CAGCggtaccATGGCAGCTGCTGCCGCAGCAGCAGAAGCCGCTTCTGCATTATATGAACATATGAATAC
AASAASAASA_KpnI_fw	CAGCggtaccATGGCTGCATCAGCAGCATCTGCAGCTTCAGCTTCTGCATTATATGAACATATGAATAC
AEAAAAAAAA_mTRAP_KpnI_fw	CAGCggtaccATGGCAGAAGCTGCCGCAGCAGCAGCTGCCGCTTCTGCATTATATGAACATATGAATAC
AGAGAGAGAG_KpnI_fw	CAGCggtaccATGGCTGGTGCAGGAGCAGGAGCCGGTGCTGGTTCTGCATTATATGAACATATGAATAC
aEagagagag_KpnI_fw	CAGCggtaccATGGCTGAAGCAGGAGCAGGAGCCGGTGCTGGTTCTGCATTATATGAACATATGAATAC
agaEaEaEag_KpnI_fw	CAGCggtaccATGGCTGGTGCAGAAGCAGAGGCCGAAGCTGGTTCTGCATTATATGAACATATGAATAC
agagaEagag_KpnI_fw	CAGCggtaccATGGCTGGTGCAGGAGCAGAAGCCGGTGCTGGTTCTGCATTATATGAACATATGAATAC
agagagagaE_KpnI_fw	CAGCggtaccATGGCTGGTGCAGGAGCAGGAGCCGGTGCTGAATCTGCATTATATGAACATATGAATAC
aeaPPPagag_KpnI_fw	CAGCggtaccATGGCTGAAGCACCTCCTCCTGCCGGTGCTGGTTCTGCATTATATGAACATATGAATAC
agEgEgEgag_KpnI_fw	CAGCggtaccATGGCTGGTGAAGGAGAGGGAGAAGGTGCTGGTTCTGCATTATATGAACATATGAATAC
Crt fw	CCGTTAATAATAAATACACGCAGTC
Crt 131 s	CACATATATGACATAAATATTTTAAAATCG

Name	Sequence with restriction site (written in lower-case characters)
GBP130mut_E90A_KpnI_fw	CAGCggtaccATGGCTGCAGGAGCAGATACGTGTGCACGAAAAGAAAAGACTACATTAAGAAAAAGTAAGCAGTCTGCATTATATGAACATATGAATAC
GBP130mut_n4-7_KpnI_fw	CAGCggtaccATGGCTGCAGCTGGTGCAGGATGTGCACGAAAAGAAAAGACTACATTAAGAAAAAGTAAGCAGTCTGCATTATATGAACATATGAATAC
GBP130mut_n5-8_KpnI_fw	CAGCggtaccATGGCTGCAGGAGCTGGTGCAGGAGCACGAAAAGAAAAGACTACATTAAGAAAAAGTAAGCAGTCTGCATTATATGAACATATGAATAC
GBP130mut_n6-9_KpnI_fw	CAGCggtaccATGGCTGCAGGAGAAGCTGGTGCAGGACGAAAAGAAAAGACTACATTAAGAAAAAGTAAGCAGTCTGCATTATATGAACATATGAATAC
GBP130mut_n8-11_KpnI_fw	CAGCggtaccATGGCTGCAGGAGAAGATACGGCTGGTGCAGGAGAAAAGACTACATTAAGAAAAAGTAAGCAGTCTGCATTATATGAACATATGAATAC
GBP130mut_n10-13_KpnI_fw	CAGCggtaccATGGCTGCAGGAGAAGATACGTGTGCAGCTGGTGCAGGAACTACATTAAGAAAAAGTAAGCAGTCTGCATTATATGAACATATGAATAC
GBP130mut_n12-15_KpnI_fw	CAGCggtaccATGGCTGCAGGAGAAGATACGTGTGCACGAAAAGCTGGTGCAGGATTAAGAAAAAGTAAGCAGTCTGCATTATATGAACATATGAATAC
GBP130mut_n12-20_KpnI_fw	CAGCggtaccATGGCTGCAGGAGAAGATACGTGTGCACGAAAAGCTGGTGCAGGAGCAGGAGCCGGTGCTGGTTCTGCATTATATGAACATATGAATAC
GFP 85 rv	ACCTTCACCCTCTCCACTGAC
GFP 272 as	CCTTCGGGCATGGCACTC
GFP 633 fw	GCCCTTTCGAAAGATCCC
KAHRP_fw_KpnI	CAGCggtaccATGGCTCAAAAACAACATGAACATCACCATCATCACCATCATCAACATCAGCATCAACATTCTGCATTATATGAACATATGAATAC
KAHRP_Q2A_fw_KpnI	CAGCggtaccATGGCTGCAAAACAACATGAACATCACCATCATCACCATCATCAACATCAGCATCAACATTCTGCATTATATGAACATATGAATAC
KAHRP_Q4A_fw_KpnI	CAGCggtaccATGGCTCAAAAAGCTCATGAACATCACCATCATCACCATCATCAACATCAGCATCAACATTCTGCATTATATGAACATATGAATAC
KAHRP_Q2A_Q4A_fw_KpnI	CAGCggtaccATGGCTGCAAAAGCTCATGAACATCACCATCATCACCATCATCAACATCAGCATCAACATTCTGCATTATATGAACATATGAATAC
KAHRP_Q2A_Q4A_E6A_fw_KpnI	CAGCggtaccATGGCTGCAAAAGCTCATGCTCATCACCATCATCACCATCATCAACATCAGCATCAACATTCTGCATTATATGAACATATGAATAC
maEtastasta_KpnI_fw	CAGCggtaccATGGCTGAAACAGCATCTACTGCTTCAACAGCATCTGCATTATATGAACATATGAATAC
MASTAETASTA_KpnI_fw	CAGCggtaccATGGCTTCAACAGCAGAAACTGCTTCAACAGCATCTGCATTATATGAACATATGAATAC
MASTASTAETA_KpnI_fw	CAGCggtaccATGGCTTCAACAGCATCTACTGCTGAAACAGCATCTGCATTATATGAACATATGAATAC
mastastasta_KpnI_fw	CAGCggtaccATGGCTTCAACAGCATCTACTGCTTCAACAGCATCTGCATTATATGAACATATGAATAC
mTRAP_rev (AvrII)	CGCGcctaggTTCGAGTGCCCAGAATTCTTCTTC
n_A7-9_KpnI_fw	CAGCggtaccATGGCTGGTGCAGGAGCAGCAGCTGCTGCTGGTTCTGCATTATATGAACATATGAATAC

Name	Sequence with restriction site (written in lower-case characters)
n_D7-9_KpnI_fw	CAGCggtaccATGGCTGGTGCAGGAGCAGACGATGATGCTGGTTCTGCATTATATGAACATATGAATAC
n_E3_G5-7_KpnI_fw	CAGCggtaccATGGCTGAAGCAGGTGGTGGAGCCGGTGCTGGTTCTGCATTATATGAACATATGAATAC
n_E4_KpnI_fw	CAGCggtaccATGGCTGGTGAAGGAGCAGGAGCCGGTGCTGGTTCTGCATTATATGAACATATGAATAC
n_E4-6_KpnI_fw	CAGCggtaccATGGCTGGTGAAGAGGAAGGAGCCGGTGCTGGTTCTGCATTATATGAACATATGAATAC
n_E5_KpnI_fw	CAGCggtaccATGGCTGGTGCAGAAGCAGGAGCCGGTGCTGGTTCTGCATTATATGAACATATGAATAC
n_E5_flanked by As_KpnI_fw	CAGCggtaccATGGCTGCAGAAGCAGGAGCCGGTGCTGGTTCTGCATTATATGAACATATGAATAC
n_E6_KpnI_fw	CAGCggtaccATGGCTGGTGCAGGTGAAGGAGCCGGTGCTGGTTCTGCATTATATGAACATATGAATAC
n_E7-9_KpnI_fw	CAGCggtaccATGGCTGGTGCAGGAGCAGAAGAGGAAGCTGGTTCTGCATTATATGAACATATGAATAC
n_E9_KpnI_fw	CAGCggtaccATGGCTGGTGCAGGAGCAGGAGCCGAAGCTGGTTCTGCATTATATGAACATATGAATAC
n_GBP4-7_KpnI_fw	CAGCggtaccATGGCTGGTGGAGAAGATACGGCCGGTGCTGGTTCTGCATTATATGAACATATGAATAC
n_GBP4-8_KpnI_fw	CAGCggtaccATGGCTGGTGGAGAAGATACGTGTGGTGCTGGTTCTGCATTATATGAACATATGAATAC
n_GBP5-8_KpnI_fw	CAGCggtaccATGGCTGGTGCAGAAGATACGTGTGGTGCTGGTTCTGCATTATATGAACATATGAATAC
n_L3_KpnI_fw	CAGCggtaccATGGCTTTAGCAGGAGCAGGAGCCGGTGCTGGTTCTGCATTATATGAACATATGAATAC
n_M7-9_KpnI_fw	CAGCggtaccATGGCTGGTGCAGGAGCAATGATGATGGCTGGTTCTGCATTATATGAACATATGAATAC
n_N3_KpnI_fw	CAGCggtaccATGGCTAATGCAGGAGCAGGAGCCGGTGCTGGTTCTGCATTATATGAACATATGAATAC
n_P7-9_KpnI_fw	CAGCggtaccATGGCTGAAGCAGGAGCACCACCACCTGCTGGTTCTGCATTATATGAACATATGAATAC
n_P11-13_KpnI_fw	CAGCggtaccATGGCTGAAGCAGGAGCAGGAGCCGGTGCTCCACCACCTTCTGCATTATATGAACATATGAATAC
n_R3_KpnI_fw	CAGCggtaccATGGCTAGAGCAGGAGCAGGAGCCGGTGCTGGTTCTGCATTATATGAACATATGAATAC
n_S3_KpnI_fw	CAGCggtaccATGGCTTCAGCAGGAGCAGGAGCCGGTGCTGGTTCTGCATTATATGAACATATGAATAC
n_Ste3-7_KpnI_fw2	CAGCggtaccATGGCTGGAACCCAAATCCATAATGGTGCTGGTTCTGCATTATATGAACATATGAATAC
n_Ste5-10_KpnI-fw	CAGCggtaccATGGCTGGTGCAGGAATCCATAATCCACATTATTCTGCATTATATGAACATATGAATAC
n_Ste9-14_KpnI-fw	CAGCggtaccATGGCTGGTGCAGGAGCAGGAGCCGGTCATTATCATAATGATCCATCTGCATTATATGAACATATGAATAC
n_Ste13-20_KpnI-fw	CAGCggtaccATGGCTGGTGCAGGAGCAGGAGCCGGTGCTGGTGATCCAGAACTCAAAGAAATAATTTCTGCATTATATGAACATATGAATAC
n_T3_KpnI_fw	CAGCggtaccATGGCTACAGCAGGAGCAGGAGCCGGTGCTGGTTCTGCATTATATGAACATATGAATAC
n2_A2_KpnI_fw	CAGCggtaccATGGCTGCAGGAGCAGGAGCCGGTGCTGGTGCTTCTGCA

Name	Sequence with restriction site (written in lower-case characters)
	TTATATGAACATATGAATAC
n2_A2_E3_KpnI_fw	CAGCggtaccATGGCTGAAGGAGCAGGAGCCGGTGCTGGTGCTTCTGCATTATATGAACATATGAATAC
n2_GAGAGAGAGA_KpnI_fw	CAGCggtaccATGGGTGCAGGAGCAGGAGCCGGTGCTGGTGCTTCTGCATTATATGAACATATGAATAC
n2_E3_KpnI_fw	CAGCggtaccATGGGTGAAGGAGCAGGAGCCGGTGCTGGTGCTTCTGCATTATATGAACATATGAATAC
n2_E3_A4_KpnI_fw	CAGCggtaccATGGGTGAAGCTGCAGGAGCCGGTGCTGGTGCTTCTGCATTATATGAACATATGAATAC
n2_E4_KpnI_fw	CAGCggtaccATGGGTGCAGAAGCAGGAGCCGGTGCTGGTGCTTCTGCATTATATGAACATATGAATAC
n2_E5_KpnI_fw	CAGCggtaccATGGGTGCAGGAGAAGGAGCCGGTGCTGGTGCTTCTGCATTATATGAACATATGAATAC
n2_E7-9_KpnI_fw	CAGCggtaccATGGGTGCAGGAGCAGGAGAAGAAGAGGGTGCTTCTGCATTATATGAACATATGAATAC
pArl 128 rv	AGCTATTTACATGCATGTGCATGCAC
pARLminus rv	CAGTTATAAATACAATCAATTGG
pArl sense 55	GGAATTGTGAGCGGATAACAATTTCACACAGG
REX21-10mTRAP/SERA733-?	CAGCggtaccATGAAAATGTATTTAGCTGAAATTTTTAGTGGTGCAAATCCAGGAGACGAACGTGAATCTTCTGCATTATATGAACATATGAATAC
REX2/SERA7$^{25-32}$_ANN_KpnI_fw	CAGCggtaccATGAAAATGTATTTAGCTGAAATTTTTAGTGCAGCTGCTGCCGATAGTACTGTTTCTGCATTATATGAACATATGAATAC
REX2/SERA7$^{25-32}$_NAN_KpnI_fw	CAGCggtaccATGAAAATGTATTTAGCTGAAATTTTTAGTAAGCCTGCAGCTGCTGCCACTGTTTCTGCATTATATGAACATATGAATAC
REX2/SERA7$^{25-32}$_NNA_KpnI_fw	CAGCggtaccATGAAAATGTATTTAGCTGAAATTTTTAGTAAGCCTCCTCCTGCAGCTGCTGCCTCTGCATTATATGAACATATGAATAC
REX2/SERA7$^{25-32}$_K_KpnI_fw	CAGCggtaccATGAAAATGTATTTAGCTGAAATTTTTAGTAAGCCTCCTCCTGATAGTACTGTTTCTGCATTATATGAACATATGAATAC
SERA723-32Q23A/REX211-?	CAGCggtaccATGGCAGAAAAGCCTCCTCCTGATAGTACTGTTTCTGGTAAAGAGTCTTTGTTATCTTTAAAGTCTGCATTATATGAACATATGAATAC
SERA7Q23A/REX2_ANN_KpnI_fw	CAGCggtaccATGGCAGAAGCAGCTGCTGCCGATAGTACTGTTTCTGGTAAAGAGTCTTTGTTATCTTTAAAGTCTGCATTATATGAACATATGAATAC
SERA7Q23A/REX2_P26-28A_KpnI_fw	CAGCggtaccATGGCAGAAAAGGCTGCTGCCGATAGTACTGTTTCTGGTAAAGAGTCTTTGTTATCTTTAAAGTCTGCATTATATGAACATATGAATAC
SERA7Q23A/REX2_NAN_KpnI_fw	CAGCggtaccATGGCAGAAAAGCCTGCAGCTGCTGCCACTGTTTCTGGTAAAGAGTCTTTGTTATCTTTAAAGTCTGCATTATATGAACATATGAATAC
SERA7Q23A/REX2_NNA_KpnI_fw	CAGCggtaccATGGCAGAAAAGCCTCCTCCTGCAGCTGCTGCCTCTGGTAAAGAGTCTTTGTTATCTTTAAAGTCTGCATTATATGAACATATGAATAC
SERA7$^{33-42}$/REX2_ANN_KpnI_fw	CAGCggtaccATGGCAGAAGCAGCTGCTGCCGGAGACGAACGTGAATCTTCTGGTAAAGAGTCTTTGTTATCTTTAAAGTCTGCATTATATGAACATATGAATAC
SERA7^{33-}	CAGCggtaccATGGCAGAAGGTGCAAATGCAGCTGCTGCCCGTGAATCT

Name	Sequence with restriction site (written in lower-case characters)
^{42}/REX2_NAN_KpnI_fw	TCTGGTAAAGAGTCTTTGTTATCTTTAAAGTCTGCATTATATGAACATATGAATAC
SERA7$^{33-42}$/REX2_NNA_KpnI_fw	CAGCggtaccATGGCAGAAGGTGCAAATCCAGGAGACGCAGCTGCTGCCTCTGGTAAAGAGTCTTTGTTATCTTTAAAGTCTGCATTATATGAACATATGAATAC
SERA7$^{33-42}$/REX2_K_KpnI_fw	CAGCggtaccATGGCAGAAGGTGCAAATCCAGGAGACGAACGTGAATCTTCTGGTAAAGAGTCTTTGTTATCTTTAAAGTCTGCATTATATGAACATATGAATAC
SSGSSGSSGS_KpnI_fw	CAGCggtaccATGTCATCTGGTTCATCAGGATCCTCTGGTTCATCTGCATTATATGAACATATGAATAC
Stev$^{44-53}$/Ste$^{54-63}$$_{scr}$_f	CAGCggtaccATGGCACAAACCCAAATCCATAATCCACATTATAAACATGATTTACACCAACATGCTAATCCTTCTGCATTATATGAACATATGAATAC
Ste44-53scr/Ste54-63_f	CAGCggtaccATGTATACAAACATCGAACAAATTCCAATAGAGCATAATGATCCAGAACTCAAAGAAATAATTTCTGCATTATATGAACATATGAATAC
Stev$^{44-53}$Q45A/Ste$^{54-63}$$_{scr}$_f_KpnI	CAGCggtaccATGGCAGCAACCCAAATCCATAATCCACATTATAAACATGATTTACACCAACATGCTAATCCTTCTGCATTATATGAACATATGAATAC
Ste$^{44-63}$_Q45A_Q47A_KpnI_fw	CAGCggtaccATGGCAGCTACCGCAATCCATAATCCACATTATCATAATGATCCAGAACTCAAAGAAATAATTTCTGCATTATATGAACATATGAATAC
Ste$^{44-63}$_45-47A_KpnI_fw	CAGCggtaccATGGCAGCTGCTGCAATCCATAATCCACATTATCATAATGATCCAGAACTCAAAGAAATAATTTCTGCATTATATGAACATATGAATAC
STEVOR$_{44-63}$-_Q45G_KpnI_fw_neu	CAGCggtaccATGGCAGGAACCCAAATCCATAATCCACATTATCATAATGATCCAGAACTCAAAGAAATAATTTCTGCATTATATGAACATATGAATAC
STEVOR$_{44-63}$-_Q45G_Q47G_KpnI_fw	CAGCggtaccATGGCAGGAACCGGTATCCATAATCCACATTATCATAATGATCCAGAACTCAAAGAAATAATTTCTGCATTATATGAACATATGAATAC

2.2 Methods

2.2.1 Sterilisation

All media, buffers, solutions, glass materials and pipette tips were autoclaved at 121°C and 1,5bar vapour pressure for 20min. Heat instable solution were sterilised by filtration using sterile filters with a pore size of 0,22µm.

2.2.2 Molecular biological methods

2.2.2.1 Production of competent *E. coli*

For the production of competent *E. coli* cells a 10ml ON culture was carefully inoculated in fresh LB medium without antibiotics using a glycerol stock of the desired bacterial strain. On the next day 4ml of the ON culture were added to 200ml prewarmed LB medium without antibiotics and shaken for approximately 2h at 37°C until the OD_{600} reaches a value of 0,5 to 0,55. Then the *E. coli* culture was immediately placed on ice and centrifuged at 2.400g and 4°C for 20min. The supernatant was aspirated. The pellet was resuspended in 60ml of cold TFBI buffer and incubated on ice for 10min. The solution was again pelleted at 2.400g and 4°C for 20min. Then the pellet was resuspended in 8ml of TFBII buffer and 100µl aliquots were prepared and stored at -80°C.

2.2.2.2 Polymerase chain reaction (PCR)

Polymerase chain reaction was used to amplify DNA fragments. Therefore genomic DNA (gDNA), complementary DNA (cDNA) or plasmid DNA was used as a template. Together with dNTPs, a thermostable polymerase and a pair of oligonucleotids, which flanks both sites of the fragment (primer), the DNA was amplified under buffered conditions. The sequences for the cloning primers as well as for the primers to sequence the plasmid afterwards, were found in the *Plasmodium* database plasmoDB. The DNA was amplified through consecutive cycles of different temperatures. First the DNA was denatured at 94°C (20sec). Next the primer annealed to the template DNA at 48-72°C (20sec), depending on the primer length. Last the synthesis of the new DNA took place at 72°C. The polymerase needed 30sec for a synthesis of 1000 base pairs. The polymerase also showed a proof reading function to avoid unwanted mutations. The mutation rate is quoted with $4x10^7$ mutations.

Standard PCR protocol	
0,3µl	template
5µl	dNTPs
10µl	5x HF buffer
2µl	forward primer
2µl	reverse primer
0,3µl	Phusion DNA polymerase
30,4µl	dH_2O

Standard PCR programme			
PCR step	Temperature	Time	Cycles
Initialising step	94°C	2min	1x
Denaturation step	94°C	30sec	30x
Annealing step	60°C	30sec	
Elongation step	72°C	0,5-3min	
Final elongation	72°C	5min	1x
Final hold	4°C	∞	

2.2.2.3 Purification of PCR products

To purify the DNA from primers, the polymerase, leftover dNTPs and salts, the NucleoSpin ExtractII Kit was used according to manufacturer's protocol.

2.2.2.4 Restriction digest

The synthesized PCR products and the vectors were digested with restriction endonucleases – depending on the used restriction sites – to prepare the constructs for the ligation. For the vector pARL_GFP the enzymes KpnI und AvrII were used, for other vectors different restriction enzymes were applied. The insert was digested in addition with DpnI, a restriction enzyme, which specifically digests methylated DNA and therefore digests the template in the insert reaction mix (Lacks & Greenberg, 1975). The digest was incubated for 1 hour at 37°C and afterwards purified using the NucleoSpin ExtractII Kit as indicated in the manual.

Standard restriction digest	
30µl	purified PCR product
6µl	10x restriction buffer
1µl	enzyme 1
1µl	enzyme 2
22µl	dH_2O

2.2.2.5 Agarose gel electrophoresis

DNA fragments were separated within an electric field depending on charge, conformation, agarose concentration and applied field intensity. For a 1% agarose gel 2g of agarose were cooked in 200ml TAE buffer, cooled down to approximately 60°C, 5µl ethidium bromide (EtBr; 0,6µg/ml) were added. Then the liquid was poured into a gel

chamber. Combs were inserted. After 20min the gel was hardened, transferred into TAE buffer and DNA was loaded. EtBr is a DNA intercalating fluorescent dye, which makes the detection of DNA fragments within an UV bench possible. The size of the DNA fragments could be detected through comparison with a used marker (1kb GeneRuler), which contains DNA fragments with defined length. The separation was carried out at 10V/cm for 20min.

2.2.2.6 Ligation of DNA fragments

For ligating the digested insert and the digested vector these two components were applied in a molar ratio of approximately 3:1 and incubated with T4 DNA ligase and T4 ligase buffer (both NEB) for 30min at RT.

Standard ligation reaction	
7µl	digested insert
1µl	digested vector
1µl	10x ligation buffer
1µl	T4 DNA ligase enzyme 2

2.2.2.7 Transformation of competent *E. coli*

For transformation an aliquot of competent *E. coli* cells was thawed on ice. Afterwards 5µl of the ligation product were added to the thawed bacteria and incubated for 30min on ice. Then a heat shock of 42°C for 45sec was carried out. The bacteria were then placed on ice for 2min, 1ml of prewarmed medium was added and the mixture was shaken at 1400rpm for 30min. The transfected *E. coli* were pelleted at 3000g for 30sec, resuspended in ~50µl medium and plated on an LB plate with ampicillin or kanamycin – depending on the resistance cassette within the vector – and incubated at 37°C ON. On the next day single colonies were picked and characterized.

2.2.2.8 Colony PCR

To screen different colonies for the right insert, a colony PCR was done. Therefore a primer pair was used, where one oligonucleotide binds within the vector and the other one within the insert. First a single colony was plated on an agar plate using a sterile

10µl pipette tip. The plate was incubated at 37°C ON. Afterwards the pipette tip was transferred into a colony PCR reaction tube to perform a colony screen.

Standard PCR protocol	
1µl	dNTPs
1µl	10x Firepol. buffer
0,6µl	$MgCl_2$
0,4µl	forward primer
0,4µl	reverse primer
0,1µl	Firepol DNA polymerase
6,5µl	dH_2O

Standard PCR programme			
PCR step	Temperature	Time	Cycles
Initialising step	94°C	2min	1x
Denaturation step	94°C	40sec	25x
Annealing step	48°C	40sec	
Elongation step	68°C	1-3min	
Final elongation	68°C	5min	1x
Final hold	4°C	∞	

2.2.2.9 Culture and storage of *E. coli*

Two to four positive colonies were picked with a sterile 100µl pipette tip from the incubated agar plate (see section 2.2.2.8) and transferred into a 2ml reaction tube filled with LB^{Amp} or LB^{Kan} medium, depending on the resistance cassette of the particular vector. The medium containing the bacteria was incubated ON at 37°C and 800rpm. For the preparation of a glycerol stabilate of the clone 1ml of the overnight culture was centrifuged at 3000g for 30sec. The pelleted bacteria were then resuspended in a glycerol-medium-mix (1:1) and stored at -80°C. The plate was stored in the fridge at 4°C for about two weeks.

2.2.2.10 Isolation of plasmid DNA (mini or midi preparation)

To isolate DNA for verification of the clones, the respective *E. coli* clones were inoculated (minis). Therefore a colony was picked as described above or a pipette tip was used to take a few µL of the glycerol stock and transfer it into a 2ml reaction tube, filled with LB^{Amp} or LB^{Kan} medium depending on the resistance cassette of the particular vector. Incubation took place ON at 37°C and 800rpm. On the next day the DNA was isolated

with the NucleoSpin Plasmid Kit as described in the manual and eluted in 30µl elution buffer. For transfecting *Plasmodium falciparum*, a larger amount of DNA was needed. For this reason 200ml medium were incubated with a single colony from a glycerol stock ON at 37°C and 800rpm. To isolate the DNA, the midi kit was used according to the manufacturer's protocol. The DNA was solved in 200µl TE buffer.

2.2.2.11 Isolation of plasmid DNA using STET buffers

When a high number of DNA mini preparations was needed, e.g. to analyse a greater number of clones for their inserts, STET minis were generated. The isolation of plasmid DNA using STET buffer is fast, but does not yield high purity of the DNA. Therefore all desired clones were inoculated (see 2.2.2.10), incubated ON and then pelleted at 6000g for 1min. Next the pellets were resuspended in 500µl STET buffer each, 50µl lysozyme (10mg/ml, in TE) were added to each reaction tube, the minis were incubated for 2 to 3min at RT and afterwards for 1 to 2min at 95°C. Next the lysates were centrifuged at 16000g for 5 to 10min and the supernatant was transferred into a new 1,5ml reaction tube. The precipitation of the DNA was done by adding 50µl of 6,5M NaAc and 500µl of isopropanol. After vortexing shortly, the DNA was centrifuged at 16000g for 5 to 10min. Afterwards the pellet was air-dried for 5min at RT and resuspended in 50µl TE (+RNAse, 50µg/ml). For verification of the plasmid DNA, it was premixed with the respective primer and sequenced, using the services of Seqlab.

2.2.2.12 Determination of DNA concentration

To determine the concentration of the isolated DNA the absorbance of a 1:100 dilution is measured in a photometer at 260nm. An OD of 1 corresponds to a concentration of double stranded DNA of 50µg/ml. To determine the purity of the DNA, the quotient of the absorption of DNA at 260nm and of proteins at 280nm was calculated. Pure DNA shows a value of approximately 1,8. Lower values point to a contamination with proteins, higher values indicate a contamination with RNA.

2.2.2.13 Precipitation of DNA

DNA was precipitated by inverting with 0,1 volumes sodium acetate (3M) and 3 volumes 100% ethanol. Afterwards the DNA was centrifuged at 16000g for 10min. Subsequently the supernatant was removed and the DNA was washed in 70% Ethanol. After centrifugation, removal of the supernatant and air-drying of the pellet, the opaque DNA pellet was resuspended in the desired volume of TE buffer.

2.2.3 Biochemical methods

2.2.3.1 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

To separate proteins the SDS Polyacrylamide gel electrophoresis (SDS-PAGE) was used. Therefore SDS gels with an acrylamide concentration of 12% or 15% were prepared. 4x SDS loading buffer was added to the protein mixes of the parasite pellets and incubated at 95°C for 5min. The separation was carried out at 200V for 1h in SDS electrophoresis buffer. By applying a marker, the molecular weight of the proteins could be detected.

2.2.3.2 Coomassie Brilliant Blue staining

To detect proteins in a polyacrylamide gel, it was stained in a Coomassie solution for 30min. Afterwards the gel was destained with water or Coomassie destainer solution until the background staining was low and the blue protein bands were visible.

2.2.3.3 Western blot analysis

Proteins separated by SDS-PAGE could also be transferred to a nitrocellulose membrane by western blotting. For this reason the membrane was soaked using transfer buffer or CAPS buffer. Three layers of chromatography paper were also soaked using transfer buffer or CAPS buffer, afterwards the nitrocellulose membrane was placed on the three layers, the gel was placed on the membrane and covered with three more moistened chromatography papers. The protein transfer was carried out at 100V for 1h or at 15V per gel ON. Next the membrane was incubated with 5ml blocking solution by rolling for 1h at RT. The primary antibody was diluted in 5ml blocking solution. The membrane was incubated with this solution by rolling for 1h at RT or ON at 4°C. After washing the

membrane 3 to 5 times with approximately 5ml washing buffer for 3 to 5min per wash, the HRP conjugated secondary antibody was diluted in blocking solution, applied to the membrane and incubated for 1h at RT. The membrane was again washed 3 to 5 times with approximately 5ml washing buffer for 3 to 5min per wash. Next the membrane was incubated with 2ml ECL solution and incubated for 5min at RT by rolling. Then the membrane was placed between two transparent overhead transparencies and laid in a developer cassette. The chemiluminescence activity was used to expose a radiographic film.

2.2.4 Cell biological and molecular procedures for *P. falciparum*

2.2.4.1 *P. falciparum* culture

Blood stages of *P. falciparum* (strain 3D7) were cultivated according to standard procedures (Trager and Jensen, 1976). The parasites were kept in 5ml or 10ml petri dishes using complete RPMI medium and human erythrocytes of the blood type 0+ (5% haematocrit) in an oxygen low atmosphere (5%CO_2, 1%O_2, 94%N_2) at 37°C. The parasite culture was diluted frequently with fresh medium and fresh blood so that it did not exceed a parasitaemia of 10%. For selection, the drugs WR99210 and/or blasticidin S were added to transgenic parasite cultures.

2.2.4.2 Giemsa staining

To determine the parasitaemia, Giemsa stained blood smears were prepared and counted by means of a Zeiss Axio Lab.A1 microscope. Therefore 0,5µl blood of a parasite culture were dropped on a glass slide. Using the edge of a second glass slide the parasites were equally distributed by smearing. After air-drying, the smear was fixed for 20sec in methanol. Next the slide was placed in a 10% Giemsa solution. By Giemsa staining the cytoplasm of the erythrocytes was stained blue, the parasites were stained red and the DNA of the parasites was stained dark blue.

2.2.4.3 Freezing of parasites (Cryopreservation)

For long-term storage of *P. falciparum*, parasite cultures exhibiting approximately 5% ring stage parasites were pelleted by centrifugation at 1.800g for 3min. Then the

medium was aspirated, the pellet was resuspended in 1ml freezing solution and transferred into a cryotube. Storage was carried out at -80°C or in liquid nitrogen at -196°C.

2.2.4.4 Thawing of parasites

The cryotube containing the frozen *P. falciparum* cultures were thawed in a 37°C water bath and afterwards pelleted by centrifugation at 1800g for 3min. The supernatant was aspirated, the blood pellet resuspended in 1ml prewarmed thawing solution and centrifuged again. The pellet was then washed in 1ml fresh medium, resuspended in 5ml fresh medium and transferred to a petri dish. The haematocrit was adjusted to 5% by adding blood.

2.2.4.5 Synchronisation

To generate a synchronous *P. falciparum* culture, showing mainly one of the different stages, the parasites (at least 5% ring stage parasites) were pelleted at 1.800g for 3min. The supernatant was aspirated afterwards and the pellet resuspended in 5 volumes of synchronisation solution and incubated in a water bath at 37°C for 5min. By using synchronisation solution the trophozoites and schizonts lyse through osmosis. Rings and uninfected red blood cells remain intact. The cells were centrifuged, washed with 10ml fresh medium, centrifuged again, resuspended in 10ml medium and transferred into a new petri dish. The haematocrit was adjusted to 5% by adding blood.

2.2.4.6 Transfection of *P. falciparum*

The transfection of *P. falciparum* blood stages is associated with some difficulties because the external DNA has to pass the erythrocyte membrane, the parasitophorous vacuole membrane, the parasite plasma membrane and the nucleus membrane. To achieve this, electroporation was performed. Here, a strong electrical field generates temporary pores in the membranes for a short time frame to let the DNA pass.

2.2.4.6.1 Transfection (ring stage parasites)

For transfecting DNA into ring stage parasites, 5 ml of a synchronous culture of ring taged parasites showing a parasitaemia of 5-10% were pelleted at 1.800g for 3min. Meanwhile 100µg of plasmid DNA was precipitated, washed with 70% ethanol and air-dried sterilely. The pellet was then resuspended in 15µl TE buffer and 385µl transfection buffer was added. The DNA solution was then added to the blood pellet and transferred into a 20mm cuvette. The electroporation was carried out at 310V und 950µF. Afterwards the cells were transferred to a new petri dish, with 12ml fresh and prewarmed medium. After 5h the medium was changed and the selection for transgenic parasites was started by adding a drug (WR99210 and /or blasticidin S). For 5 days the medium was changed daily, then every second day. Transgenic parasites could be detected 3 to 12 weeks after successful transfection in Giemsa-stained blood smears.

2.2.4.6.2 Transfection (schizont stage parasites)

For transfecting DNA in schizonts, 10ml of an asynchronous schizont culture showing a parasitaemia of 5-10% were pelleted at 1.800g for 3min. The blood pellet was resuspended in 8ml fresh medium and carefully added on top of 4ml of a 60% percoll solution (see 2.2.4.9). The cells were centrifuged at 2000g for 8min. After centrifugation a layer of schizonts could be detected, taken off and washed once in medium (approximately 25µl of schizonts). During centrifugation 50µg of plasmid DNA was precipitated, washed with 70% ethanol and air-dried sterilely. The pellet was then resuspended in 10µl TE buffer and 100µl transfection buffer II was added. 12µl of schizonts were then added to the DNA solution, well resuspended and transferred into a 20mm cuvette. The electroporation was carried out in an Amaxa Biosystems Nucleofector II from Lonza using the programme U-033. Afterwards the transfected cells were transferred into a 1,5ml reaction tube with 300µl of fresh blood and 500µl of prewarmed medium. After shaking at 600rpm at 37°C for 30min the cells were transferred into a new petri dish with 12ml fresh prewarmed medium. After 24h the medium was changed and the selection for transgenic parasites started by adding a drug (WR99210 and /or blasticidin S). For 5 days the medium was changed daily, then every second day. Transgenic parasites could be detected 8 days to 4 weeks after successful transfection in Giemsa-stained blood smears.

2.2.4.7 Saponin lysis

To isolate *P. falciparum* parasites from a blood culture, saponin was used. Saponin lyses the erythrocyte membrane as well as the parasitophorous vacuole membrane but not the parasite plasma membrane. To achieve this 10ml of a parasite culture was centrifuged at 1.800g for 3min, the supernatant was removed and the pellet was resuspended in 4ml saponin lysis buffer and incubated on ice for 5 to 20min. Then the lysate was centrifuged at 16000g for 5min and the parasite pellet was washed 3 times in 1x PBS. 2µl protease inhibitor were added and the pellet was resuspended in 20 to 200µl parasite lysis buffer, depending on the size of the pellet. The pellet was frozen at -20°C.

2.2.4.8 Isolation of genomic DNA

After the saponin lysis of parasites genomic DNA could be isolated. This was done by means of the Qiagen QIAamp DNA Mini Kit according to manufacturer's protocol and eluted in 200µl elution buffer.

2.2.4.9 Percoll gradient

By means of a percoll gradient uninfected erythrocytes as well as the different stages of *P. falciparum* parasites could be separated. Therefore 10ml of a parasite culture were centrifuged at 1.800g for 3min, the supernatant was removed and the pellet was washed once in 1x PBS. Meanwhile 550µl of an 80% percoll solution were filled in a 2ml reaction tube, followed by 600µl of a 60% percoll solution and 450µl of a 40% percoll solution. A careful layering of the percoll solutions led to a stable separation of the different phases. The blood pellet was resuspended in 300µl 1x PBS and transferred on top of the percoll gradient. After immediate centrifugation at 16000g for 5min, the desired parasite stages were transferred into a 1,5ml reaction tube and washed with 1x PBS until all percoll solution was removed.

To perform a percoll with a higher amount of parasite culture, a 50ml centrifugation tube was filled carefully with 10ml 80%, 10ml 60% and 10ml 40% percoll solution consecutively. The parasite culture was centrifuged, washed and added slowly. Centrifugation at 4000g was performed for 10min and the desired parasite stages were

washed until all percoll solutions were removed. The parasites were cultured, e.g. for transfection of parasites with DNA or analysis by SDS-PAGE.

2.2.4.10 Quantification of protein export rates in *P. falciparum*

For quantifying the export rates of different transgenic parasite cell lines, living parasites were analysed (see 2.2.5.1). The investigated protein was GFP-tagged and therefore visible by fluorescence microscopy. 50 cells per cell line were counted on three occasions by two counters in a blinded manner. The GFP tagged protein was graded into three groups: parasites, which showed no export (protein confined to the parasite or its PV), partial export (GFP signal in the parasite and/or in the PV and in the host cell) or full protein export (GFP signal only in the host cell). Statistical analyses were performed using Microsoft Excel.

2.2.5 Microscopy

2.2.5.1 Live cell imaging

For fluorescence microscopy analysis of living parasites 500µl of a parasite blood culture were incubated with DAPI (1µg/ml) for 10min at RT. Then one drop was pipetted on a glass slide, covered with a cover slip and immediately analysed by fluorescence microscopy. Therefore the Zeiss Axioskop M1 or M2 microscope with a 100x/1,4 immersion oil lamp, a Hamamatsu Orca-ER C4742-80 digital camera or a Hamamatsu Orca-R_2 C10600 digital camera and the Zeiss Axiovision programme were used. The pictures were edited using Corel Photo Paint.

3 Results

3.1 The reporter system to analyse export promoting capacity in N-termini, and definitions

At the second position of the N-terminus after PEXEL motif cleavage (E, Q or D), the primary (1°) positive signal is located. This signal mediates export of the protein (see section 1.3.5 and Figure 1.6) but export in N-termini where this motif was mutated indicated additional export promoting sequences further downstream (Grüring *et al*., 2012; Tarr *et al*., 2013). These downstream signals we here refer to as the secondary (2°) positive signal. To analyse the capacity of N-termini to promote export, a truncated mTRAP-GFP fusion reporter was previously used (Figure 3.1A; Haase *et al*., 2009; Grüring *et al*., 2012; Heiber *et al*., 2013). This reporter harbours the transmembrane domain of the PNEP REX2 (Haase *et al*., 2009). Fusion proteins were episomally expressed in blood stage parasites of *P. falciparum*. The N-terminus of exported proteins N-terminally fused to the reporter resulted in the export of the protein, whereas N-termini of non-exported proteins did not promote export of the fusion construct (Grüring *et al*., 2012). The localisation of the reporter can be assessed by fluorescence microscopy based on its GFP tag. To receive a more differentiated picture of the export promoting capacity of each N-terminus, the export levels of each construct can be quantified by counting the proportion of cells showing no export of the reporter (fluorescence confined to the parasite or PV), partial export (GFP signal in the parasite/PV and in the host cell) and full export (fluorescence only in the host cell). In this work, scoring of cells was done on three occasions by two counters (each counting at least 50 cells per occasion) in a blinded manner (see section 2.2.4.10).

To obtain a standardised definition of amino acid positions in the N-termini here analysed, the individual positions in the N-termini were numbered starting with the amino acid after the methionine (which usually is removed after a protein is synthesised). This also allows a direct comparison with the actual mature N-termini after signal peptide cleavage and PEXEL processing, where no such methionine is present. According to that, position 1 (P1) in the N-terminus represents the first amino acid after the methionine, P2 the second and so forth.

3.2 The blocking sequence within SERA7

To identify the common core export domain of PEXEL proteins and PNEPs a bioinformatics analysis was conducted including all N-termini of known exported proteins so far. As this *in silico* analysis did not reveal a consensus sequence of exported proteins (see Appendix), the investigation of N-termini and their properties regarding export was carried out based on previous experiments.

N-termini with a bona fide 1° export signal can harbour a blocking region, preventing export (see section 1.3.5 and Figure 1.6). The PV protein SERA7 is such an example, as previously demonstrated (Grüring *et al*., 2012), and its N-terminus was chosen to identify such a sequence. The SERA7 N-terminus contains a perfect primary signal in form of a glutamic acid (E) at P2 after the predicted signal peptide cleavage site. After cleavage it would consequently mimic a processed PEXEL motif, corresponding to position P23 of the uncleaved protein. However, the amino acid before the E, a Q at P1, was not typical for PEXEL proteins and therefore previously replaced by an alanine, which however did not result in export (Grürig *et al*., 2012).

To narrow down which other part in this N-terminus might prevent export (i.e. the blocking region), it was first assessed whether this block-inducing sequence was in the first or the second 10 amino acids of this N-terminus. To do this, the first ten amino acids of the SERA N-terminus (already carrying the Q to A mutation in position 1, corresponding to Q23A in the uncleaved SERA7 sequence) were fused to the amino acids 10-20 of the PNEP REX2, a region not essential for the export of the protein (Haase *et al*., 2009), resulting in the N-terminus SERA7Q23A$^{23\text{-}32}$/REX2$^{11\text{-}20}$. This N-terminus N-terminally fused to the truncated mTRAP reporter (Figure 3.1A) resulted in only very poor export (Figure 3.1B), suggesting that the first amino acids of the mature SERA N-terminus contain an export blocking region. However, when the second part of the SERA7 mature N-terminus (amino acids 10-20) was tested by fusing it behind the first export promoting region of REX2 (REX2$^{1\text{-}10}$/SERA7$^{33\text{-}42}$), this construct was also poorly exported (Figure 3.1C). This indicates that this region also contains an export blocking sequence that had a negative influence on the normally exporting REX2$^{1\text{-}10}$ sequence. As both constructs did not result in an export, both regions of SERA7 appeared to contain sequences blocking export in backgrounds expected to be exported.

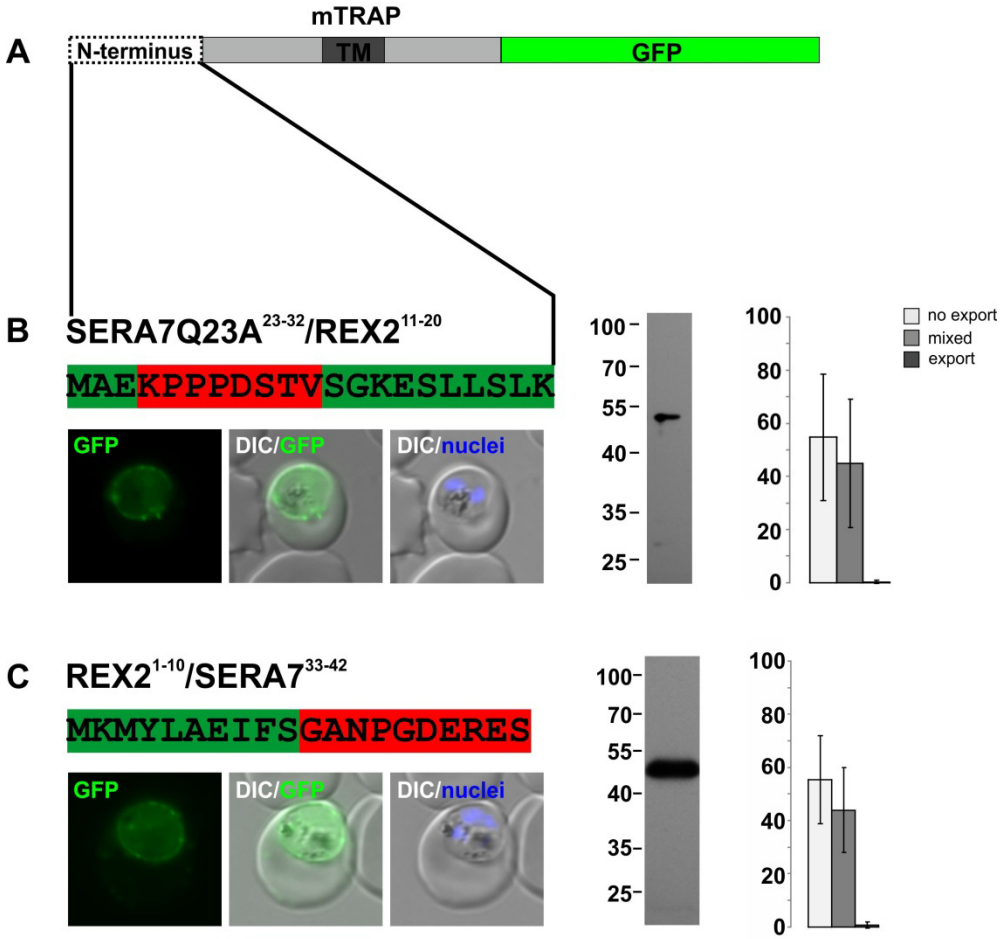


Figure 3.1 – Both the first and the second 10 amino acids of the SERA7 mature N terminus contain regions preventing export. (A) Schematic of the mTRAP reporter (TM, transmembrane domain). The dashed box labelled 'N-terminus' shows the position of the appended N-termini displayed in the cell lines shown below. **(B-C)** Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the N-terminus shown above each panel (sequences belonging to SERA7_Q23A are highlighted in red, sequences belonging to REX2, as well as the primary positive signal AE, in green). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5µm. Western blots to the right of the image panels with extracts of the cell lines corresponding to the fusion construct probed with α-GFP. The size standard is indicated in kDa. Both constructs showed the expected size. The graphs to the very right show the proportion of cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3. Amino acids are shown in one letter code.

3.2.1 Narrowing down the export-blocking sequence in the N-terminus of SERA7

As the first 10 amino acids of the mature SERA7 N-terminus seem to be effective directly after the primary positive signal, this region was first analysed in more detail. Therefore region 10-20 of the mature SERA7 N-terminus was shifted to the very N-terminus with a bona fide PEXEL 1° signal (AE). Surprisingly, this resulted in export that was close to what usually was observed in N-termini with perfect primary signals (Figure 3.2B1). To more precisely identify the region involved, an alanine scan was done in the

corresponding region. Overlapping quadruplets of the first 10 amino acids of the SERA7 mature N-terminus were replaced by alanine stretches (AAAA) to neutralise the particular regions. As a perfect primary export signal was inserted, this covered P3 to P10 of this new N-terminus. The inserted neutral regions marginally increased the export rate further with the alanines immediately after the primary signal (Figure 3.2B2) and directly in front of the REX2 sequence (Figure 3.2B2, B4), resulting in very efficient export. Replacement of the amino acids in the middle of the SERA7 stretch did not improve export (Figure 3.2B3) compared to the parental construct (Figure 3.2B1). As the alanine scan of the second 10 amino acids of the mature SERA7 N-terminus located at the very N-terminus improved export, it may be possible that quadruple alanines independently enhance export if placed immediately after a perfect 1° signal.

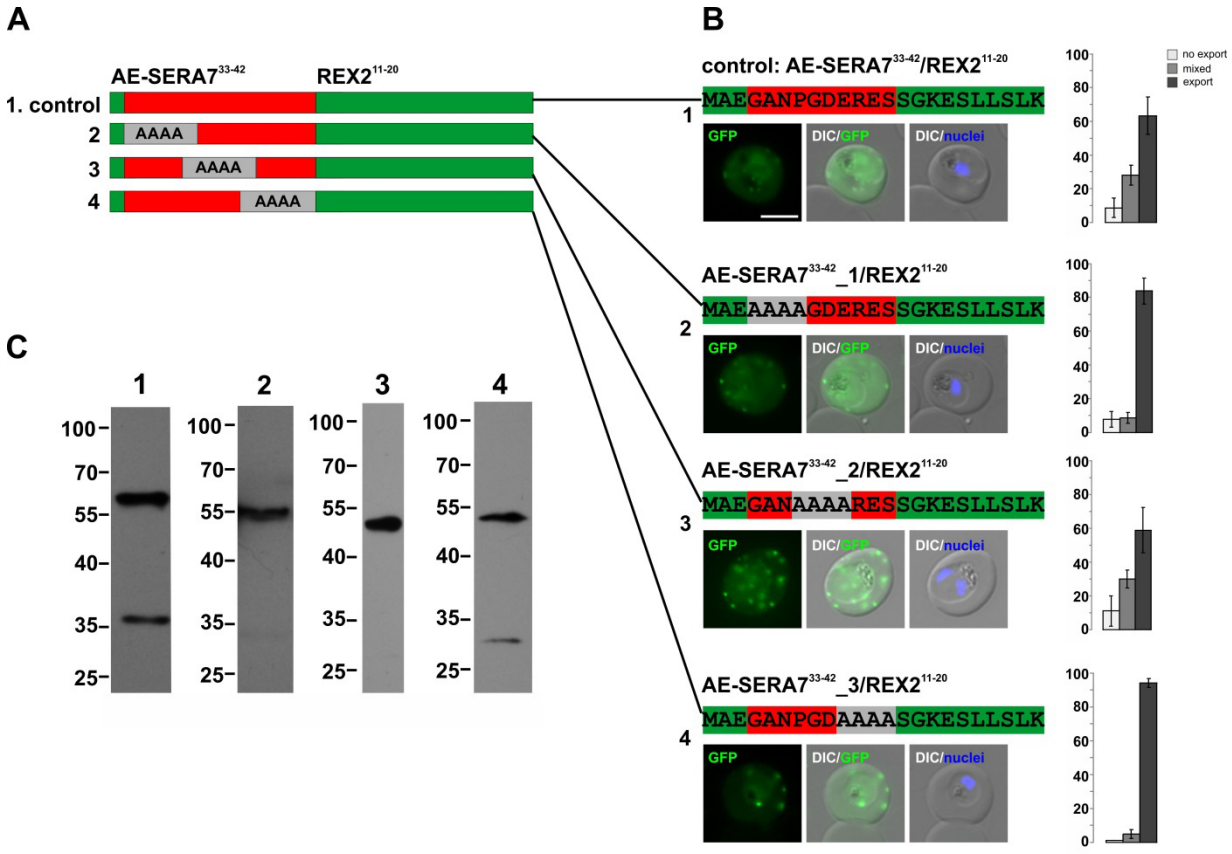


Figure 3.2 – Placing the second 10 amino acids of the N-terminus of SERA7 (AE-SERA7$^{33-42}$) in front results in export of the protein. AE-SERA7$^{33-42}$ was fused to the second part of the N-terminus of REX2 (REX2$^{11-20}$). **(A)** Alanine stretches were inserted into the part of the N-terminus belonging to SERA7. (1) represents the control, in (2) P3 – P6, in (3) P6 – P9 and in (4) P9 – P12 of the SERA7 N-terminus were mutated to alanines. This numbering refers to (B) and (C). **(B)** Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the N-terminus shown above each panel (sequences belonging to SERA7 are highlighted in red, sequences belonging to REX2, as well as the primary positive signal AE, in green, and the AAAA stretches in grey). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5µm. The graphs to the very right show the proportion of cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3. **(C)** Western blots with extracts of the cell lines corresponding to the fusion construct numbered according to A probed

with α-GFP. The size standard is indicated in kDa. All constructs showed the expected size. Amino acids are shown in one letter code.

As not only the position in the N-terminus alone but also the amino acid sequence itself seems to be decisive for protein export, the first 10 amino acids of the mature SERA7 N-terminus were investigated by means of an alanine scan. This region likely contains a blocking signal (Figure 3.1B). As the primary export signal was not changed, the investigated region covered P3 to P10. The alanine scan was done in the construct where the first part of the N-terminus of SERA7 was fused to the second part of the N-terminus of REX 2 (SERA7$^{23-32}$Q23A/REX2$^{11-20}$) (Figure 3.3A). It resulted in 3 constructs, where the alanines covered regions P3-P6, P5-P8 and P7-P10 of the mature N-terminus of SERA7. Unexpectedly all three of these constructs shifted the export profile from a none-exported phenotype (Figure 3.1B) to export (Figure 3.3B). This indicated that the entire region encompassing P3 to P10 was required to cause the export block and that already a partial disturbance abolished the blocking activity of this region.

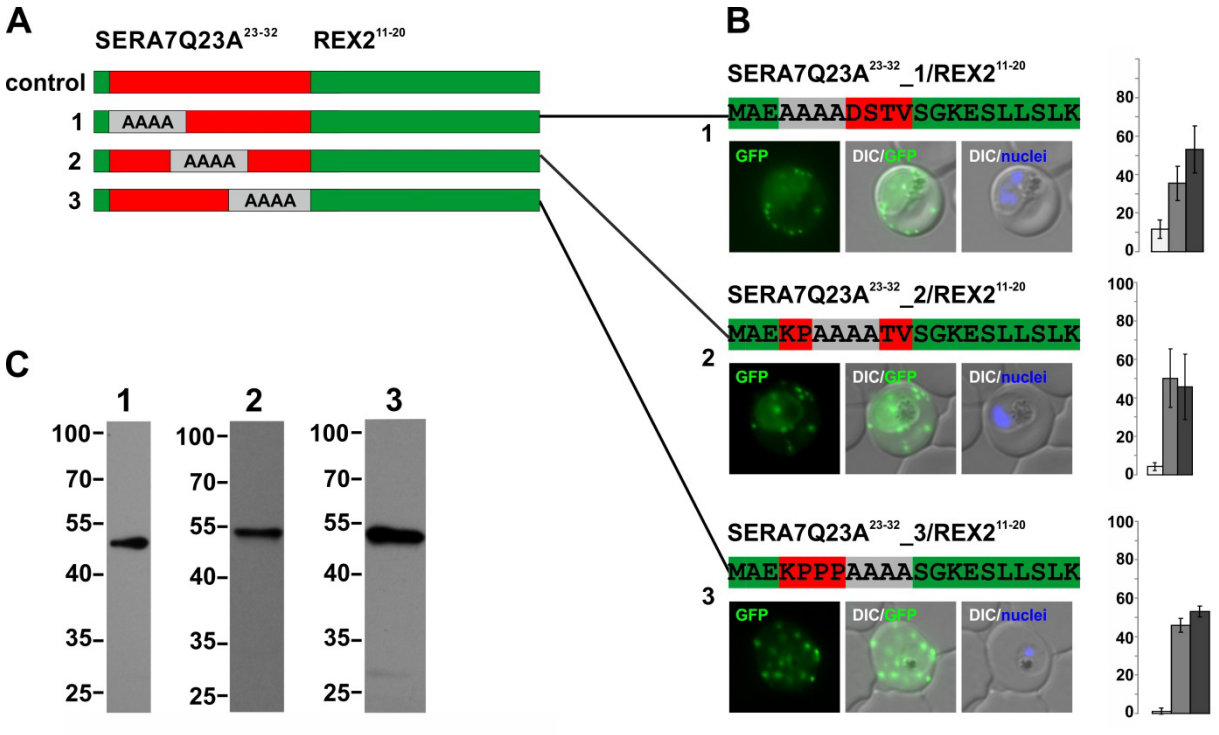


Figure 3.3 – Mutating regions in the first part of the N-terminus of SERA7 (SERA7$^{23-32}$_Q23A) results in higher export rates (of the reporter protein). SERA7$^{23-32}$_Q23A was fused to the second part of the N-terminus of REX2 (REX2$^{11-20}$). **(A)** Alanine stretches were inserted into the part of the N-terminus belonging to SERA7. The schematic of the control is shown at the top, results are represented in Figure 3.1B. In (1) P3 – P6, in (2) P5 – P8 and in (3) P7 – P10 of the SERA7 N-terminus were mutated to alanines. This numbering refers to (B) and (C). **(B)** Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the N-terminus shown above each panel (sequences belonging to SERA7 are highlighted in red, sequences belonging to REX2, as well as the primary positive signal AE, in green, and the AAAA stretches in grey). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges

are shown for DIC with GFP and DIC with DAPI. Size bar: 5μm. The graphs to the very right show the proportion of cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3. **(C)** Western blots with extracts of the cell lines corresponding to the fusion construct numbered according to A probed with α-GFP. The size standard is indicated in kDa. All constructs showed the expected size. Amino acids are shown in one letter code.

As the position within the N-terminus affected the blocking capacity of the second part (amino acids 10-20) of the mature SERA7 N-terminus (compare Figure 3.1B with Figure 3.2B1), it was tested next whether the blocking region in the first 10 amino acids of the mature SERA7 N-terminus was also functional further downstream. To do this, the first part of the N-terminus of SERA7 without the initial AE was placed after the export mediating first 10 amino acids of REX2 (REX2$^{1-10}$/SERA7$^{25-32}$, Figure 3.4A1). In accordance with the blocking capacity of that SERA7 region, this construct showed poor export, with 55% mixed export and 45% no export (Figure 3.4B1). This indicated that in contrast to the second part of the mature SERA7 N-terminus (compare Figure 3.1B with Figure 3.2B1), the first part can exert its blocking capacity in different positions after the primary signal, although to different extents. Next, the same alanine replacements, used for the constructs when the SERA7 region was N-terminal, were done in the REX2$^{1-10}$/SERA7$^{25-32}$ construct. In the first two constructs the replacement of parts of SERA7 by alanines moderately increased the rate of export compared to the parent (first six to eight amino acids of SERA7; Figure 3.4B2, B3). The last construct, holding the inserted alanines at the end of the N-terminus, showed even less export than the control (Figure 3.4B4). These results indicated that the first three to eight amino acids of SERA7 harbour a sequence that blocks protein export if placed further downstream.

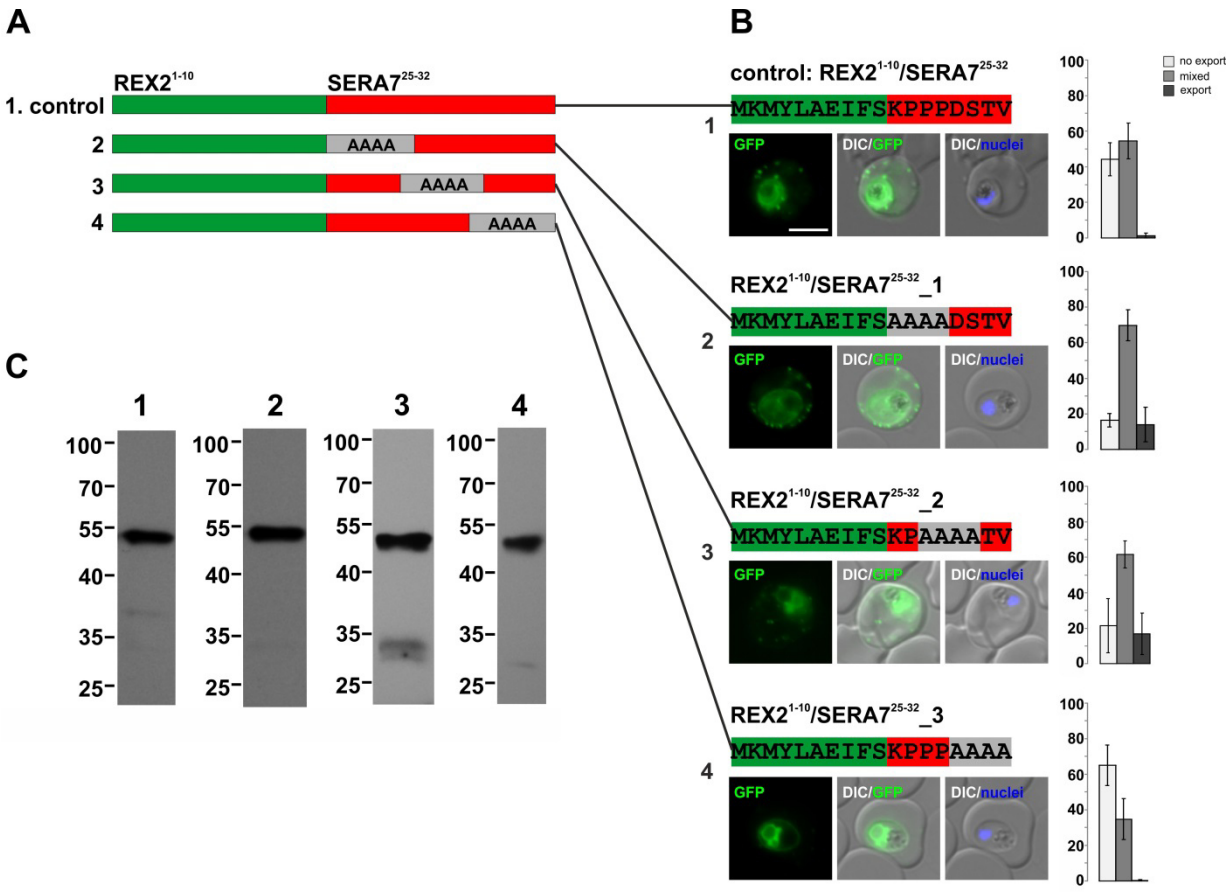


Figure 3.4 – Mutating the first six amino acids of SERA7$^{25-32}$ to alanines removes an export blocking region when placed further downstream in the N-terminus. The first part of the N-terminus of REX2 (REX2$^{1-10}$) was prepended to the first part of the N-terminus of SERA7 (SERA7$^{25-32}$) without the initial QE. **(A)** Alanine stretches were inserted into the part of the N-terminus belonging to SERA7. (1) represents the control, in (2) P10 – P13, in (3) P12 – P15 and in (4) P14 – P17 of the N-terminus were mutated to alanines. This numbering refers to (B) and (C). **(B)** Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the N-terminus shown above each panel (sequences belonging to REX2 are highlighted in green, sequences belonging to SERA7 in red and the AAAA stretches in grey). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5µm. The graphs to the very right show the proportion of cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3. **(C)** Western blots with extracts of the cell lines corresponding to the fusion construct numbered according to A probed with α-GFP. The size standard is indicated in kDa. All constructs showed the expected size. Amino acids are shown in one letter code.

3.2.2 Prolines as an export blocker

The most notable sequence in the export blocking region within the first three to eight amino acids of the N-terminus of SERA7 were three consecutive prolines, which are located exactly in the overlapping region of the alanine quadruplets restoring some export (Figure 3.3 and Figure 3.4). Thus the possibility was tested, if the prolines alone may be responsible for the block of export. Prolines are known to strongly affect the folding of proteins. The same background as before was used (Figure 3.1B), but this time only the three prolines were exchanged with alanines resulting in the construct

SERA7$^{23-32}$Q23A_P26-28A/REX2$^{11-20}$. In comparison to the control, which showed a mostly blocked exported phenotype (Figure 3.1B), the replacement of the prolines by alanines resulted in full export of the protein in more than 80% of the cells (Figure 3.5). Hence, removing the prolines alone was sufficient to change the N-terminus of SERA7 into an export promoting region. However, as all other changes in this region restored export (Figure 3.3), it was unclear if the prolines by themselves were sufficient to prevent the export of N-termini with perfect primary signals.

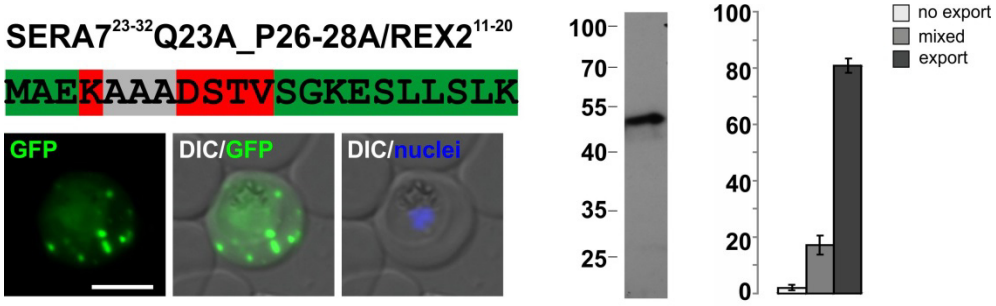


Figure 3.5 – Mutation of prolines within the N-terminus of SERA7 results in export of the protein. Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the N-terminus shown above the image panel (sequences belonging to SERA7 are highlighted in red, sequences belonging to REX2, as well as the primary positive signal AE, in green, and the replacement of the prolines by alanines is highlighted in grey). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5µm. The western blots to the right of the image panel with the extract of the cell line corresponding to the fusion construct probed with α-GFP. The size standard is indicated in kDa. The construct showed the expected size. The graph to the very right shows the proportion of cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3. Amino acids are shown in one letter code.

3.3 Design of a neutral N-terminus 'n'

The results in search of the blocking region demonstrated that the amino acid arrangement, the neighbouring amino acids and the position within the N-terminus significantly influence the export promoting or blocking capacity of an N-terminus. To enable the investigation of different amino acids and their position in isolation, the design of a neutral N-terminus was attempted (Figure 1.6E). This would make possible to add specific amino acids into a neutral background to then assess the effect on export. Alanine is considered the most neutral amino acid, as its side chain is small but large enough to not negatively affect protein folding. First, an N-terminus consisting of ten alanines in a row was examined (AAAAAAAAAA). Again, this N-terminus was N-terminally appended to the truncated mTRAP reporter (Figure 3.1A). However, this construct was almost undetectable by fluorescence microscopy, indicating poor

expression or rapid degradation (data not shown). For this reason, the alanines were next combined with other amino acids. First serines (S) and threonines (T; amino acids considered not to promote export but frequently present in exporting N-termini) were stringed together with alanines (A) in consecutive repeats, resulting in an N-terminus with the sequence mastastasta. However, as this N-terminus promoted export of the protein (Figure 3.6A), it was not suitable for further investigations. Next different N-termini using alanine (A), glycine (G) and serine (S), avoiding the larger threonine (T), were tested in different arrangements. This comprised N-termini consisting of repeats of either, AAS, SSG or AG with a total length of ten amino acids. The N-terminus consisting of AAS-repeats showed partial to full export with some cells showing a prominent pool in the parasite's ER and was therefore not appropriate as a neutral N-terminus (Figure 3.6B). The N-terminus consisting of SSG-repeats, as well as the N-terminus consisting of AG-repeats, did not promote export of the protein reporter (Figure 3.6C, D). However, the construct with the SSG-repeats showed only soluble staining within the parasite (Figure 3.6C). This indicated that the protein did not enter the secretory pathway. This construct was therefore also not suitable as a neutral N-terminus. The AG-repeats led to an accumulation of the GFP fluorescence around the parasite, typical for transport of the fusion protein to the PV (Figure 3.6D). This indicated efficient recruitment into the secretory pathway but absence of other trafficking signals, as the PV is the default destination of such proteins (Waller *et al*., 2000; Deponte *et al*., 2012). Thus the AG-N-terminus appeared to be truly neutral.

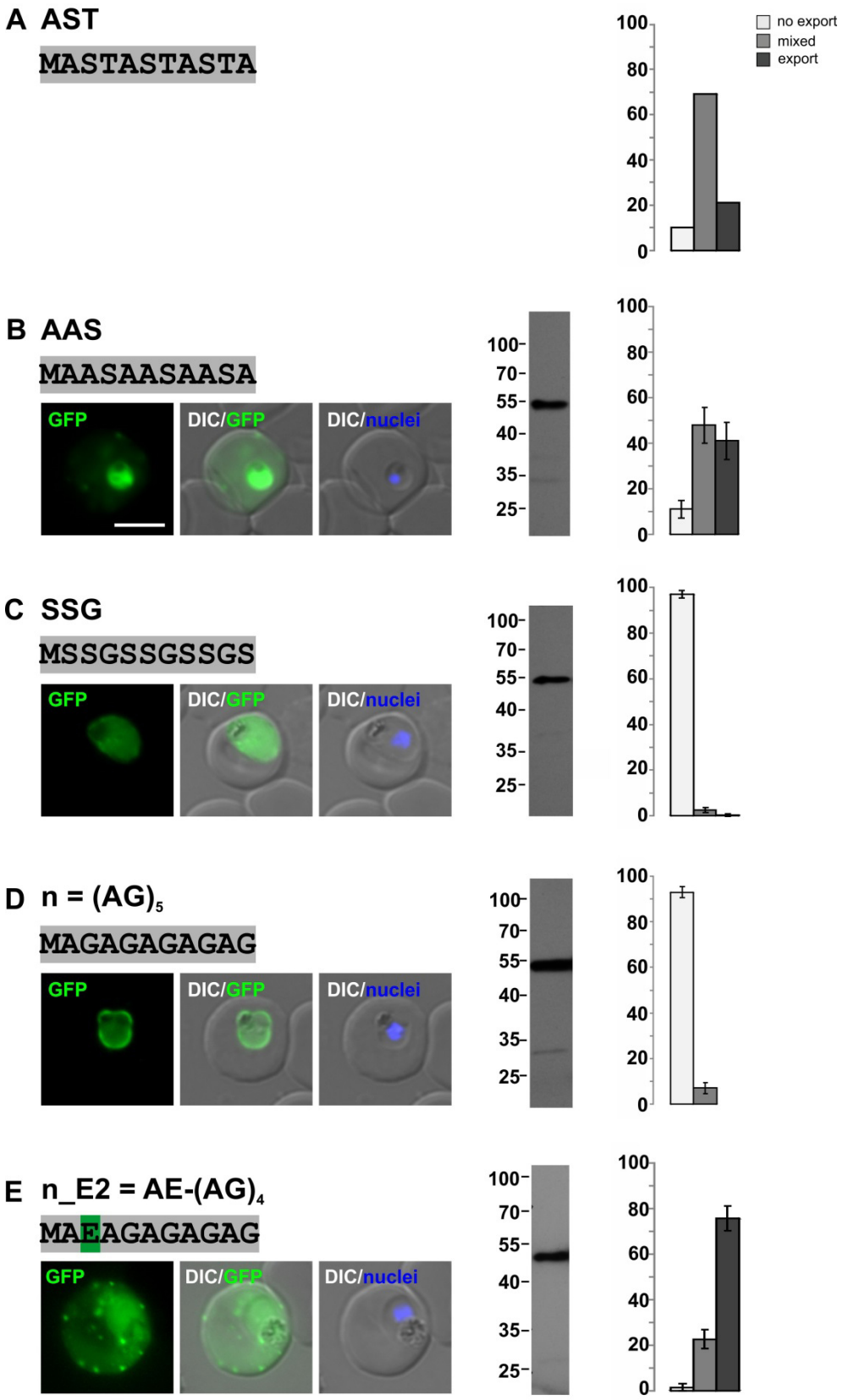


Figure 3.6 – A screen for an N-terminus lacking trafficking signals (A-D) and the validation of the neutral N-terminus by inserting a primary positive signal (E). Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the N-terminus shown above each panel (the neutral N-terminus is highlighted in grey, the primary positive signal in green). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5µm. Western blots to the right of the image panels with extracts of the cell lines corresponding to the fusion construct probed with α-GFP. The size standard is indicated in

kDa. All constructs showed the expected size. The graphs to the very right show the proportion of cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3, except for (A), which was only counted once. M: methionine; A: alanine; S: serine; G: glycine.

To test whether this N-terminus permitted export and did not contain blocking regions that may equally have led to the observed phenotype, a bona fide PEXEL primary signal was inserted, in form of a glutamic acid (E) at the second position after the methionine, replacing a glycine. This construct was referred to as n_E2 (following the nomenclature ignoring the start methionine, see section 3.1). The resulting construct was fully exported (Figure 3.6E). Thus, addition of a primary export signal promoted export, indicating that the AG-N-terminus did not harbour a blocking signal further downstream of the inserted E. It can hence be concluded that the n_E2 N-terminus is truly neutral and that it can be used to test the effect of different amino acids, their position and charge on export.

The AG-repeat N-terminus creates different neighbouring amino acids (A or G) depending on the position an amino acid is inserted. As this might be a confounding factor, a similar N-terminus with a reversed order of As and Gs was designed – GAGAGAGAGA (instead of AGAGAGAGAG). This N-terminus would make possible to distinguish positional from amino acid environmental effects by inserting changes into the respective position of this reverse ordered N-terminus. This GAGAGAGAGA N-terminus did also not promote export of the reporter protein and was found in the PV (Figure 3.7A). However, addition of an E in position 2 of the N-terminus showed that this was due to a blocking capacity of this N-terminus, as the primary signal was incapable to mediate export of this construct (Figure 3.7B). Hence the arrangement of amino acids within this N-terminus seems to prevent export and no such control N-terminus was available for subsequent experiments.

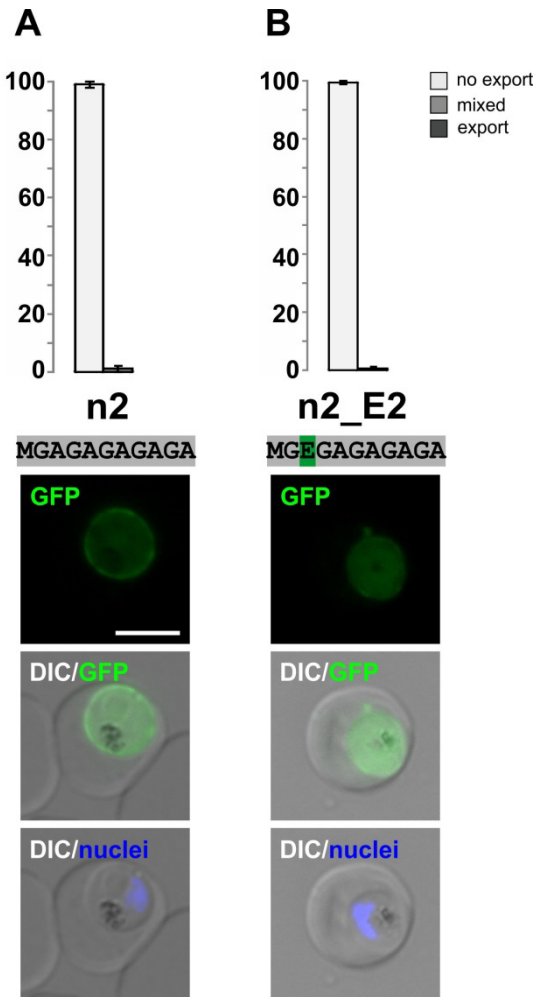


Figure 3.7 – The inverted order of the amino acids of the neutral N-terminus does not promote export, also when a perfect primary positive signal is present. (A) n2, **(B)** n2_E2. Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the respective N-terminus shown above the panels (the neutral N-terminus is highlighted in grey, the replaced amino acid in green). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5µm. The graphs above the panels show the proportion of cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3. M: methionine; A: alanine; G: glycine; E: glutamic acid.

3.3.1 Prolines prevent export of n_E2

Firstly, the capacity of prolines as an export blocker was examined in the neutral N-terminus harbouring a primary positive signal (n_E2; Figure 3.6E). As a triple proline was necessary to prevent SERA7 from being exported (Figure 3.5), it was now tested if this triplet is sufficient to abolish export of the otherwise exported protein n_E2. To test this, three prolines were inserted into the n_E2 N-terminus at position P4, P5 and P6, corresponding to the positions these residues are found in the SERA7 mature N-terminus. In contrast to n_E2, this construct was not exported (Figure 3.8A), indicating that a triple proline can be sufficient as an export blocker. Next triple proline were

inserted further downstream in the N-terminus (P6-8 and 8-10, respectively) to check for their export preventing capacity in dependency to the distance to the very N-terminus. Both of these constructs showed a high export rate (Figure 3.8B, C), comparable to the parent n_E2 (Figure 3.6E). This indicates that a proline triplet can be a highly effective export blocker when located in close proximity to the N-terminus but does not appear to have much effect if placed further downstream.

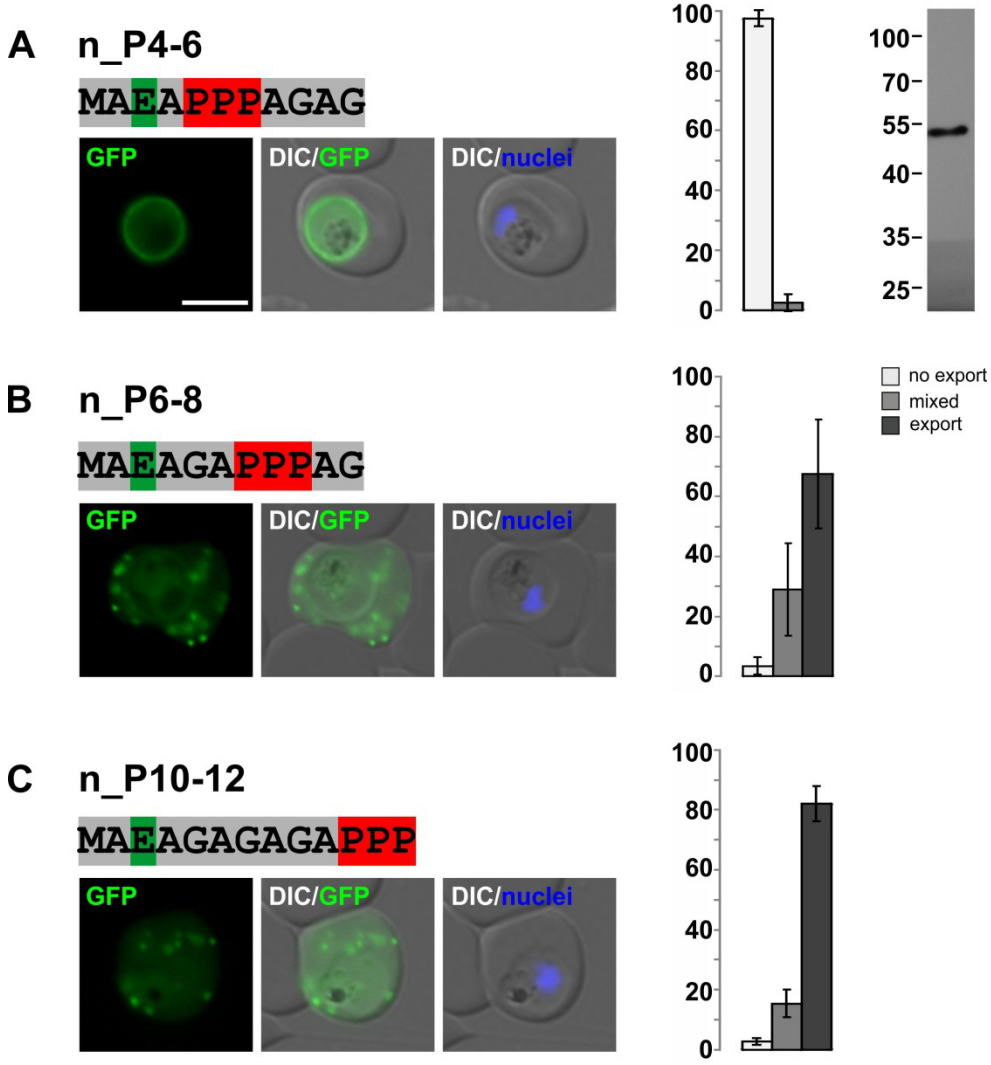


Figure 3.8 – Three prolines in a row can abolish protein export when placed at P4 to P6. (A) n_P4-6, **(B)** n_P6-8, **(C)** n_P10-12. Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the N-terminus shown above each panel (the neutral N-terminus is highlighted in grey, the primary positive signal in green and the blocking region PPP in red). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5µm. The graphs to the right of the image panels show the proportion of cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3. The western blot to the right of the image panel with the extract of the cell line corresponding to the fusion construct probed with α-GFP. The size standard is indicated in kDa. The construct showed the expected size. M: methionine; A: alanine; G: glycine; E: glutamic acid; P: proline.

3.3.2 Plasticity of the primary positive signal

Although defined by the PEXEL motif consensus, the primary positive signal has never been systematically analysed. It was therefore investigated next, if other amino acids than the typical glutamic acid (E) can also serve as a primary positive signal. Besides E, glutamine (Q) and aspartic acid (D) are the two amino acids annotated as the second and third most prominent amino acid residues at position 2 (in the mature N-terminus after plasmepsin V cleavage of PEXEL proteins), respectively (Hiller *et al*., 2004; Marti *et al*., 2004; Boddey *et al*., 2013). These two amino acids were therefore tested first for their export promoting capacity as a primary signal. To this end, they were inserted at position 2 of the neutral N-terminus (n) fused to the truncated mTRAP reporter and expressed in *P. falciparum* parasites. The N-terminus with the Q (n_Q2) showed a high export rate (Figure 3.9A), whereas the construct with D at P2 in the N-terminus (n_D2) showed almost no export (Figure 3.9B). Hence Q in position 2 of the mature N-terminus can act as a true primary positive signal, whereas D cannot.

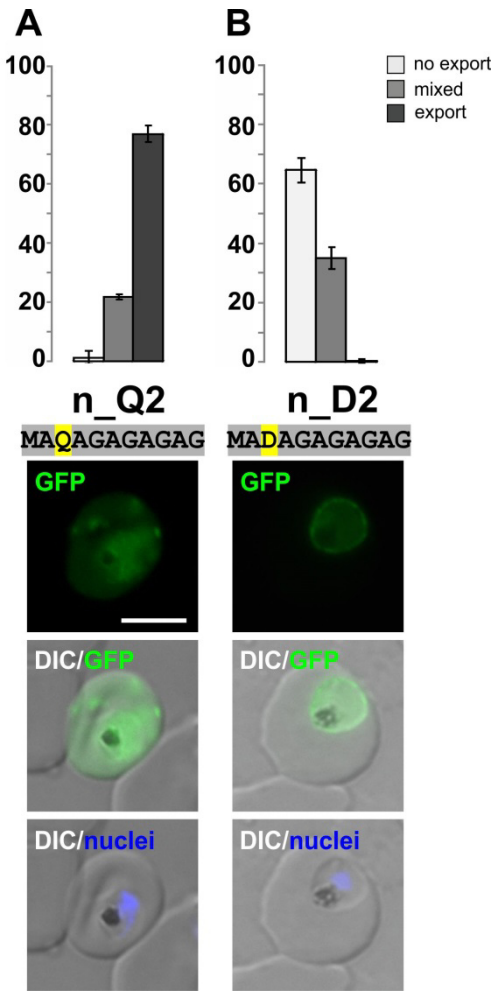


Figure 3.9 – The PEXEL consensus residues present at P2 after PEXEL cleavage (1° positive signal) differ in their export promoting capacity. (A) n_Q2, **(B)** n_D2. Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the respective N-terminus shown above the panels (the neutral N-terminus is highlighted in grey, replaced amino acids in yellow). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5µm. The graphs above the panels show the proportion of cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3. M: methionine; A: alanine; G: glycine; Q: glutamine; D: aspartic acid.

Next amino acids were inserted with different characteristics and properties. Methionine (M) and leucine (L) are nonpolar and show a high helix forming propensity, whereas serine (S) and threonine (T) are polar, but uncharged and show a comparably low helix forming propensity (Creighton, 1992). All the four amino acids promoted export to differing extents when positioned as a potential primary positive signal in P2 of the neutral N-terminus (n) (Figure 3.10). Out of these four, the N-terminus harbouring threonine resulted in the best export rate, comparable to the N-termini with E or Q in the same position (Figure 3.10A). Almost similar levels were achieved using leucine at P2 (Figure 3.10B). M and S resulted in a predominantly mixed export phenotype, although close to 40% of the cells still showed full export (Figure 3.10C, D).

This leads to the conclusion that the primary positive signal is more flexible than anticipated from the current PEXEL consensus and that there is a conflict in the PEXEL consensus as only E and Q but not D at P2 promote export.

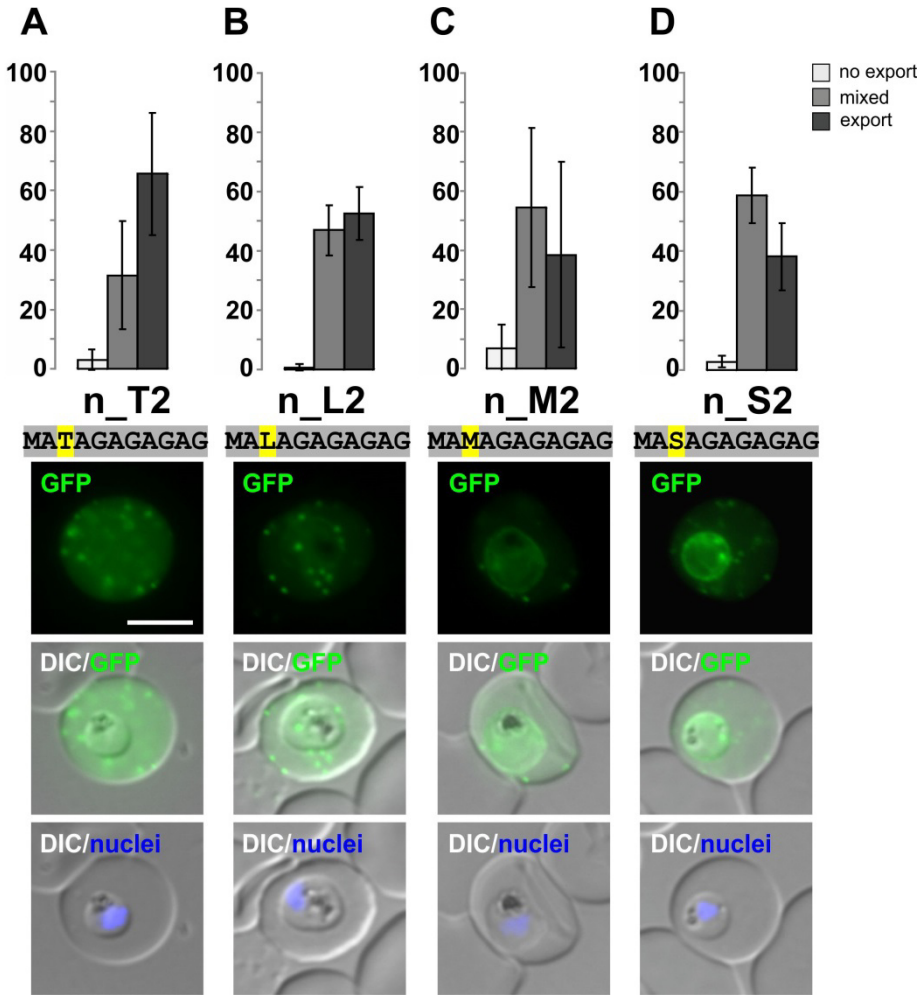


Figure 3.10 – The amino acids T, L, M and S promote export to different levels as a primary export signal (i.e. when inserted at P2) in the neutral N-terminus. (A) n_T2, **(B)** n_L2, **(C)** n_M2, **(D)** n_S2. Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the N-terminus shown above each panel (the neutral N-terminus is highlighted in grey, replaced amino acids in yellow). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5µm. The graphs above each panel show the proportion of cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3. M: methionine; A: alanine; G: glycine; T: threonine; L: leucine; S: serine.

Next, Asparagine (N) – an amino acid with a polar side chain that is very similar to the export promoting Q but smaller and has a lower helix forming propensity – was tested in position 2. This construct (n_N2) showed a predominantly mixed export phenotype (Figure 3.11A) and had a lower export promoting propensity than all the previous tested residues (with the exception of D). An even lower export phenotype was observed when arginine (R), a positively charged amino acid with a comparably low helix forming

propensity, was used (Figure 3.11B). Inserting an alanine resulted in the lowest export rate of all investigated amino acids (Figure 3.11C). Hence, alanine seems to be a truly neutral amino acid at this position in the neutral background. Overall, the results show that more amino acids than expected can promote export as a primary positive signal and hence this position shows great plasticity.

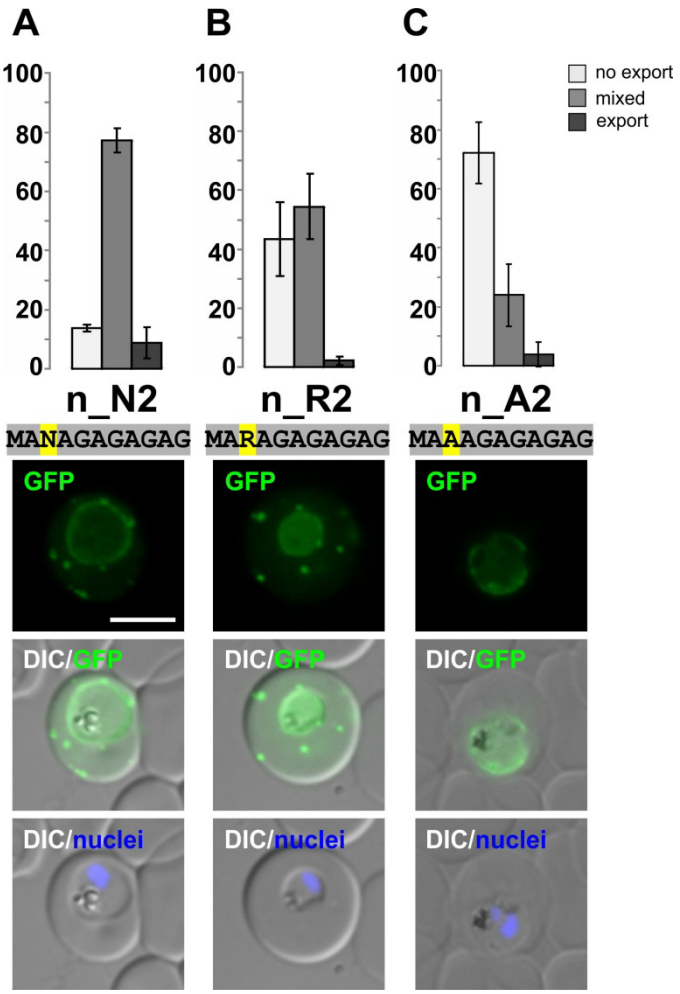


Figure 3.11 – N, R and A as a primary export signal (at P2) do not mediate high export levels. (A) n_N2, **(B)** n_R2, **(C)** n_A2. Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the N-terminus shown above each panel (the neutral N-terminus is highlighted in grey, replaced amino acids in yellow). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5µm. The graphs above each panel show the proportion of cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3. M: methionine; A: alanine; G: glycine; N: asparagine; R: arginine.

3.3.3 Secondary positive signal investigation

3.3.3.1 Influence of the 1° export signal residue E as a 2° export signal

3.3.3.1.1 Single glutamic acid residues are poor secondary export signals

Previous work showed that when the primary positive signal in PEXEL N-termini was mutated, secondary signals further downstream can mediated protein export (Grüring *et al*., 2012; Tarr *et al*., 2013; see section 1.3.5). In a first approach to characterise these secondary export signals the neutral N-terminus (n) was exploited to test whether the residues acting as a primary signal could also act as secondary signal when placed further downstream in the N-terminus. For this analysis E was used, as it is the most frequent amino acid of the corresponding position of the PEXEL motif (P2) and promoted full export as a primary signal (Figure 3.6E). Single Es were inserted at different positions in the neutral terminus, specifically at P3, P4, P5, P6 and P10 (resulting in constructs n_E3, n_E4, n_E5, n_E6, n_E10, respectively). In general, these constructs were not well exported (Figure 3.12B-F) if compared to n_E2, which was included once more for comparison reasons (Figure 3.12A). Generally, the further downstream the glutamic acids were inserted, the lower was the export promoting capacity and consequently the export rate. The only exception was n_E3, where the export rate was close to zero even though E residues in the neighbouring positions (n_E2 and n_E4) led to a high and a moderate export rate, respectively (compare Figure 3.12B with Figure 3.12A and C). Thus, it seems that the flanking amino acids and/or – similar to the triple prolines (Figure 3.8) – the exact relative position to the N-terminus at least in part influence the export rate of the protein. However, overall it can be concluded that a single E, although efficiently promoting export as a primary signal, is a poor secondary signal.

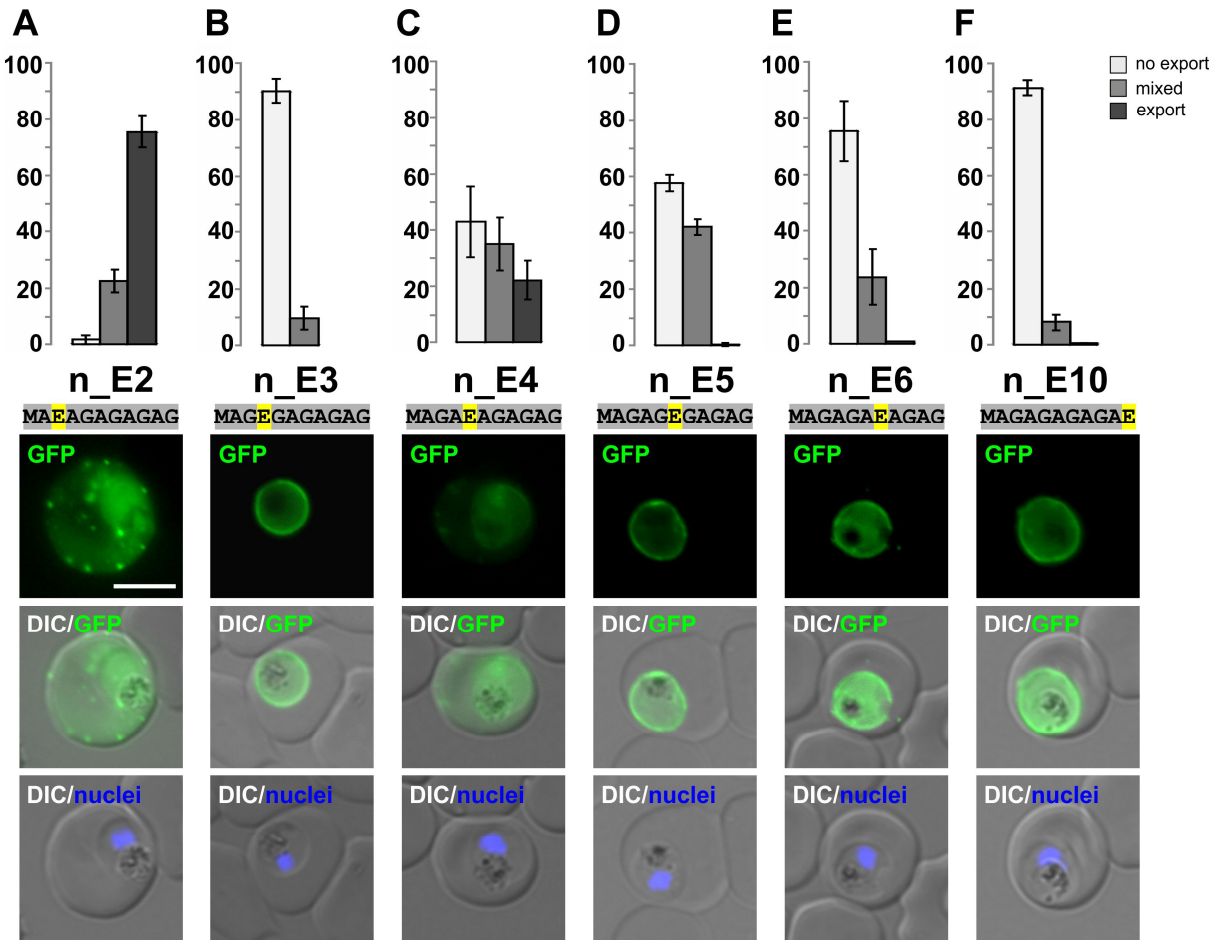


Figure 3.12 – A single E is a poor secondary signal. (A) n_E2, **(B)** n_E3, **(C)** n_E4, **(D)** n_E5, **(E)** n_E6, **(F)** n_E10. Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the N-terminus shown above each panel (the neutral N-terminus is highlighted in grey, replaced amino acids in yellow). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5µm. The graphs above each panel show the proportion of cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3. M: methionine; A: alanine; G: glycine; E: glutamic acid.

Due to that reason, an N-terminus with the amino acid sequence A**AEA**GAGAG was designed (n_E3$^{A\text{-}flank}$), to check if the flanking amino acids influence the protein export. The protein harbouring this N-terminus showed a more mixed phenotype (Figure 3.13) and was consequently much better exported than the n_E3 construct, where the E was flanked by glycines (Figure 3.12B). Hence the flanking amino acids influence the capacity of E residues to promote export in this position close to the primary signal.

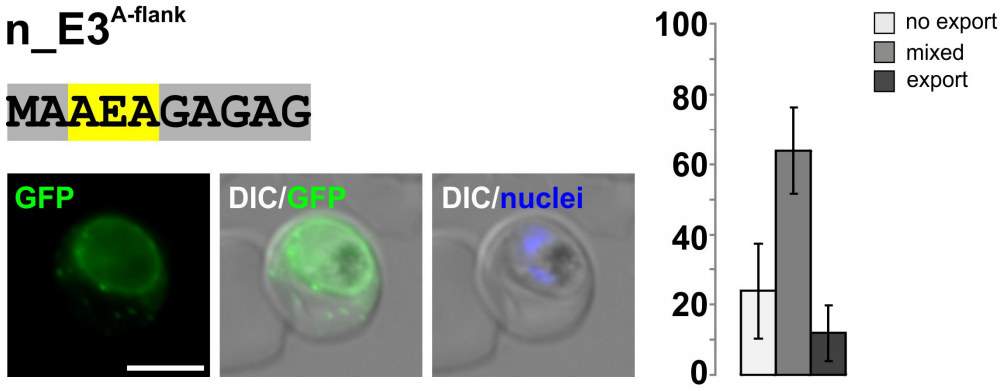


Figure 3.13 – Alanines flanking the glutamic acid in position 3 (n_E3$^{A\text{-}flank}$) partially rescue export. Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the N-terminus shown above the image panel (the neutral N-terminus is highlighted in grey, replaced amino acids in yellow). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5μm. The graph to the right of the image panel shows the proportion of cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the fusion construct. The values are mean ± standard deviation (s.d.), n = 3. M: methionine; A: alanine; G: glycine; E: glutamic acid.

3.3.3.1.2 Capacity of multiple glutamic acid residues to act as a secondary positive export signal

The results so far indicated that insertion of the primary export signal residue E did not well mediate export if placed further downstream in the N-terminus (Figure 3.12). As it however mediated export to some extent and it previously was suggested that negative charge in an N-terminus can positively influence export (Bhattacharjee *et al.*, 2006; Saridaki *et al.*, 2009), it was tested next, if multiple E residues can increase the export rate and serve as a better secondary export signal. This was done either by inserting three glutamic acids (E) at every other residue or consecutively in the neutral N-terminus. First, two constructs alternatingly replacing either the amino acids at positions P3, P5 and P7 or at positions P4, P6 and P8 in the neutral N-terminus, were generated. The first of these constructs (n_E3_E5_E7) showed a mixed to exported phenotype (Figure 3.14A), with more export than any single E replacement (Figure 3.12). In this construct the alternating E residues were flanked by glycines (G). Slightly less export was seen with the construct n_E4_E6_E8 (Figure 3.14B). Here the Es were inserted one position further downstream from the N-terminus and were flanked by alanines (A). These data show that multiple E residues can act as a 2° positive signal. Again, position and flanking region appeared to influence export, although only to a small extent in this case.

Next the impact of consecutive E residues was examined. For this, three glutamic acids were inserted in a row into the neutral N-terminus at positions P3-P5 and P6-P8 (constructs n_E3-5 and n_E6-8, respectively). n_E3-5 showed a low export rate, even though the triple E was located close to the N-terminus (Figure 3.14C). In this case the E residues were flanked by glycines. However, the construct, where the E residues were inserted further downstream in the N-terminus, but were flanked by alanines, showed a much higher export rate (Figure 3.14D). Consequently, multiple Es can serve as a secondary positive signal. And again the flanking amino acids and the position of the secondary positive signal may influence the export levels of the protein.

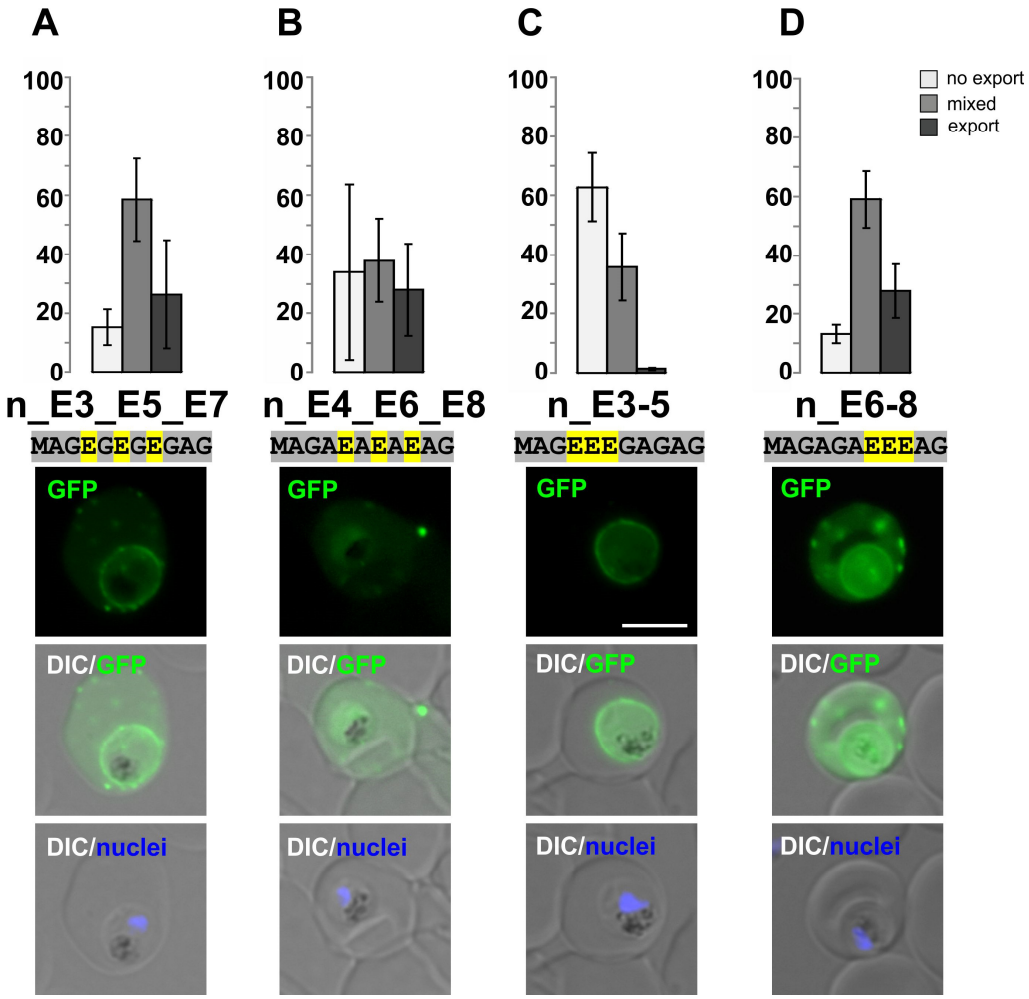


Figure 3.14 – Multiple E residues can act as a secondary export signal. (A) n_E3_E5_E7, **(B)** n_E4_E6_E8, **(C)** n_E3-5, **(D)** n_E6-8. Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the respective N-terminus shown above each panel (the neutral N-terminus is highlighted in grey, replaced amino acids in yellow). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5µm. The graphs above the panels show the proportion of cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3. M: methionine; A: alanine; G: glycine; E: glutamic acid.

3.3.3.2 Capacity of other residues to act as secondary positive export signal

As multiple Es downstream of the location of the primary signal can act as a secondary export signal, it was next investigated whether multiples of other residues can fulfil this function as well. Different residues were chosen based on the previous results: alanine (A) and methionine (M) as residues with comparably high helix forming propensities but without the negative charge of E. In addition, D was also used due to its annotation as the third most prominent amino acid at PEXEL position 5 (Hiller *et al*., 2004; Marti *et al*., 2004; Boddey *et al*., 2013; corresponds to P2 in the neutral N-terminus). It is negatively charged, has a low helix forming propensity and was a poor primary signal. Triplets of these residues were inserted into the neutral N-terminus in position 6 to 8, a position that resulted in export when three Es were used as replacing residues (Figure 3.14D). Analysis of the cell lines expressing the corresponding construct revealed that insertion of alanines led to poor detectability of the fusion protein. As a consequence a quantification of the export phenotype was not possible for this construct (n_A6-8). However, a small number of cells with detectable fluorescence indicated poor export (Figure 3.15A). The aspartic acid triplet (D) did not promote export of the fusion construct (Figure 3.15B). This was in line with the finding that D is unsuitable as a primary positive signal (Figure 3.9B). The construct with the methionine triplet showed a mixed phenotype but overall export was only moderate (Figure 3.15C). Methionine was functional as a primary positive signal with moderate to good efficiency (Figure 3.10C). Thus, amino acids which can serve as a primary positive signal can apparently as triplets also act as a secondary positive signal further downstream, although then the export levels are lower.

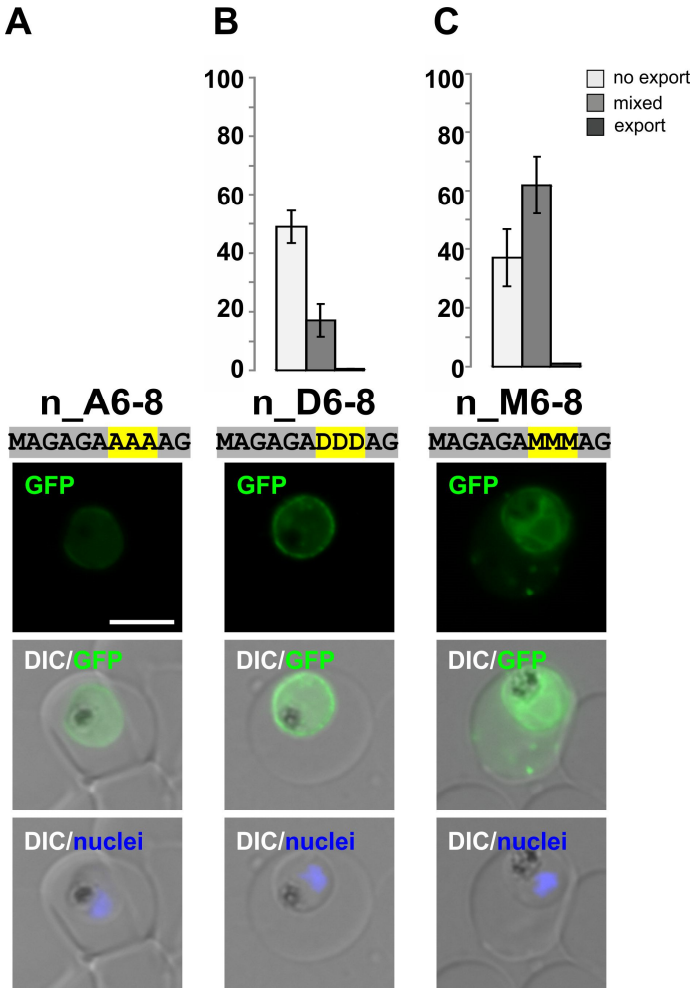


Figure 3.15 – Triplets of alanines, aspartic acids and methionines are poor or moderate secondary export signals. (A) n_A6-8, **(B)** n_D6-8, **(C)** n_M6-8. Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the N-terminus shown above each panel (the neutral N-terminus is highlighted in grey, replaced amino acids in yellow). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5µm. The graphs above the panels show the proportion of cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3. M: methionine; A: alanine; G: glycine; D: aspartic acid.

3.4 Natural secondary export signals

3.4.1 Investigating possible positive signals in STEVOR

To identify the secondary positive signal of natural proteins, the PEXEL protein STEVOR was investigated first. The first 20 amino acids of its mature N-terminus (after cleavage of the PEXEL motif) are sufficient to promote export when fused to the mTRAP reporter. Its 1° positive signal mediates export of the protein but mutation of this signal indicated the presence of a 2° positive signal within the first 20 amino acids of the N-terminus (Grüring *et al.*, 2012). Scrambling the N-terminus of STEVOR abolished export of the

protein (Grüring *et al*., 2012). To narrow down the export promoting region, hybrids of the N-terminus of STEVOR with parts of the scrambled N-terminus were made. For that the first ten amino acids of the N-terminus of wild type (wt) STEVOR were fused to the second ten amino acids of the scrambled N-terminus of STEVOR, resulting in the construct $Ste^{44\text{-}53}/Ste^{54\text{-}63}{}_{scr}$ (Figure 3.16A). The reverse construct containing the first 10 amino acids of the scrambled sequence of STEVOR fused with amino acids 10-20 of the wt N-terminus was also generated ($Ste^{44\text{-}53}{}_{scr}/Ste^{54\text{-}63}$; Figure 3.16B). The $Ste^{44\text{-}53}/Ste^{54\text{-}63}{}_{scr}$ construct still showed full export (Figure 3.16A), whereas no export was observed with $Ste^{44\text{-}53}{}_{scr}/Ste^{54\text{-}63}$ (Figure 3.16B). Inserting a primary positive signal at P2 in the scrambled STEVOR background rescued protein export (Grüring *et al*., 2012) what shows that this N-terminus did not harbour a blocking region. This indicated that no secondary export mediating region lies in amino acids 10-20 of STEVOR and that the 2° positive signal must lie in the first 10 amino acids of the N-terminus, which also contains the primary positive signal. In order to find the 2° export signal in these first 10 amino acids, the primary positive signal was mutated, i.e. the glutamine at position 2 in the mature N-terminus (corresponding to position 45 of the uncleaved N-terminus) was mutated to an alanine ($Ste^{44\text{-}53}Q45A/Ste^{54\text{-}63}{}_{scr}$). The corresponding protein $Ste^{44\text{-}53}Q45A/Ste^{54\text{-}63}{}_{scr}$ was exported (Figure 3.16C), confirming that a secondary positive signal lies between position 3 and 10 of the mature N-terminus. The fourth position in this N-terminus is also a glutamine. To test whether this Q could possibly substitute for the Q at position 2, and hence represent a secondary positive signal, it was next changed to alanine, to obtain a double mutant (Q45A/Q47A). However, the corresponding construct was obtained in the wt N-terminus not the hybrid N-terminus. Nevertheless, this should be representative, as the region 10-20 was already excluded as a possibility for holding the secondary positive region (compare Figure 3.16B and C). The resulting protein, ($Ste^{44\text{-}63}$_Q45A_Q47A) was still exported (Figure 3.16D). Even when amino acids P2, P3 and P4 all were mutated to alanines, the construct was still fully exported (Figure 3.16E). These results lead to the conclusion that position P5 to P10 is necessary for the secondary positive signal in the N-terminus of STEVOR to be active.

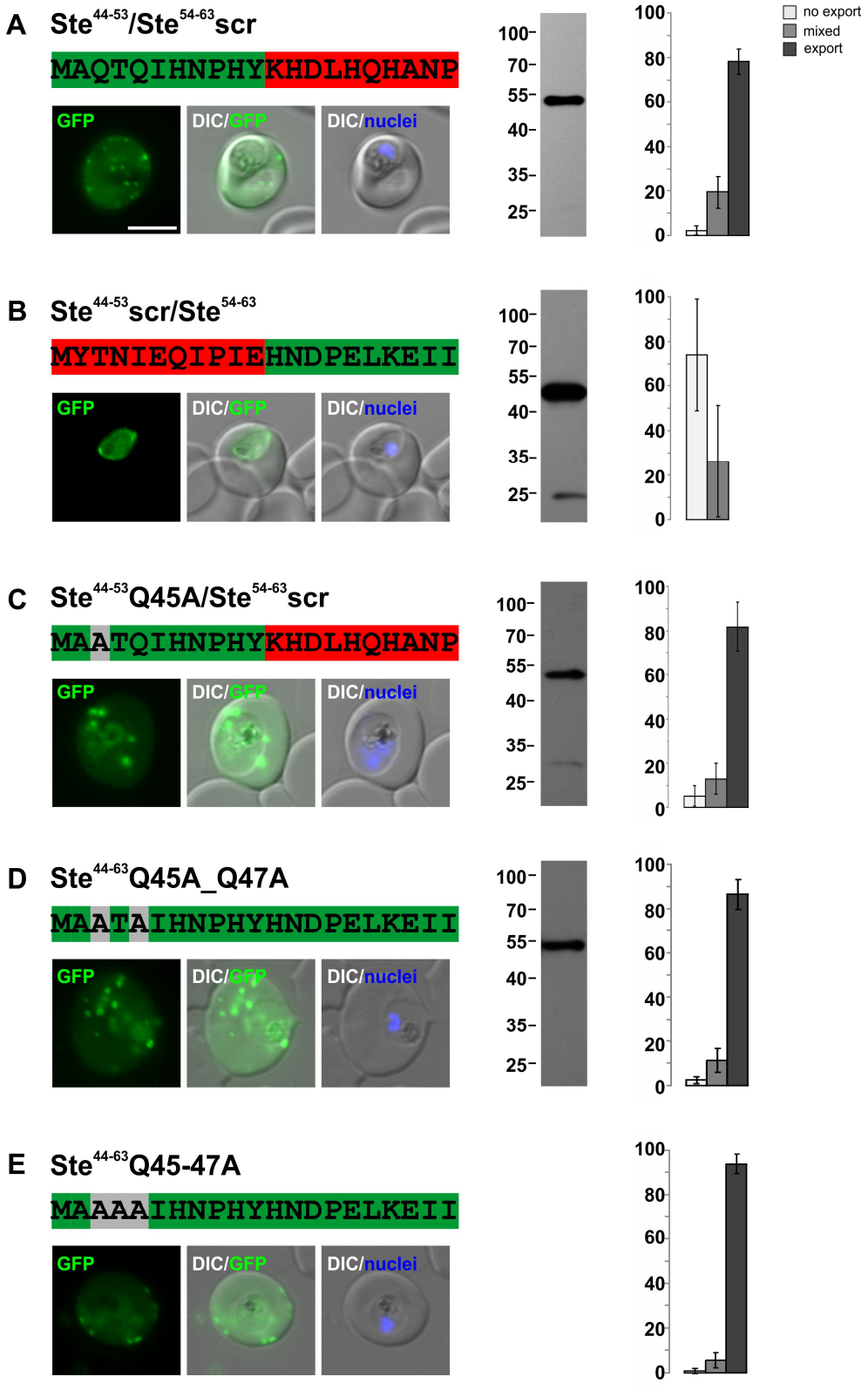


Figure 3.16 – Investigating the secondary positive signal of the N-terminus of STEVOR. (A) $Ste^{44\text{-}53}/Ste^{54\text{-}63}{}_{scr}$, **(B)** $Ste^{44\text{-}53}{}_{scr}/Ste^{54\text{-}63}$, **(C)** $Ste^{44\text{-}53}Q45A/Ste^{54\text{-}63}{}_{scr}$, **(D)** $Ste^{44\text{-}63}Q45A\_Q47A$, **(E)** $Ste^{44\text{-}63}Q45\text{-}47A$. Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the N-terminus shown above each panel (STEVOR sequences are highlighted in green, scrambled sequences in red and mutated amino acids in grey). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5μm. Western blots to the right of the image panels with extracts of the cell lines corresponding to the fusion construct probed with α-GFP. The size standard is indicated in kDa. All constructs showed the expected size. The graphs to the very right show the proportion of cells in % with either no export (light grey bar), partial export (grey

bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3. Amino acids are shown in one letter code.

3.4.1.1 Testing potential natural secondary export signals of STEVOR in the neutral background

As the secondary positive signal of STEVOR requires position 5 to 10 of the mature N-terminus (corresponding to position 48 to 53 in the uncleaved N-terminus of STEVOR; Figure 3.16C-E), it was tested whether this region was also sufficient to act as a 2° positive signal. To do this, parts of the N-terminus of STEVOR were inserted in the neutral N-terminus (n) and investigated for their export promoting capacity. For unknown reasons, P3 to P7 of the STEVOR N-terminus inserted at the exact same position in the neutral N-terminus (agTQIHNgag) showed no detectable GFP signal, which prevented to judge the export promoting capacity of this part of STEVOR (data not shown). Amino acids 5 to 10 of STEVOR, replacing the respective amino acids in the neutral N-terminus, did not promote protein export (Figure 3.17A). The amino acids found at position 9 to 14 in the mature N-terminus of STEVOR did also not mediate export, when inserted in the neutral N-terminus (Figure 3.17B). The amino acids 13 to 20 of the mature STEVOR N-terminus were inserted at position 11 to 18 of the neutral N-terminus (n_Ste13-20). This was done, as the original neutral N-terminus was also only 10 amino acids long and served as the reference. Unexpectedly, n_Ste13-20 led to a mixed export phenotype of the fusion protein (Figure 3.17C). Overall these results do not well agree with the data obtained with the hybrid constructs but indicate that although amino acids 5-10 in the STEVOR N-terminus are necessary for the secondary export signal to be active, they are by themselves not sufficient to exert this function in the neutral N-terminus. These findings point to complicated properties of the secondary export signal in the STEVOR mature N-terminus.

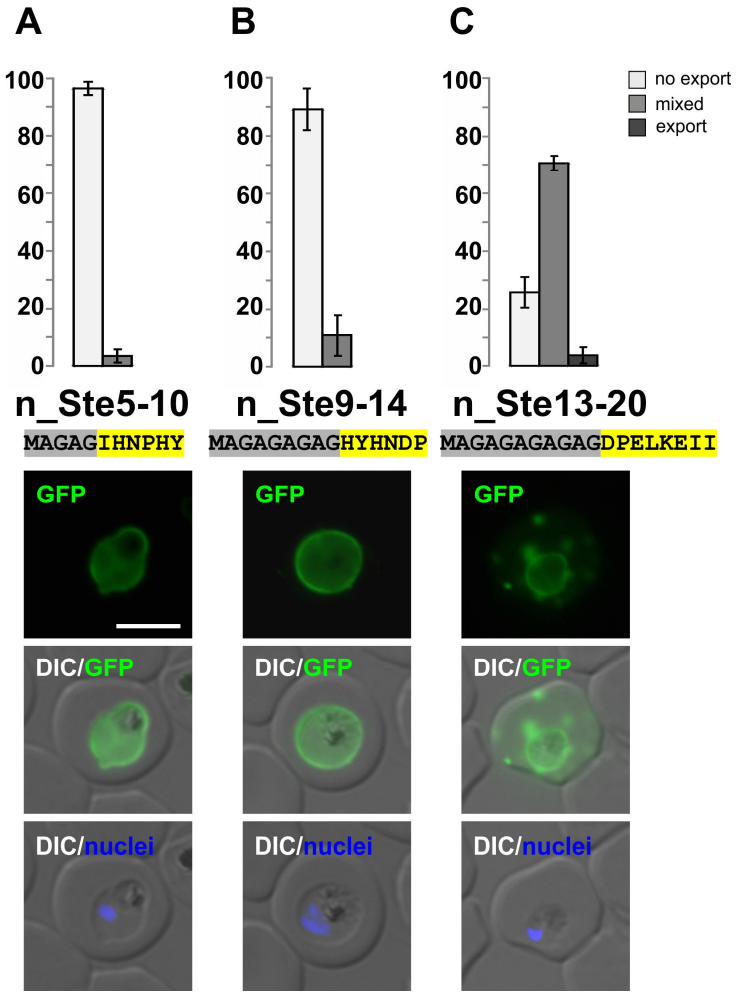


Figure 3.17 – The last part of STEVOR is sufficient to act as a moderate secondary positive signal. (A) n_Ste5-10, **(B)** n_Ste9-14, **(C)** n_Ste13-20. Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the N-terminus shown above each panel (the neutral N-terminus is highlighted in grey, replaced amino acids by parts of the STEVOR N-terminus in yellow). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5µm. The graphs above each panel show the proportion of cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3. Amino acids are shown in one letter code.

3.4.2 Localising secondary positive signals in the mature N-terminus of GBP130

As the data regarding the secondary positive export signal of STEVOR were inconclusive, the mature N-terminus of the PEXEL protein GBP (Boddey *et al.*, 2009) was chosen as a second region to attempt to locate a natural secondary positive signal. It was selected because the first 20 amino acids of the mature N-terminus of this PEXEL protein (corresponding to position 87 to 106 in the unprocessed N-terminus of GBP) promoted full export (MAEGEDTCARKEKTTLRKSKQ; Grüring *et al.*, 2012). And after mutating the

glutamic acid (E) at position P2 to alanine (MA**A**GEDTCARKEKTTLRKSKQ) the protein was still exported, indicating the presence of a secondary export signal downstream of the primary signal (Grüring *et al*., 2012). This mutated N-terminus (GBP130_E88A, termed $GBP130_{mut}$) was used as the basis for investigating the secondary positive signal of the mature GBP130 N-terminus. First, the glutamic acid (E) at P4 (corresponding to position 90 in the uncleaved N-terminus of GBP130) was mutated to see if this residue substituted for the primary signal. This mutation in the GBP_{mut} background almost completely abolished export (Figure 3.18). Consequently this glutamic acid seems to be necessary for the secondary positive signal activity in GBP130. As this construct accumulated in the ER, not the PV, it is however also possible that export unrelated factors caused this phenotype as export maybe would have occurred if the construct could have reached the PV.

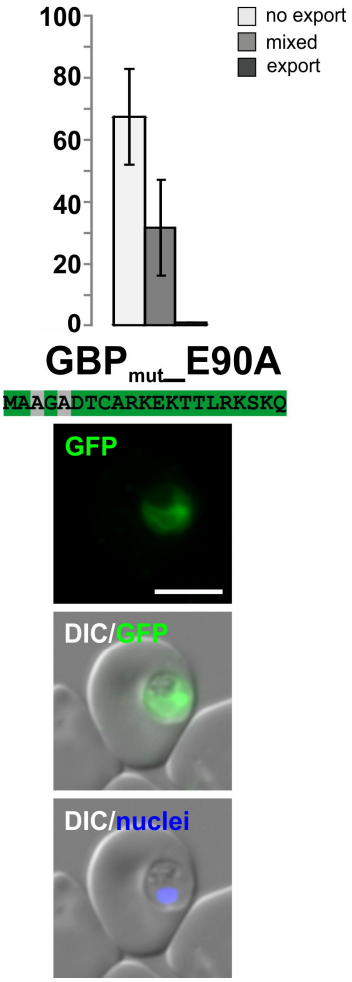


Figure 3.18 – The E at P4 is necessary for the export of $GBP130_{mut}$. Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the N-terminus shown above the panel (the N-terminus of $GBP130_{mut}$ is highlighted in green, replaced amino acids in grey). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5µm. The graph above the panel shows the proportion of

cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3. Amino acids are shown in one letter code.

To test whether the region including the E at P4 was sufficient to promote export, the respective parts of the GBP130 N-terminus were inserted at the corresponding positions in the neutral N-terminus. However, none of the three examined N-termini (region 3-6, 3-7 and 4-7) promoted protein export (Figure 3.19). Hence, the tested parts alone are not sufficient to promote protein export in the neutral N-terminus and are by themselves incapable to act as a secondary positive signal in GBP130. Thus, similar to the situation in STEVOR, the region necessary for export is by itself not sufficient to promote export in the neutral N-terminus. It can therefore be assumed that E at P4 acts as a 2° export signal only in combination with other regions in the N-terminus of GBP130.

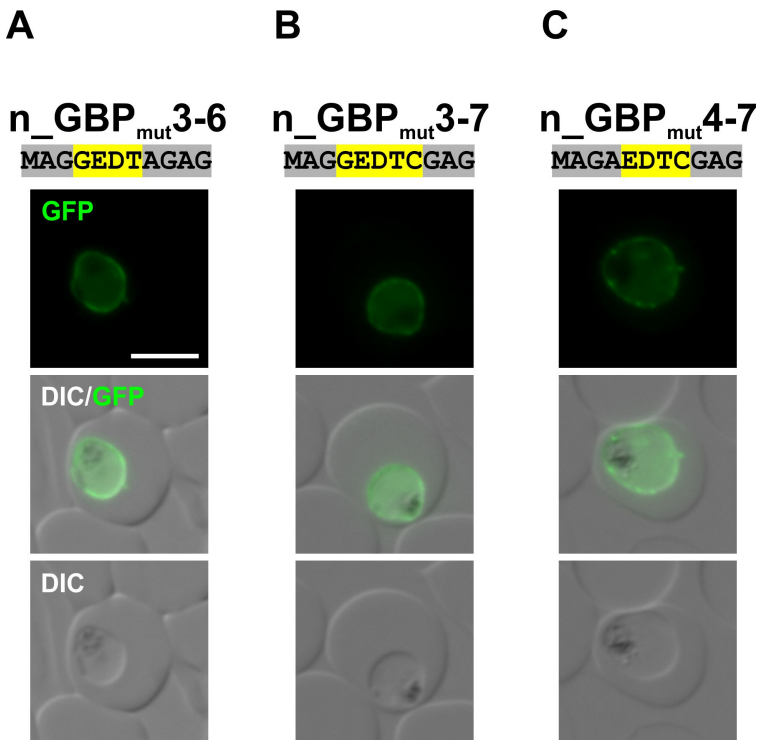


Figure 3.19 – The second E in the GBP mature N-terminus and its flanking amino acids are alone not sufficient to promote protein export. (A) n_GBP130$_{mut}$3-6, **(B)** n_GBP130$_{mut}$3-7, **(C)** n_GBP130$_{mut}$4-7. Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the N-terminus shown above each panel (the neutral N-terminus is highlighted in grey, replaced amino acids by parts of the GBP130 N-terminus in yellow). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. A merge is shown for DIC with GFP. Size bar: 5µm. Amino acids are shown in one letter code.

3.4.3 The PEXEL protein KAHRP and its positive signals

In a third attempt to find a secondary positive signal in a natural context, the first 20 amino acids of the N-terminus of the PEXEL protein KAHRP were investigated. As the second part of this region consists of a histidine (H) stretch and histidine-glutamine (HQ) repeats only, the first part of this region was examined. First, the primary signal was mutated to A (KAHRP_Q2A). As a control the wild type mature N-terminus was used. The control showed a fully exported phenotype (Figure 3.20A) and mutating the primary signal resulted in no considerable reduction of the export rate (Figure 3.20B). This showed that also in the KAHRP N-terminus a secondary positive signal must exist.

Besides the 1° export signal (Q at position P2), the KAHRP N-terminus contains also a second Q at P4 and a glutamic acid (E) at P6. To test whether these residues could serve as a 2° positive signals, they were mutated in KAHRP_Q2A. Mutation of the glutamine in position 4 in addition to the primary signal (P2 and P4, KAHRP_Q2A_Q4A) clearly reduced the protein export and resulted in 30% cells showing no export and a higher rate of cells exhibiting a mixed phenotype (Figure 3.20C). This implies that the glutamine at P4 partially contributes to the 2° export signal. It should be noted that the pool of protein that was not exported in KAHRP_Q2A_Q4A was predominantly found in the ER, which was previously observed for some unexported KAHRP variants (Boddey *et al*., 2009). Mutating also the E in position 6, in addition to P2 and P4 (KAHRP_Q2A_Q4A_E6A), did not further reduce the export efficiency but showed a marginally higher export rate than the double mutant (Figure 3.20D). Hence, a second Q, closely downstream of the Q at P2 serving as the 1° export signal, contributes to the 2° positive export signal in this N-terminus, similar to the situation in GBP.

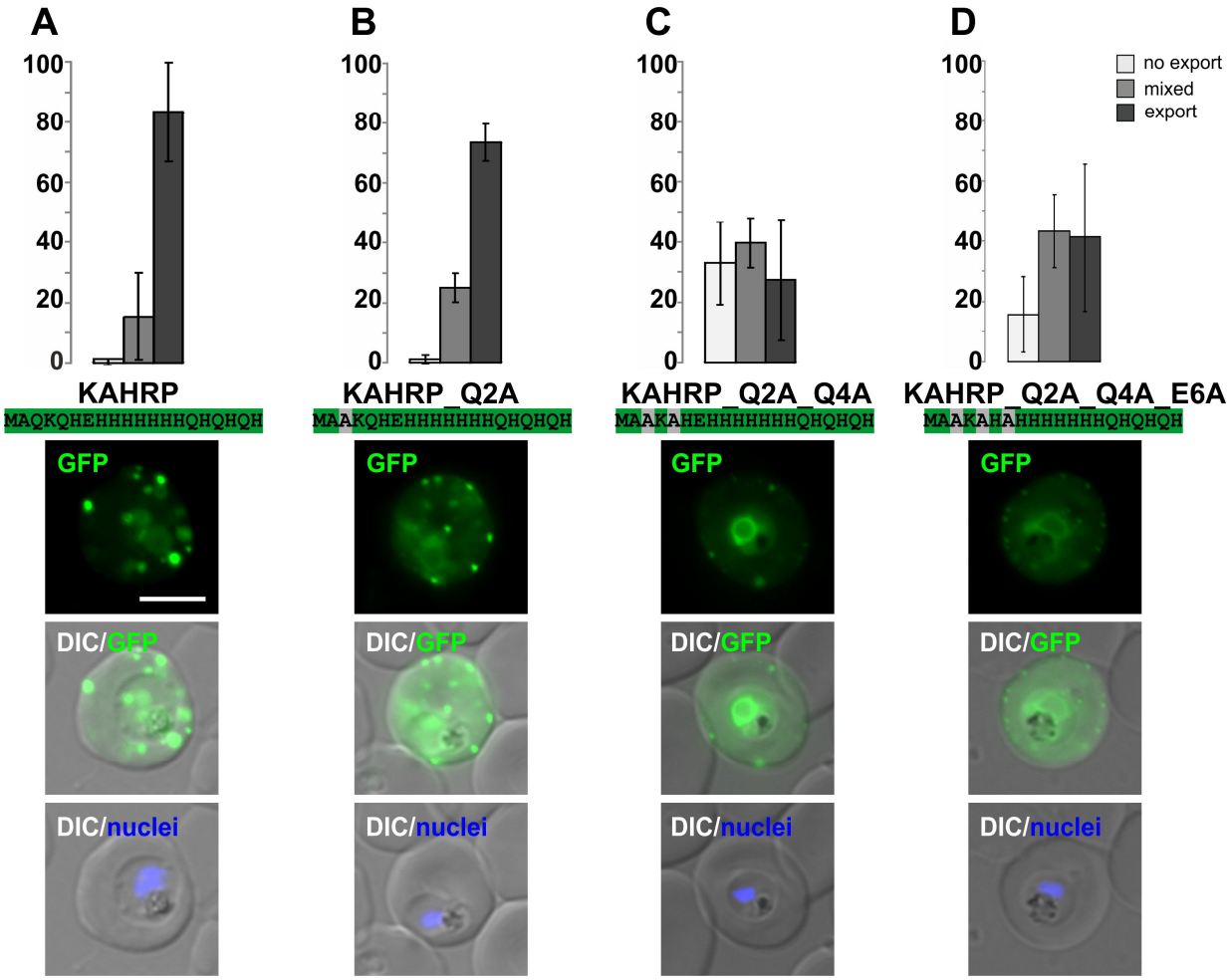


Figure 3.20 – The glutamine at position 4 of the KAHRP mature N-terminus is partially necessary for the 2° export signal. (A) KAHRP, **(B)** KAHRP_Q2A, **(C)** KAHRP_Q2A_Q4A, **(D)** KAHRP_Q2A_Q4A_E6A. Representative fluorescence images of live *P. falciparum* parasites expressing the mTRAP reporter with the N-terminus shown above each panel (KAHRP sequences are highlighted in green, mutated amino acids in grey). DIC, differential interference contrast; GFP, green fluorescent protein (green); DAPI, parasite nuclei staining (blue). The GFP signal shows the fusion construct. Merges are shown for DIC with GFP and DIC with DAPI. Size bar: 5µm. The graphs above each panel show the proportion of cells in % with either no export (light grey bar), partial export (grey bar) or full export (dark grey bar) of the respective fusion construct. The values are mean ± standard deviation (s.d.), n = 3. Amino acids are shown in one letter code.

As none of these investigated N-termini (STEVOR, GBP130, KAHRP) resulted in the identification of a simple sequence that was necessary and sufficient to act as a 2° positive export signal, the uncovering of such regions in neutral N-termini seems to be challenging and indicates complicated properties of the secondary export signal.

4 Discussion

Plasmodium falciparum causes the most severe form of human malaria. To survive within its host red blood cell, the parasite needs to remodel this host cell extensively. This is done by exporting proteins across the two parasite membranes – the parasite plasma membrane (PPM) and the parasitophorous vacuole membrane (PVM) (Spillman *et al*., 2015). The Maurer's clefts (MC's) are a major structure involved in protein export of the parasite into its host cell (Lanzer *et al*., 2006). These parasite-derived membranous structures transfer proteins from the parasite to their destination in the host cell, mainly the erythrocyte membrane (Mundwiler-Pachlatko & Beck, 2013). One such protein is the major virulence factor PfEMP1, which transiently localises to the Maurer's clefts until it is anchored into the so-called knobs. These knobs are located underneath the surface of the erythrocyte membrane, with PfEMP1 displayed on the surface (Kriek *et al*., 2003). Up to now roughly 450 proteins have been identified belonging to the *Plasmodium falciparum* exportome (Boddey *et al*., 2013). The majority of these proteins contains a conserved motif (RxLxE/Q/D), termed the *Plasmodium* export element (PEXEL; Marti *et al*., 2004; or HT signal, Hiller *et al*., 2004). This motif is found approximately 20 amino acids downstream of a signal peptide (Marti *et al*., 2004; Hiller *et al*., 2004; Sargeant *et al*., 2006; van Ooij *et al*., 2008). After cleavage in the endoplasmic reticulum (ER) by plasmepsin V, the mature N-terminus xE/Q/D is released and sufficient to promote export of the respective protein (Boddey *et al*., 2009). A few exported proteins do not harbour such a motif and are therefore referred to as PEXEL negative exported proteins (PNEPs; Spielmann & Gilberger, 2010). Most of these PNEPs show a transmembrane domain and no signal peptide, but different constellations are possible (Heiber *et al*., 2013). The N-termini of PNEPs without a signal peptide were found to be a common region mediating export of these proteins (Grüring *et al*., 2012). This region was found to be exchangeable in its capacity to promote export with the mature N-terminus of PEXEL proteins. This indicated that PEXEL proteins and PNEPs likely share a common core export domain (Grüring *et al*., 2012). The properties guiding export in this common core export domain are so far only poorly defined, but previous data indicated that a combination of positive and negative signals within the N-termini of the respective proteins decided on the export promoting capacity of the particular N-terminus (Marti & Spielmann, 2013).

The aim of this study was to characterise the positive and negative export signals to ultimately obtain rules for the prediction of further (or all) exported proteins in malaria parasites and to obtain the means to assess where in the export pathway the N-terminal export domain acts.

4.1 Primary and secondary positive signal

To study the sequence properties affecting the export promoting capacity of the N-terminal export domain of exported proteins, a well-known reporter was used. In previous studies it was shown that a truncated version of the micronemal protein mTRAP (Baum *et al*., 2006) was an ideal reporter to assess the capacity of an N-terminus to promote export (Grüring *et al*., 2012). To do this, the N-terminus to be tested was N-terminally appended to the reporter, placing this region into the position it is found in the mature protein. The reporter itself was found to be neutral, but export also depended on the presence of a PNEP transmembrane domain (REX2TM) that however by itself did not promote export (Grüring *et al*., 2012). Most crucially, the findings with this reporter containing a mature PEXEL N-terminus were reconciled when the N-terminus tested was placed back into a PEXEL context (Grüring *et al*., 2012), suggesting that even in the absence of Plasmepsin V cleavage, this reporter is a reliable system to test the export promoting capacity of mature N-termini.

However, while the mTRAP reporter can be considered neutral, the current work clearly demonstrates that when looking for the effect of individual sequences affecting export, all natural N-termini appended to this reporter were too complex to obtain a simple result. It was therefore also instrumental to design a neutral N-terminus to assess the effect of small sequence changes on export. In fact the data with the neutral N-terminus may be the most clear cut from this thesis. Such a neutral N-terminus should not hold any export promoting or blocking regions and therefore should not be exported. The most obvious choice to serve as a neutral N-terminus, ten alanines in a row appended to the mTRAP reporter, did not result in a detectable protein, likely due to degradation of this construct. Although alanine in theory should lead to a comparably long half-life of its respective protein compared to other amino acids (Bachmair *et al*., 1986), this did not seem to be the case in *P. falciparum*. One possibility is that oligoA leads to a fold that is rapidly recognised for degradation, leading to degradation of the fusion protein in the

respective parasite line. After further tests it was found that an N-terminus containing five alanine-glycine-repeats was appropriate. It did not hold any export stimulating or prohibiting information – which was validated by inserting a perfect 1° export signal into it – and is therefore truly neutral. Surprisingly, an N-terminus with a reversed order of the A and Gs (GA), contained export blocking properties as even with the perfect primary signal it did not lead to export of this construct. This finding already hinted at complex properties that can affect the effectiveness of the different signals in the export promoting N-termini, here in particular in regards to the position of G residues.

Using the neutral N-terminus, it was shown that the primary positive signal is more flexible than expected. Glutamic acid (E) and glutamine (Q) were the amino acids annotated as the first and second most prominent residues at position P2 after plasmepsin V cleavage of the PEXEL motif, respectively (Hiller *et al*., 2004; Marti *et al*., 2004; Sargeant *et al*., 2006; Boddey *et al*., 2013). These amino acids indeed promoted export when located at this position. Glutamic acid is negatively charged and it was previously postulated that negative charges might influence export (Saridaki *et al*., 2009) although this was not the case in REX2 (Haase *et al*., 2009) and would have had to be position-specific, as scrambling of N-termini (which does not change the overall charge; Grüring *et al*., 2012) did abolish export. The negatively charged aspartic acid (D), the amino acid predicted to be the third most frequently found at P2 after cleavage, did not result in export when tested as a 1° export signal. Hence, negative charge alone is not the deciding factor for the 1° export signal. This is also supported by the effectiveness of Q, which is also not negatively charged but related to E. Instead, other amino acids, not found in the PEXEL consensus position 5 and not considered to promote export, caused export albeit most with somewhat lower export rates than E or Q at P2. Threonine (T), leucine (L), methionine (M) and serine (S) mediated high (T and L) to mediate export rates (M and S; Table 4.1). As these amino acids are not negatively charged but uncharged, the hypothesis that charge is the decisive factor in protein export was not confirmed. However, in the case of PEXEL position 5 the independence of charge was already expected, as Q is not a negatively charged residue. But much less expected was the finding that other residues can promote export in this position what may be possibly related to the helix-forming propensity of the respective amino acid (Table 4.1 and Table 4.3).

The range of amino acids acting as a possible 1° positive signal is a very relevant issue, as this could indicate that the current PEXEL consensus may not be accurate. It is already known that the first part the PEXEL motif (RxL) is necessary for plasmepsin V cleavage (Boddey *et al*., 2009; Russo *et al*., 2010), whereas the second part (xE/Q/D) is necessary for exporting the protein (Boddey *et al*., 2009, 2010, 2013; Russo *et al*., 2010). There are also indications that the first part required for cleavage is rather invariable (Boddey *et al*., 2013; Sleebs *et al*., 2014) but that the fifth position is more tolerant (Boddey *et al*., 2009). However, this tolerance in the fifth position, based on its mutation to A, may also derive from the fact that the 2° positive signal, now well documented (Grüring *et al*., 2012; Tarr *et al*., 2013), may has taken over. The findings of this thesis however clearly indicate that the fifth position of the PEXEL motif needs to be revisited. Mal7P1.170 (PF3D7_0730800.1) for instance is an exported protein classified as a PNEP with the PEXEL like sequence RxLxS ~20 amino acids downstream of the signal peptide. As the 5th position of the PEXEL motif now seems to be variable, this protein may possibly need to be reclassified as a PNEP (Alexandra Blancke Soares, Dissertation 2016). However, processing of this PEXEL-like sequence first need to be experimentally demonstrated before this is warranted.

The next step should now be to place the unexpected residues found to act as the 1° export signal into a bona fide PEXEL protein and then assess export. It should be noted that most of the unexpectedly active residues still showed less export than E or Q. It may therefore be that although partially active, these amino acids may not be found in natural proteins, as they promote less efficient export. Further, proteins with an imperfect PEXEL motif could also be only partially exported, as they maybe fulfil more than one function. The phenomenon of proteins exhibiting more than one function is called 'protein moonlighting' or 'gene sharing' (Jeffery, 2003) and is to date not well characterised in *P. falciparum*. It is however also not clear why there is such a strong selection on this position in PEXEL proteins if its mutation has no effect due to 2° positive export signals further downstream. It is also important to note that the PEXEL consensus derives from a motif identified in a smaller group of exported proteins (Marti *et al*., 2004; Hiller *et al*., 2004) and no systematic experimental analysis was carried out for position 5. This work clearly highlights that such an analysis is overdue and provides first leads that however still need to be tested in a PEXEL context. An appropriately adapted PEXEL consensus based on this analysis could result in a more comprehensive prediction of PEXEL proteins and may lead to a more complete exportome.

It should however also be noted that a strong overrepresentation of certain residues (more frequent occurrence than expected from frequency this residue is normally found in *P. falciparum* proteins) in positon 5 should have been detected previously. High frequency PEXEL motifs with differing position 5 residues are therefore not expected. It may even be that these residues are only tolerated in certain contexts, as was previously shown for position 1 of the PEXEL motif. Here initial data indicated that only R but not K (as originally suspected) is tolerated (Boddey *et al*., 2013; Sleebs *et al*., 2014) but later on it turned out that depending on the particular motif, K can be equally functional (Schulze *et al*., 2015).

Maybe the most puzzling finding was the fact that D did not promote export as a 1° export signal. However, this residue was included in the PEXEL motif mostly due to its presence of the PEXEL like sequence in PfEMP1 (Sargeant *et al*., 2006), a motif now known not to be cleaved by Plasmepsin V (Boddey *et al*., 2013). Interestingly, D at the 5th position of the PEXEL motif is overrepresented in *P. falciparum* (Schulze *et al*., 2015), indicating that some of these motifs may be functional which makes the data obtained in this thesis even more striking. Nevertheless, this needs to be experimentally verified in a PEXEL context but it may be that D can be removed from the consensus, is present in PEXEL motifs of a particular function that may even be unrelated to export, or is found in only a subset of PEXEL motifs with an appropriate surrounding where this amino acid is tolerated.

Amino acid	export promoting capacity						α-helix or β-sheet
	as 1° signal			as 2° signal			
	export	mixed	no ex.	export	mixed	no ex.	
Glu (E)	x			x	x		✓
Gln (Q)	x						✓
Thr (T)	x						✓
Leu (L)	x						✓
Met (M)		x			x		✓
Ser (S)		x					✗
Asn (N)		x	x				✗
Arg (R)		x	x				✗
Asp (D)			x			x	✗
Ala (A)			x			x	✓
Gly (G) → neutral			x				✗

Table 4.1 – List of amino acids according to their export promoting capacity. The amino acids are arranged regarding to their export promoting capacity as a 1° or 2° positive signal (middle) and their secondary structure forming propensity. Dark grey: High export promoting capacity/high secondary structure (α-helix or β-sheet) forming propensity; grey: medium export promoting capacity; light grey: low export promoting capacity/low secondary structure forming propensity. Amino acids are shown in one and three letter code.

The secondary positive signal also appeared to be variable which might have been one reason it could not be sufficiently defined to arrive at a consensus. This is however less surprising, as the sequence variance downstream of the 1° export signal in mature PEXEL N-termini and PNEP N-termini in general is considerable, without a simple consensus between these regions (Table 4.2; Hiss *et al.*, 2008; Haase *et al.*, 2009; Saridaki *et al.*, 2009; Bhattacharjee *et al.*, 2009; Grüring *et al.*, 2012; Tarr *et al.*, 2013). As shown in Table 4.2 the N-termini causing export through a 2° positive signal are very diverse regarding their amino acid arrangement. Interestingly it appeared that positive signals further downstream in the N-terminus can act as an export signal and therefore take over the export promoting capacity in place of the primary positive signal (Table 4.1).

Based on the experimental findings, two signal types acting as 2° positive signals were observed – firstly the 1° export signal type residues alone close to the 1° signal itself (e.g. GBP). This type was found based on mutations in natural N-termini. However, these

signals did not seem to be sufficient to drive export by themselves. The other type were multiples of 1° export type residues further downstream. However, the efficiency was lower and such sequences were not found in naturally occurring N-termini. Strong positional and neighbouring effects may confound this and indicate that the consensus regarding positive signals within N-termini of exported proteins is more complex than this. It should also be noted that these 1° export signal residues placed further downstream are quite artificial and may be very imperfect if compared to the 2° export signals found in natural proteins. In these proteins the secondary positive signal appears to be diffuse – a region where the amino acid arrangement or the overall charge is decisive for the export phenotype instead of a special amino acid alone. As the bioinformatics analysis (see Appendix) as well as the experiments did not uncover a clear sequence promoting export, apart from the already known 1° positive signal, the 2° positive signal seems to be complex. Thus even though a common core export domain for PEXEL proteins and PNEPs is likely (Grüring *et al*., 2012), an exact categorisation or localisation of a secondary positive signal and therefore an overall rule could not be deduced here. This is maybe due to the fact that a clear-cut secondary positive signal (like the primary positive signal) does not exist (Table 4.2). Instead of that several different secondary positive signals may exist, but differ remarkably in their location, length, composition and folding. Assuming also that multiples of several residues can act as 2° positive signal, they will likely not occur the way it was tested but may be mixtures of these residues (i.e. instead of triple E it may be ETQ or, as they do not need to be consecutive, e.g. in the ExExE situation, such a signal may be difficult to spot). However, if the export promoting capacity for all residues would be known, it might be possible to calculate the probability of a given N-terminus to be exporting or not.

Exported proteins through 2° positive signal	
name	**amino acid sequence**
MAHRP1$^{1\text{-}20}$ *1/(*3) ^	AEQAAVQPESVPTVGTVPQ
MAHRP2$^{1\text{-}20}$ *1	QPCPYDVYNQINHVGHWA
SBP1$^{1\text{-}26}$ *1	CSAARAFDFFTDLADEPTQLQDAP
REX1$^{1\text{-}38}$ *1 ^	ADYSSNEEETPKEEKKISKL EDMQSPFDYKRFFRKYT
REX2$^{1\text{-}20}$ *1/*4	KMYLAEIFSSGKESLLSLK
GBP130$^{87\text{-}106}$ E88A *1	AAGEDTCARKEKTTLRKSKQ
STEVOR $^{44\text{-}63}$Q45A *1	AATQIHNPHYHNDPELKEII
PfEMP3$_{63\text{-}82}$_Q2A *2	AAVLGNTRLSSRGVRDPRTK
PFI1755c$_{1\text{-}61}$:GFP_EP2A *2	SAPVVEEQDLKKT
PFI1755c$_{1\text{-}61}$_mutant2 *2	SAAAAEEQDLKKT
PFI1755c$_{1\text{-}61}$_mutant3 *2	SAPVVEEQDAAAA
KAHRP$_{1\text{-}69}$	AQKQHEHHHHHHHQHQHQH
KAHRP_Q2A	AAKQHEHHHHHHHQHQHQH
KAHRP_Q2A_Q4A	AAKAHEHHHHHHHQHQHQH
KAHRP_Q2A_Q4A_E6A	AAKAHAHHHHHHHQHQHQH
Ste$^{44\text{-}53}$Q45A/Ste$^{54\text{-}63}_{scr}$	AATQIHNPHYKHDLHQHANP
Ste$^{44\text{-}63}$Q45A_Q47A	AATAIHNPHYHNDPELKEII
Ste$^{44\text{-}63}$45-47A	AAAAIHNPHYHNDPELKEII
n_Ste13-20	AGAGAGAGAGDPELKEII (mixed)
REX2$^{1\text{-}10}$mTRAP/SERA7$^{25\text{-}32}$_1	KMYLAEIFSAAAADSTV (mixed)
REX2$^{1\text{-}10}$mTRAP/SERA7$^{25\text{-}32}$_2	KMYLAEIFSKPAAAATV (mixed)
REX2$^{1\text{-}10}$mTRAP/SERA7$^{25\text{-}32}$_c	KMYLAEIFSKPPPDSTV (mixed)
AAS	AASAASAASA (mixed)
AST	ASTASTASTA (mixed)
n_E3$^{A\text{-}flank}$	AAEAGAGAG (mixed)
n_E3_E5_E7	AGEGEGEGAG (mixed)
n_E4_E6_E8	AGAEAEAEAG (mixed)
n_E6-8	AGAGAEEEAG (mixed)
n_M6-8	AGAGAMMMAG (mixed)

Table 4.2 – N-termini promoting export due to a 2° positive signal. The upper part of the table (above bold line) shows already known N-termini and below N-termini analysed or generated in this work are shown. On the left, the name of the protein/construct is noted; on the right the amino acid sequence of the respective N-terminus is shown. The N-termini mediated full export except for those labelled 'mixed' in brackets where export was mixed. Amino acids are shown in one letter code. Color code for amino acids: green, polar, including histidine; blue, acidic; red, hydrophobic; magenta, basic. *1: Grüring *et al.*, 2012; *2: Tarr *et al.*, 2013; *3: Saridaki *et al.*, 2009; *4: Haase *et al.*, 2009. ^: PNEPs resembling a 1° positive signal.

Regarding the PEXEL protein STEVOR, the very last part of its N-terminus (P13 to P20 after the PEXEL cleavage site) seems to hold a secondary positive signal, which can cause export of the protein. This would explain why, in contrast to e.g. GBP, mutating the positive signals (in terms of glutamines at P2 and P4 after the PEXEL cleavage site) to alanines did not abolish the export of the protein, as the secondary positive signal at the end of the N-terminus remained undisturbed. But as these two glutamines (at P2 and

P4) are clearly positive signals (although there is no evidence based on the investigated constructs that these Qs would be sufficient to promote export in this case), it seems likely that more than one secondary positive signal can be located within one N-terminus (Figure 4.1). When mutating the mentioned Qs to glycines instead of alanines (SteQ45G/Q47G), the export was however fully abolished (data not shown). This could lead to the conclusion that the alanines promoted the export instead of neutralising the signals. But as this is very unlikely due to comparisons with several other N-termini where alanine indeed neutralised positive signals similar to these two, the Gs probably acted as an export blocker instead. Hence, it is possible that in exported proteins export signals, multiple positive signals or even the entire sequence need to fulfil certain properties to be compatible with export.

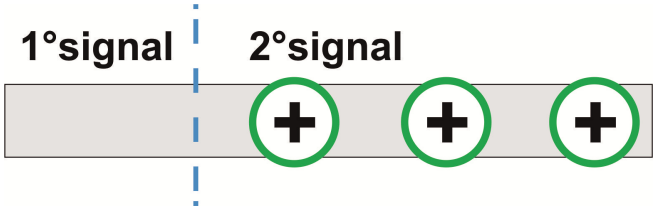


Figure 4.1 – Schematic of an N-terminus (grey bar) harbouring several secondary positive signals (green positive signs). Blue dashed line: Boundary separates primary and secondary positive signal.

4.2 Localisation of export signals

The effectivity of signals affecting export seems to be position-dependent. This is likely due to the fact that the proline (P) triplet blocked protein export only in the beginning of the N-terminus, but not when located further downstream. This assumption is further supported by the export promoting capacity of glutamic acid, which was lower the further downstream this amino acid was inserted in the N-terminus. Hence it is likely that the export promoting or preventing capacity of a signal is dependent of its location – the further downstream in the N-terminus, the lower is the impact of the signal. As Haase *et al*. postulated the first ten amino acids of an N-terminus seem to be sufficient to decide about the export of the protein (Haase *et al*., 2009). At least this is the case for certain signals, independent of whether they are export promoting or preventing in nature. However, the parts of some N-termini do not promote or block export when divided into 10 amino acids each, although the full N-terminus consisting of 20 amino acids promoted or prevented the export of the respective protein. This might be the case

for the combination of the soluble PV protein SERA7 with the PNEP REX2, where the intersection of the first 10 amino acids of REX2 and the appended region from SERA7 generated a blocking region at the interface of the two regions that both sequences alone did not contain. It is also possible that there was some cross talk between the regions – for instance if the SERA7 region negatively affected the folding of the REX2 first 10 amino acids and this normally was essential for export. Hence, the overall amino acid sequence could be important and could overrule individual sequence signals, e.g. if the overall charge or folding of the N-terminus is important (Christof Grüring, Dissertation 2011).

4.3 Folding

As charge and localisation of different amino acids do not explain the export phenotypes of all shown and known N-termini, not the primary structure of the amino acids but also the secondary structure was considered as an influencing factor. For this reason the helix forming propensity of each amino acid was taken into account. This parameter gives the likelihood of the respective amino acid to form an α-helix (Chakrabartty *et al*., 1994; Creamer & Rose, 1994; Pace & Scholtz, 1998) which in turn is essential for the folding of the emerging protein. The α-helical conformation is a common secondary structure. It is a stable, naturally occurring conformation of a linear amino acid sequence (Anfinsen, 1972; Scholtz & Baldwin, 1992). It depends on the interplay of amino acids in a sequence and follows complex rules, which can make it challenging to predict such a secondary structure with certainty. Secondary structures again can form tertiary structures, which will lead to the overall 3D shape of a protein. Several protein subunits can further build the quaternary structure (Schellman & Schellman, 1997). The 3D structure of a protein is decisive for its (unique) function (Anfinsen, 1973). It is built through hydrogen bonds between particular amino acids, which lead to a singular structure of every individual protein (Pauling *et al*., 1951; Kabsch & Sander, 1983). Another secondary structure comparable to the α-helix is a β-sheet. In comparison to the coiled α-helix, the β-sheet looks accordion-like (Anfinsen, 1973). A further secondary structure are the so called turns. There the C^{α} atoms of two residues are in close proximity to each other, but do not form a real secondary structure (Rose *et al*., 1985). Thus this structure might be considered having more of a spacer function.

Amino acids have different probabilities to form one of these secondary structures (Table 4.3). Different classifications exist in regards to the α-helix, β-sheet and turn forming propensities of the different amino acids (Chakrabartty *et al*., 1994; Pace & Scholtz, 1998). For this work the classification by Creighton was used (Table 4.3; Creighton, 1992). By the means of a secondary structure prediction tool, an even more different and complex scheme could be drawn.

Amino acid	α-helix	β-sheet	Turn
Ala	**1.29**	0.90	0.78
Cys	**1.11**	0.74	0.80
Leu	**1.30**	1.02	0.59
Met	**1.47**	0.97	0.39
Glu	**1.44**	0.75	1.00
Gln	**1.27**	0.80	0.97
His	**1.22**	1.08	0.69
Lys	**1.23**	0.77	0.96
Val	0.91	**1.49**	0.47
Ile	0.97	**1.45**	0.51
Phe	1.07	**1.32**	0.58
Tyr	0.72	**1.25**	1.05
Trp	0.99	**1.14**	0.75
Thr	0.82	**1.21**	1.03
Gly	0.56	0.92	**1.64**
Ser	0.82	0.95	**1.33**
Asp	1.04	0.72	**1.41**
Asn	0.90	0.76	**1.28**
Pro	0.52	0.64	**1.91**
Arg	0.96	0.99	0.88

Table 4.3 – Relative frequencies of amino acid residues in secondary structures. The amino acids are categorised regarding their preference to form α-helices (top), β-sheets (middle), or turns (bottom) based on the frequency these residues are found in the respective structures. Arginine is not classified, as it shows no significant preference for any of the structures. Amino acids are shown in three letter code. After T. E. Creighton, *Proteins: Structures and Molecular Properties,* 2d ed. (W. H. Freeman and Company, 1992), p. 256.

Many programs and calculators are known to predict the secondary structure of a protein based on its amino acid sequence. However the secondary structure prediction is extremely difficult as the protein can theoretically take many different possible conformations, depending on the charge, hydrophobicity and polarity of the different amino acids within the sequence and therefore depending on the actual hydrogen bonds formed (Anfinsen 1972 & 1973; Sikder & Zomaya, 2005). One program, which shows comparably reliable results, is PEP-FOLD (Thevenet *et al*., 2012). This program was used to predict the secondary structure of the neutral N-terminus (AGAGAGAGAG), which did not promote export, and its positive control (A**E**AGAGAGAG), which promoted export. The neutral N-terminus did not show a defined secondary structure such as an α-helix or a β-sheet and is probably arranged as a random coil (Figure 4.2A). On the other hand its positive control (harbouring a glutamic acid (E) at P2, which is the perfect primary signal) is predicted to show an α-helix at the beginning of its N-terminus (Figure 4.2B). As the only difference on the sequence level is the E at P2, this amino acid is clearly responsible for inducing the helix in this amino acid chain. Glutamic acid (E) shows a high helix forming propensity and is apparently sufficient to induce such a structure in this context. This α-helix located in close proximity to the N-terminus may contribute to the export promoting capacity of this N-terminus.

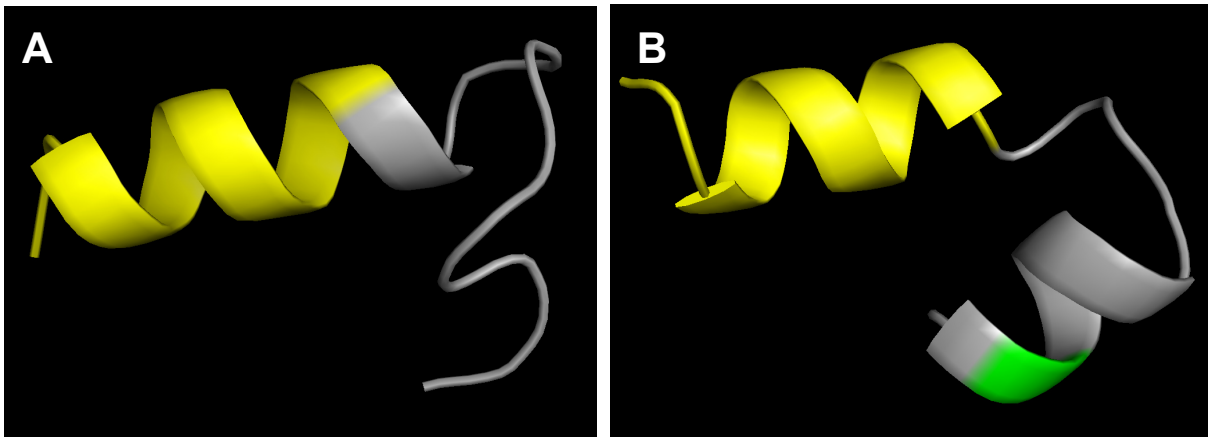


Figure 4.2 – The secondary structure prediction of the neutral N-terminus (A) and its positive control (B). The N-terminus is represented in grey, the positive signal (E) in green and the appended mTRAP reporter (first 10 amino acids) in yellow. Prediction program: PEP-FOLD, Thevenet *et al*., 2012.

Interestingly, the glutamine (Q, annotated as the second most prominent amino acid at P2), which resulted in a high export rate of its respective protein, shows in theory also a high helix forming propensity but was not predicted to induce an α-helix (see Appendix). Aspartic acid (D, annotated as the third most prominent amino acid at P2

(Hiller *et al*., 2004; Marti *et al*., 2004; Boddey *et al*., 2013)), which however did not promote protein export, shows a low helix forming propensity and was predicted to form a random coil (see Appendix), similar to the also not exported neutral N-terminus. Previously it was believed that the charge of amino acids is mainly responsible for the export profile of the respective protein. But as E is negatively charged and exported, Q is uncharged but exported and D is negatively charged but not exported (Table 4.4), it seems likely that helix forming propensity (in combination with or instead of charge) and the export promoting capacity are associated.

	charge	**helix prop.**	**export**
E	**neg.**	✓	✓
Q	**/**	✓	✓
D	**neg.**	✗	✗

Table 4.4 – The amino acids E, Q and D and their properties related to export. neg.: negatively charged; /: uncharged; ✓: high helix forming propensity/protein exported; ✗: low helix forming propensity/protein not exported.

Most of the other tested amino acids, serving as a primary positive signal, also fit this hypothesis. Leucine (L) and methionine (M) show comparably high to moderate export rates, respectively. Both amino acid residues show a high helix forming propensity in theory but did not appear to induce a helix in the neutral N-terminus when predicted by PEP-FOLD (see Appendix). Threonine (T) showed the highest theoretical export rate of all investigated amino acids at P2 (aside from E and Q), although also here the prediction of the actual protein did not show a helix (see Appendix). This amino acid is commonly found in antiparallel β-sheets. As other amino acids frequently located in β-sheets have not been investigated as a primary positive signal, no concrete conclusion can be drawn based on this finding. But it is in principle possible that β-sheets are – like α-helices – beneficial for the export of the respective protein. As E and Q resulted in higher export rates than T, it seems unlikely that a β-sheet forming propensity is more beneficial than α-helices, but maybe a β-sheet can provide a similar structural requirement that is beneficial for export. Serine (S) also resulted in a moderate export rate, but is mainly found in turns. As this amino acid can form bonds that prevent the formation of an α-helix, it is usually classified as a helix breaker (Lewis *et al*., 1970). At least in this case this does not prevent protein export. In accordance with this observation, this N-

terminus is predicted to hold an α-helix (see Appendix). This could be an explanation for the export of the respective protein, as it is possible that in this context S really induces the formation of an α-helix. In accordance with the results seen for n_S2, the N-terminus AASAASAASA also promoted export and was predicted to induce an α-helix. Taken together this highlights the fact that the helix forming propensity of an amino acid alone may be a poor predictor of the effect of this individual amino acid on the secondary structure of a peptide sequence. Asparagine (N) and arginine (R) resulted in a moderate export phenotype. Asparagine is uncharged and strongly related to the export promoting Q, whereas arginine is positively charged. These two amino acids are – like the negatively charged aspartic acid (D) – mainly found in turns. However, according to PEP-FOLD prediction N does not induce the formation of an α-helix, whereas R does (see Appendix). This further indicates that not the charge but other features are decisive for protein export. As the amino acids with the tendency to form a turn do not really form a structure but act more like a spacer to link α-helices and β-sheets in the tertiary structure, real structure building amino acids seem to be beneficial for protein export, whereas turn promoting residues seem to be less compatible with export. Alanine (A) seems to be a special case regarding its function as a primary positive signal. It caused nearly no export of the reporter and showed the lowest export rate compared to all tested primary positive signals. This is surprising as this small amino acid shows a high helix forming capacity, comparable to that of glutamine (Q), and was also predicted to induce an α-helix (see Appendix). Maybe in the context of the neutral N-terminus (AG-repeats) alanine at P2 cannot form a helix. The most likely explanation is – maybe beside the effect due to the surrounding sequence – that not helix forming propensity alone but other features of the amino acid, such as a certain structure or size of the side chain, also have to fit so that export is promoted. This may the reason why E, Q or T work well whereas A, S, D and N, which all are smaller, do not promote high export levels. As the sidechains of D and E as well as N and Q are very similar, it seems that size rather than the type of sidechain is also deciding. It should however also be noted that it is unclear if a single residue can influence the folding of the N-terminus. Another possibility is that – besides the formation of a secondary structure (like an α-helix or a β-sheet) – other properties such as polarity or hydrophobicity can also influence the export promoting capacity.

One striking observation was the fact that glutamic acid (E) could act as a secondary positive signal. The E was inserted at several different positions and the export rate was lower the further downstream of the N-terminus the E was located. Only when inserting the E at P3 (AG**E**GAGAGAG) the respective protein was not exported at all, although this position is in very close proximity to the N-terminus (which was generally found to be beneficial for protein export). The difference compared to many other constructs is the flanking by glycines (G) instead of alanines (A). Glycine is a very small amino acid, which makes the local peptide structure flexible. This is a unique feature within amino acids. G can be found in different confirmations (3_{10} helix, β-hairpin loops) but favours the unfolded state and shows therefore a high turn forming propensity (Rauscher *et al*., 2006; Cheng *et al*., 2010). This property was previously connected with a high helix breaking capacity and therefore a low export rate of the respective protein in most cases (D, R, N). Thus, two flanking glycines seem to act as a helix breaker what could correspond to a blocking sequence that abolishes the effect of E, an amino acid with a high helix forming propensity. As alanine has an annotated tendency to form (or even support the formation of) helices, this would also explain the difference of the flanking amino acids – alanines flanking an amino acid with a high helix forming propensity (e.g. E) support (or at least do not prevent) protein export, whereas flanking glycines block the export of the respective protein. This was confirmed by a construct where the flanking glycines were mutated to alanines (AG**E**GAGAGAG → A**AEA**GAGAG), which resulted in export of the protein. Thus mutating the flanking glycines rescued protein export. The construct, where E is inserted at P5 is the only one where this hypothesis does not fit completely (AGAG**E**GAGAG). Here, the export rate is low, but not lower than in the construct with E at P3, even though the E is inserted further downstream in the N-terminus and in addition flanked by Gs. The reason for this may be that the effect further downstream is different or that folding is affected in a different way in this position, as the secondary structure is influenced by a complex interplay of the amino acid chain in this region.

As the impact of a positive export signal (e.g. E) differs depending on its position in the N-terminus and the flanking amino acids, a triple E was inserted further downstream and resulted in export of the protein when flanked by alanines. Maybe a helix can form more easily when located at the beginning of an N-terminus (and therefore only one amino acid can be sufficient to induce it in a suitable background), whereas further downstream a stronger positive signal is needed (enhancing the positive signal in terms

of the glutamic acid from one E to a triplet of Es) to form a helix in this background. This assumption is also supported by the PEP-FOLD analysis regarding these two constructs, as n_E7 is not predicted to show a helix, whereas n_E7-9 is predicted to C-terminally append the already existing helix at the end of the N-terminus (see Appendix). Quite contrary to the possible detrimental role of Gs, both constructs harbouring Es in alternating positions, one flanked by glycines, one by alanines (AG**E**G**E**G**E**GAG and AGA**E**A**E**A**E**AG), resulted in similarly mixed phenotypes. It may be possible that the E residues, even if all are flanked by glycines, are sufficient to induce a helix, irrespective of the glycines, although PEP-FOLD predicts a helix for AGA**E**A**E**A**E**AG but not for AG**E**G**E**G**E**GAG (see Appendix). A positive signal covering a greater part of the N-terminus hence may overrule the negative export effect due to flanking amino acids (e.g. glycines), which may not be the case in the primary signal where a single residue has a much smaller effect on the overall secondary structure and may at best tip the balance towards a certain fold. Interestingly, constructs consisting of an N-terminus with GA repeats (instead of AG-repeats, GA**E**AGAGAGA (data not shown) and G**E**GAGAGAGA) did not result in export, although the E is found in the perfect primary position, indicating that the Gs may influence the overall folding of the N-terminus depending on small differences in order and position in the N-terminus.

The triple proline in the N-terminus of SERA also prevents the protein from being exported. As proline is the amino acid with the highest helix avoiding propensity and instead shows the highest turn forming propensity, it is a distinct helix breaker (Pace & Scholtz, 1998). Although the likelihoods to form helices and turns are almost equal for prolines and glycines, these two amino acids nevertheless hold contrary properties. Proline shows a cyclic side chain, which gives this amino acid an exceptional rigidity (Rauscher *et al*., 2006; Cheng *et al*., 2010). It further breaks hydrogen bonds, or at least cannot act as a hydrogen bond donor and is therefore not appropriate to form secondary structures. Hence, too much flexibility (glycine) as well as too much rigidity (proline) are usually detrimental for the formation of secondary structures such as α-helices or β-sheets and thus may negatively affect the export promoting capacity of N-termini.

In general the export promoting capacity of the different amino acids acting as a 1° positive export signal was ranked E>Q>T>L>M>S>N>R>D>A>G, from E showing the highest to G showing the lowest export of the respective fusion construct. Prolines were

not tested, as their blocking capacity was already validated. In comparison the helix forming propensity of the respective amino acids was M>E>L>A>Q>D>R>N>T/S>G, from M exhibiting the highest to G exhibiting the lowest helix forming propensity. As the export promoting and the helix forming capacity of the amino acids in comparison to each other do not coincide, the helix forming propensities of the different amino acids alone cannot be decisive for protein export. Further, theoretical helix-forming propensity and secondary structure prediction were not consistent in all investigated N-termini. As only the helix forming propensity of one amino acid within an N-terminus is not sufficient to influence to the secondary structure of the entire N-terminus, a reliable prediction is not possible. By the means of secondary structure prediction tools, which are not very accurate, drawing of a concrete conclusion regarding the secondary structure of the investigated N-termini is even more difficult. Further structural analysis such as CD and NMR of different N-termini would provide a better picture regarding the secondary structure of the respective proteins and possible correlations to the export promoting capacity of individual N-termini.

The reasons why folding of a protein is important for protein export, can be diverse. One reason could be that the secondary structure of the protein is recognised by another protein. This could be an adaptor during sorting of exported proteins in the secretory pathway or, as a known key component that may decide whether a protein remains in the PV lumen or is exported into the host cell, the PTEX translocon (de Koning-Ward, 2009). PTEX is located in the PVM, the membrane surrounding the parasite in the host cell. One of its essential key components is the chaperone HSP101 (de Koning-Ward *et al*., 2009), which is essential for protein export (Beck *et al*., 2014; Elsworth *et al*., 2014). As unfolding is important for the export process (Gehde *et al*., 2009; Grüring *et al*., 2012; Heiber *et al*., 2013; Mesén-Ramírez *et al*., 2016), HSP101 may bind the folded substrate protein, unfold it and under the use of ATP may pass it through the actual pore at the PVM into the host cell. Thus secondary structures like α-helices and β-sheets (which can further build up tertiary and quaternary structures) could be recognised by this chaperone and may be needed to initiate transport of the protein through the translocon. Preliminary binding assays using HA-tagged HSP101, purified from the parasite, with the neutral N-terminus (which is not exported and does probably not harbour an α-helix or β-sheet) compared to the positive control n_E2 (which is exported

and probably holds an α-helix) did not detect a binding preference for n_E2 (data not shown). It therefore remains unclear where the N-terminus acts and what recognises it.

The findings in this thesis indicate that no simple, generalised rule explains the capacity of the common core export domain to promote export. It rather seems that a combination of folding, amino acid charge, size and other sidechain properties as well as the exact position in the N-terminus influence export. The influence of folding complicates matters considerably, as the secondary structure of a protein is unique and not reliably predictable. Hence, a general export forecasting based on the amino acid sequence of an N-terminus may not possible. But with the new insights from this thesis the tendency of a protein to be exported or blocked, may at least in part be gleaned and for certain residues, such as E, this is likely reliable. Further analyses assessing all the so far not tested residues may lead to a scoring map based on the export promoting capacity of each residue depending on its position that may give a net export score for the entire N-terminus. This may in the end provide a prediction tool. However, this tool would likely also have to be extensively tested.

References

Achan, J., Talisuna, A.O., Erhart, A., Yeka, A., Tibenderana, J.K., Baliraine, F.N., Rosenthal, P.J., and D'Alessandro, U. (2011). Quinine, an old anti-malarial drug in a modern world: role in the treatment of malaria. Malar J *10*, 144.

Aikawa, M., Miller, L.H., and Rabbege, J. (1975). Caveola--vesicle complexes in the plasmalemma of erythrocytes infected by Plasmodium vivax and P cynomolgi. Unique structures related to Schuffner's dots. Am J Pathol *79*, 285-300.

Anfinsen, C.B. (1972). The formation and stabilization of protein structure. Biochem J *128*, 737-749.

Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. Science *181*, 223-230.

Arnot, D.E., and Gull, K. (1998). The Plasmodium cell-cycle: facts and questions. Ann Trop Med Parasitol *92*, 361-365.

Atkinson, C.T., and Aikawa, M. (1990). Ultrastructure of malaria-infected erythrocytes. Blood Cells *16*, 351-368.

Atkinson, C.T., Aikawa, M., Rock, E.P., Marsh, K., Andrysiak, P.M., Campbell, G.H., Collins, W.E., and Howard, R.J. (1987). Ultrastructure of the erythrocytic stages of Plasmodium malariae. J Protozool *34*, 267-274.

Attaran, A., and Maharaj, R. (2000). Ethical debate: doctoring malaria, badly: the global campaign to ban DDT. BMJ *321*, 1403-1405.

Bachmair, A., Finley, D., and Varshavsky, A. (1986). In vivo half-life of a protein is a function of its amino-terminal residue. Science *234*, 179-186.

Banerjee, R., Liu, J., Beatty, W., Pelosof, L., Klemba, M., and Goldberg, D.E. (2002). Four plasmepsins are active in the Plasmodium falciparum food vacuole, including a protease with an active-site histidine. Proc Natl Acad Sci U S A *99*, 990-995.

Bannister, L.H., Hopkins, J.M., Fowler, R.E., Krishna, S., and Mitchell, G.H. (2000). A brief illustrated guide to the ultrastructure of Plasmodium falciparum asexual blood stages. Parasitol Today *16*, 427-433.

Baum, J., Richard, D., Healer, J., Rug, M., Krnajski, Z., Gilberger, T.W., Green, J.L., Holder, A.A., and Cowman, A.F. (2006). A conserved molecular motor drives cell invasion and gliding motility across malaria life cycle stages and other apicomplexan parasites. J Biol Chem *281*, 5197-5208.

Beard, J., and Australian Rural Health Research, C. (2006). DDT and human health. Sci Total Environ *355*, 78-89.

Beck, J.R., Muralidharan, V., Oksman, A., and Goldberg, D.E. (2014). PTEX component HSP101 mediates export of diverse malaria effectors into host erythrocytes. Nature *511*, 592-595.

Bhattacharjee, S., Hiller, N.L., Liolios, K., Win, J., Kanneganti, T.D., Young, C., Kamoun, S., and Haldar, K. (2006). The malarial host-targeting signal is conserved in the Irish potato famine pathogen. PLoS Pathog *2*, e50.

Binka, F., and Akweongo, P. (2006). Prevention of malaria using ITNs: potential for achieving the millennium development goals. Curr Mol Med *6*, 261-267.

Blisnick, T., Morales Betoulle, M.E., Barale, J.C., Uzureau, P., Berry, L., Desroses, S., Fujioka, H., Mattei, D., and Braun Breton, C. (2000). Pfsbp1, a Maurer's cleft Plasmodium falciparum protein, is associated with the erythrocyte skeleton. Mol Biochem Parasitol *111*, 107-121.

Blobel, G., and Dobberstein, B. (1975). Transfer of proteins across membranes. II. Reconstitution of functional rough microsomes from heterologous components. J Cell Biol *67*, 852-862.

Blobel, G., and Dobberstein, B. (1975). Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. J Cell Biol *67*, 835-851.

Boddey, J.A., Carvalho, T.G., Hodder, A.N., Sargeant, T.J., Sleebs, B.E., Marapana, D., Lopaticki, S., Nebl, T., and Cowman, A.F. (2013). Role of plasmepsin V in export of diverse protein families from the Plasmodium falciparum exportome. Traffic *14*, 532-550.

Boddey, J.A., and Cowman, A.F. (2013). Plasmodium nesting: remaking the erythrocyte from the inside out. Annu Rev Microbiol *67*, 243-269.

Boddey, J.A., Hodder, A.N., Gunther, S., Gilson, P.R., Patsiouras, H., Kapp, E.A., Pearce, J.A., de Koning-Ward, T.F., Simpson, R.J., Crabb, B.S.*, et al.* (2010). An aspartyl protease directs malaria effector proteins to the host cell. Nature *463*, 627-631.

Boddey, J.A., Moritz, R.L., Simpson, R.J., and Cowman, A.F. (2009). Role of the Plasmodium export element in trafficking parasite proteins to the infected erythrocyte. Traffic *10*, 285-299.

Boddey, J.A., O'Neill, M.T., Lopaticki, S., Carvalho, T.G., Hodder, A.N., Nebl, T., Wawra, S., van West, P., Ebrahimzadeh, Z., Richard, D.*, et al.* (2016). Export of malaria proteins requires co-translational processing of the PEXEL motif independent of phosphatidylinositol-3-phosphate binding. Nat Commun *7*, 10470.

Bray, P.G., Hawley, S.R., and Ward, S.A. (1996). 4-Aminoquinoline resistance of Plasmodium falciparum: insights from the study of amodiaquine uptake. Mol Pharmacol *50*, 1551-1558.

Brown, E.M., Dower, H.J., and Greenberg, R. (1988). Effects of reductive alkylation on peptide retention times. J Chromatogr *443*, 247-254.

Capanna, E. (2006). Grassi versus Ross: who solved the riddle of malaria? Int Microbiol *9*, 69-74.

Chakrabartty, A., Kortemme, T., and Baldwin, R.L. (1994). Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. Protein Sci *3*, 843-852.

Chang, H.H., Falick, A.M., Carlton, P.M., Sedat, J.W., DeRisi, J.L., and Marletta, M.A. (2008). N-terminal processing of proteins exported by malaria parasites. Mol Biochem Parasitol *160*, 107-115.

Cheng, S., Cetinkaya, M., and Grater, F. (2010). How sequence determines elasticity of disordered proteins. Biophys J *99*, 3863-3869.

Chin, W., Contacos, P.G., Coatney, G.R., and Kimball, H.R. (1965). A Naturally Acquited Quotidian-Type Malaria in Man Transferable to Monkeys. Science *149*, 865.

Coatney, G.R. (1963). Pitfalls in a discovery: the chronicle of chloroquine. Am J Trop Med Hyg *12*, 121-128.

Coatney, G.R. (1971). The simian malarias: zoonoses, anthroponoses, or both? Am J Trop Med Hyg *20*, 795-803.

Cogswell, F.B. (1992). The hypnozoite and relapse in primate malaria. Clin Microbiol Rev *5*, 26-35.

Collins, W.E., and Jeffery, G.M. (2005). Plasmodium ovale: parasite and disease. Clin Microbiol Rev *18*, 570-581.

Collins, W.E., and Jeffery, G.M. (2007). Plasmodium malariae: parasite and disease. Clin Microbiol Rev *20*, 579-592.

Cooke, B.M., Buckingham, D.W., Glenister, F.K., Fernandez, K.M., Bannister, L.H., Marti, M., Mohandas, N., and Coppel, R.L. (2006). A Maurer's cleft-associated protein is

essential for expression of the major malaria virulence antigen on the surface of infected red blood cells. J Cell Biol *172*, 899-908.

Cooke, B.M., Lingelbach, K., Bannister, L.H., and Tilley, L. (2004). Protein trafficking in Plasmodium falciparum-infected red blood cells. Trends Parasitol *20*, 581-589.

Cowman, A.F., and Crabb, B.S. (2006). Invasion of red blood cells by malaria parasites. Cell *124*, 755-766.

Cox, F.E. (2002). History of human parasitology. Clin Microbiol Rev *15*, 595-612.

Cox-Singh, J., and Singh, B. (2008). Knowlesi malaria: newly emergent and of public health importance? Trends Parasitol *24*, 406-410.

Crabb, B.S., Cooke, B.M., Reeder, J.C., Waller, R.F., Caruana, S.R., Davern, K.M., Wickham, M.E., Brown, G.V., Coppel, R.L., and Cowman, A.F. (1997). Targeted gene disruption shows that knobs enable malaria-infected red cells to cytoadhere under physiological shear stress. Cell *89*, 287-296.

Creamer, T.P., and Rose, G.D. (1994). Alpha-helix-forming propensities in peptides and proteins. Proteins *19*, 85-97.

Creighton, T.E. (1992). Proteins: Structures and Molecular Properties, 2d ed. (W. H. Freeman and Company, 1992), p. 256.

Croft, A.M. (2007). A lesson learnt: the rise and fall of Lariam and Halfan. J R Soc Med *100*, 170-174.

Cyrklaff, M., Sanchez, C.P., Frischknecht, F., and Lanzer, M. (2012). Host actin remodeling and protection from malaria by hemoglobinopathies. Trends Parasitol *28*, 479-485.

de Koning-Ward, T.F., Gilson, P.R., Boddey, J.A., Rug, M., Smith, B.J., Papenfuss, A.T., Sanders, P.R., Lundie, R.J., Maier, A.G., Cowman, A.F*., et al.* (2009). A newly discovered protein export machine in malaria parasites. Nature *459*, 945-949.

de Niz, M., Ullrich*, A., Heiber*, A, Blancke Soares, A., Pick, C., Lyck, R., Keller, D., Kaiser, G., Prado, M., Flemming, S., del Portillo, H., Janse, C.J., T Heussler*, V., and Spielmann*, T. (2016). The machinery underlying malaria parasite virulence is conserved between rodent and human malaria parasites. Nat Commun.

de Vries, P.J., and Dien, T.K. (1996). Clinical pharmacology and therapeutic potential of artemisinin and its derivatives in the treatment of malaria. Drugs *52*, 818-836.

de Zulueta, J. (1998). The end of malaria in Europe: an eradication of the disease by control measures. Parassitologia *40*, 245-246.

Decherf, G., Egee, S., Staines, H.M., Ellory, J.C., and Thomas, S.L. (2004). Anionic channels in malaria-infected human red blood cells. Blood Cells Mol Dis *32*, 366-371.

Deitsch, K.W., and Wellems, T.E. (1996). Membrane modifications in erythrocytes parasitized by Plasmodium falciparum. Mol Biochem Parasitol *76*, 1-10.

Deponte, M., Hoppe, H.C., Lee, M.C., Maier, A.G., Richard, D., Rug, M., Spielmann, T., and Przyborski, J.M. (2012). Wherever I may roam: protein and membrane trafficking in P. falciparum-infected red blood cells. Mol Biochem Parasitol *186*, 95-116.

Desai, S.A., Bezrukov, S.M., and Zimmerberg, J. (2000). A voltage-dependent channel involved in nutrient uptake by red blood cells infected with the malaria parasite. Nature *406*, 1001-1005.

Dhangadamajhi, G., Kar, S.K., and Ranjit, M. (2010). The survival strategies of malaria parasite in the red blood cell and host cell polymorphisms. Malar Res Treat *2010*, 973094.

Dixon, M.W., Hawthorne, P.L., Spielmann, T., Anderson, K.L., Trenholme, K.R., and Gardiner, D.L. (2008). Targeting of the ring exported protein 1 to the Maurer's clefts is mediated by a two-phase process. Traffic *9*, 1316-1326.

Dower, G.E., Yakush, A., Nazzal, S.B., Jutzy, R.V., and Ruiz, C.E. (1988). Deriving the 12-lead electrocardiogram from four (EASI) electrodes. J Electrocardiol *21 Suppl*, S182-187.

Dower, K.M., Petney, T.N., and Horak, I.G. (1988). The developmental success of Amblyomma hebraeum and Amblyomma marmoreum on the leopard tortoise, Geochelone pardalis. Onderstepoort J Vet Res *55*, 11-13.

Dower, S.M., Piccinini, L.A., Roman, S.H., Safirstein, R., and Davies, T.F. (1988). Allospecific DR gene expression on the human thyroid cell. Autoimmunity *1*, 37-44.

Dower, W.J., Miller, J.F., and Ragsdale, C.W. (1988). High efficiency transformation of E. coli by high voltage electroporation. Nucleic Acids Res *16*, 6127-6145.

Eastman, R.T., and Fidock, D.A. (2009). Artemisinin-based combination therapies: a vital tool in efforts to eliminate malaria. Nat Rev Microbiol *7*, 864-874.

El Tahir, A., Malhotra, P., and Chauhan, V.S. (2003). Uptake of proteins and degradation of human serum albumin by Plasmodium falciparum-infected human erythrocytes. Malar J *2*, 11.

Elsworth, B., Matthews, K., Nie, C.Q., Kalanon, M., Charnaud, S.C., Sanders, P.R., Chisholm, S.A., Counihan, N.A., Shaw, P.J., Pino, P.*, et al.* (2014). PTEX is an essential nexus for protein export in malaria parasites. Nature *511*, 587-591.

Escalante, A.A., and Ayala, F.J. (1995). Evolutionary origin of Plasmodium and other Apicomplexa based on rRNA genes. Proc Natl Acad Sci U S A *92*, 5793-5797.

Flick, K., and Chen, Q. (2004). var genes, PfEMP1 and the human host. Mol Biochem Parasitol *134*, 3-9.

Foquet, L., Hermsen, C.C., van Gemert, G.J., Van Braeckel, E., Weening, K.E., Sauerwein, R., Meuleman, P., and Leroux-Roels, G. (2014). Vaccine-induced monoclonal antibodies targeting circumsporozoite protein prevent Plasmodium falciparum infection. J Clin Invest *124*, 140-144.

Francis, S.E., Sullivan, D.J., Jr., and Goldberg, D.E. (1997). Hemoglobin metabolism in the malaria parasite Plasmodium falciparum. Annu Rev Microbiol *51*, 97-123.

Frey, S.G., Chelo, D., Kinkela, M.N., Djoukoue, F., Tietche, F., Hatz, C., and Weber, P. (2010). Artesunate-mefloquine combination therapy in acute Plasmodium falciparum malaria in young children: a field study regarding neurological and neuropsychiatric safety. Malar J *9*, 291.

Gallup, J.L., and Sachs, J.D. (2001). The economic burden of malaria. Am J Trop Med Hyg *64*, 85-96.

Garcia, L.S. (2010). Malaria. Clin Lab Med *30*, 93-129.

Gehde, N., Hinrichs, C., Montilla, I., Charpian, S., Lingelbach, K., and Przyborski, J.M. (2009). Protein unfolding is an essential requirement for transport across the parasitophorous vacuolar membrane of Plasmodium falciparum. Mol Microbiol *71*, 613-628.

Gerald, N., Mahajan, B., and Kumar, S. (2011). Mitosis in the human malaria parasite Plasmodium falciparum. Eukaryot Cell *10*, 474-482.

Gilson, P.R., and Crabb, B.S. (2009). Morphology and kinetics of the three distinct phases of red blood cell invasion by Plasmodium falciparum merozoites. Int J Parasitol *39*, 91-96.

Glenister, F.K., Coppel, R.L., Cowman, A.F., Mohandas, N., and Cooke, B.M. (2002). Contribution of parasite proteins to altered mechanical properties of malaria-infected red blood cells. Blood *99*, 1060-1063.

Glushakova, S., Yin, D., Li, T., and Zimmerberg, J. (2005). Membrane transformation

during malaria parasite release from human red blood cells. Curr Biol *15*, 1645-1650.

Gold, D.A., Kaplan, A.D., Lis, A., Bett, G.C., Rosowski, E.E., Cirelli, K.M., Bougdour, A., Sidik, S.M., Beck, J.R., Lourido, S.*, et al.* (2015). The Toxoplasma Dense Granule Proteins GRA17 and GRA23 Mediate the Movement of Small Molecules between the Host and the Parasitophorous Vacuole. Cell Host Microbe *17*, 642-652.

Goldberg, D.E., and Cowman, A.F. (2010). Moving in and renovating: exporting proteins from Plasmodium into host erythrocytes. Nat Rev Microbiol *8*, 617-621.

Goldberg, D.E., Slater, A.F., Beavis, R., Chait, B., Cerami, A., and Henderson, G.B. (1991). Hemoglobin degradation in the human malaria pathogen Plasmodium falciparum: a catabolic pathway initiated by a specific aspartic protease. J Exp Med *173*, 961-969.

Goldberg, D.E., Slater, A.F., Cerami, A., and Henderson, G.B. (1990). Hemoglobin degradation in the malaria parasite Plasmodium falciparum: an ordered process in a unique organelle. Proc Natl Acad Sci U S A *87*, 2931-2935.

Greenwood, B.M., Fidock, D.A., Kyle, D.E., Kappe, S.H., Alonso, P.L., Collins, F.H., and Duffy, P.E. (2008). Malaria: progress, perils, and prospects for eradication. J Clin Invest *118*, 1266-1276.

Gruring, C., Heiber, A., Kruse, F., Flemming, S., Franci, G., Colombo, S.F., Fasana, E., Schoeler, H., Borgese, N., Stunnenberg, H.G.*, et al.* (2012). Uncovering common principles in protein export of malaria parasites. Cell Host Microbe *12*, 717-729.

Gruring, C., Heiber, A., Kruse, F., Ungefehr, J., Gilberger, T.W., and Spielmann, T. (2011). Development and host cell modifications of Plasmodium falciparum blood stages in four dimensions. Nat Commun *2*, 165.

Haase, S., Herrmann, S., Gruring, C., Heiber, A., Jansen, P.W., Langer, C., Treeck, M., Cabrera, A., Bruns, C., Struck, N.S.*, et al.* (2009). Sequence requirements for the export of the Plasmodium falciparum Maurer's clefts protein REX2. Mol Microbiol *71*, 1003-1017.

Hamada, Y. (2016). A novel N-terminal degradation reaction of peptides via N-amidination. Bioorg Med Chem Lett *26*, 1690-1695.

Hanssen, E., Carlton, P., Deed, S., Klonis, N., Sedat, J., DeRisi, J., and Tilley, L. (2010). Whole cell imaging reveals novel modular features of the exomembrane system of the malaria parasite, Plasmodium falciparum. Int J Parasitol *40*, 123-

134.

Hanssen, E., Dekiwadia, C., Riglar, D.T., Rug, M., Lemgruber, L., Cowman, A.F., Cyrklaff, M., Kudryashev, M., Frischknecht, F., Baum, J.*, et al.* (2013). Electron tomography of Plasmodium falciparum merozoites reveals core cellular events that underpin erythrocyte invasion. Cell Microbiol *15*, 1457-1472.

Hanssen, E., Sougrat, R., Frankland, S., Deed, S., Klonis, N., Lippincott-Schwartz, J., and Tilley, L. (2008). Electron tomography of the Maurer's cleft organelles of Plasmodium falciparum-infected erythrocytes reveals novel structural features. Mol Microbiol *67*, 703-718.

Hawthorne, P.L., Trenholme, K.R., Skinner-Adams, T.S., Spielmann, T., Fischer, K., Dixon, M.W., Ortega, M.R., Anderson, K.L., Kemp, D.J., and Gardiner, D.L. (2004). A novel Plasmodium falciparum ring stage protein, REX, is located in Maurer's clefts. Mol Biochem Parasitol *136*, 181-189.

Hay, S.I., Guerra, C.A., Tatem, A.J., Noor, A.M., and Snow, R.W. (2004). The global distribution and population at risk of malaria: past, present, and future. Lancet Infect Dis *4*, 327-336.

Heiber, A., Kruse, F., Pick, C., Gruring, C., Flemming, S., Oberli, A., Schoeler, H., Retzlaff, S., Mesen-Ramirez, P., Hiss, J.A.*, et al.* (2013). Identification of new PNEPs indicates a substantial non-PEXEL exportome and underpins common features in Plasmodium falciparum protein export. PLoS Pathog *9*, e1003546.

Henrich, P., Kilian, N., Lanzer, M., and Cyrklaff, M. (2009). 3-D analysis of the Plasmodium falciparum Maurer's clefts using different electron tomographic approaches. Biotechnol J *4*, 888-894.

Hiller, N.L., Bhattacharjee, S., van Ooij, C., Liolios, K., Harrison, T., Lopez-Estrano, C., and Haldar, K. (2004). A host-targeting signal in virulence proteins reveals a secretome in malarial infection. Science *306*, 1934-1937.

Hiss, J.A., Przyborski, J.M., Schwarte, F., Lingelbach, K., and Schneider, G. (2008). The Plasmodium export element revisited. PLoS One *3*, e1560.

Horrocks, P., Pinches, R.A., Chakravorty, S.J., Papakrivos, J., Christodoulou, Z., Kyes, S.A., Urban, B.C., Ferguson, D.J., and Newbold, C.I. (2005). PfEMP1 expression is reduced on the surface of knobless Plasmodium falciparum infected erythrocytes. J Cell Sci *118*, 2507-2518.

Jeffery, C.J. (2003). Moonlighting proteins: old proteins learning new tricks. Trends Genet *19*, 415-417.

Johnson, A.E., and van Waes, M.A. (1999). The translocon: a dynamic gateway at the ER membrane. Annu Rev Cell Dev Biol *15*, 799-842.

Johnson, D., Gunther, K., Ansorge, I., Benting, J., Kent, A., Bannister, L., Ridley, R., and Lingelbach, K. (1994). Characterization of membrane proteins exported from Plasmodium falciparum into the host erythrocyte. Parasitology *109 (Pt 1)*, 1-9.

Johnston, S.P., Pieniazek, N.J., Xayavong, M.V., Slemenda, S.B., Wilkins, P.P., and da Silva, A.J. (2006). PCR as a confirmatory technique for laboratory diagnosis of malaria. J Clin Microbiol *44*, 1087-1089.

Jones, M.L., Kitson, E.L., and Rayner, J.C. (2006). Plasmodium falciparum erythrocyte invasion: a conserved myosin associated complex. Mol Biochem Parasitol *147*, 74-84.

Josling, G.A., and Llinas, M. (2015). Sexual development in Plasmodium parasites: knowing when it's time to commit. Nat Rev Microbiol *13*, 573-587.

Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers *22*, 2577-2637.

Kara, U.A., Stenzel, D.J., Ingram, L.T., and Kidson, C. (1988). The parasitophorous vacuole membrane of Plasmodium falciparum: demonstration of vesicle formation using an immunoprobe. Eur J Cell Biol *46*, 9-17.

Knuepfer, E., Rug, M., and Cowman, A.F. (2005). Function of the plasmodium export element can be blocked by green fluorescent protein. Mol Biochem Parasitol *142*, 258-262.

Kozarsky, K.F., Call, S.M., Dower, S.K., and Krieger, M. (1988). Abnormal intracellular sorting of O-linked carbohydrate-deficient interleukin-2 receptors. Mol Cell Biol *8*, 3357-3363.

Krafts, K., Hempelmann, E., and Skorska-Stania, A. (2012). From methylene blue to chloroquine: a brief review of the development of an antimalarial therapy. Parasitol Res *111*, 1-6.

Kriek, N., Tilley, L., Horrocks, P., Pinches, R., Elford, B.C., Ferguson, D.J., Lingelbach, K., and Newbold, C.I. (2003). Characterization of the pathway for transport of the cytoadherence-mediating protein, PfEMP1, to the host cell surface in malaria parasite-infected erythrocytes. Mol Microbiol *50*, 1215-1227.

Kulzer, S., Charnaud, S., Dagan, T., Riedel, J., Mandal, P., Pesce, E.R., Blatch, G.L., Crabb, B.S., Gilson, P.R., and Przyborski, J.M. (2012). Plasmodium falciparum-encoded

exported hsp70/hsp40 chaperone/co-chaperone complexes within the host erythrocyte. Cell Microbiol *14*, 1784-1795.

Kulzer, S., Rug, M., Brinkmann, K., Cannon, P., Cowman, A., Lingelbach, K., Blatch, G.L., Maier, A.G., and Przyborski, J.M. (2010). Parasite-encoded Hsp40 proteins define novel mobile structures in the cytosol of the P. falciparum-infected erythrocyte. Cell Microbiol *12*, 1398-1420.

Kupper, T.S., Horowitz, M., Birchall, N., Mizutani, H., Coleman, D., McGuire, J., Flood, P., Dower, S., and Lee, F. (1988). Hematopoietic, lymphopoietic, and proinflammatory cytokines produced by human and murine keratinocytes. Ann N Y Acad Sci *548*, 262-270.

Kupper, T.S., Lee, F., Birchall, N., Clark, S., and Dower, S. (1988). Interleukin 1 binds to specific receptors on human keratinocytes and induces granulocyte macrophage colony-stimulating factor mRNA and protein. A potential autocrine role for interleukin 1 in epidermis. J Clin Invest *82*, 1787-1792.

Kyes, S., Horrocks, P., and Newbold, C. (2001). Antigenic variation at the infected red cell surface in malaria. Annu Rev Microbiol *55*, 673-707.

Lacks, S., and Greenberg, B. (1975). A deoxyribonuclease of Diplococcus pneumoniae specific for methylated DNA. J Biol Chem *250*, 4060-4066.

Lanzer, M., Wickert, H., Krohne, G., Vincensini, L., and Braun Breton, C. (2006). Maurer's clefts: a novel multi-functional organelle in the cytoplasm of Plasmodium falciparum-infected erythrocytes. Int J Parasitol *36*, 23-36.

Lauer, S.A., Rathod, P.K., Ghori, N., and Haldar, K. (1997). A membrane network for nutrient import in red cells infected with the malaria parasite. Science *276*, 1122-1125.

Laveran, C.L. (1982). Classics in infectious diseases: A newly discovered parasite in the blood of patients suffering from malaria. Parasitic etiology of attacks of malaria: Charles Louis Alphonse Laveran (1845-1922). Rev Infect Dis *4*, 908-911.

Leech, J.H., Barnwell, J.W., Aikawa, M., Miller, L.H., and Howard, R.J. (1984). Plasmodium falciparum malaria: association of knobs on the surface of infected erythrocytes with a histidine-rich protein and the erythrocyte skeleton. J Cell Biol *98*, 1256-1264.

Lewis, P.N., Go, N., Go, M., Kotelchuck, D., and Scheraga, H.A. (1970). Helix probability profiles of denatured proteins and their correlation with native structures. Proc Natl Acad Sci U S A *65*, 810-815.

Lingelbach, K., and Przyborski, J.M. (2006). The long and winding road: protein trafficking mechanisms in the Plasmodium falciparum infected erythrocyte. Mol Biochem Parasitol *147*, 1-8.

Lingelbach, K., and Przyborski, J.M. (2006). The long and winding road: protein trafficking mechanisms in the Plasmodium falciparum infected erythrocyte. Mol Biochem Parasitol *147*, 1-8.

Maier, A.G., Cooke, B.M., Cowman, A.F., and Tilley, L. (2009). Malaria parasite proteins that remodel the host erythrocyte. Nat Rev Microbiol *7*, 341-354.

Maier, A.G., Rug, M., O'Neill, M.T., Brown, M., Chakravorty, S., Szestak, T., Chesson, J., Wu, Y., Hughes, K., Coppel, R.L., *et al.* (2008). Exported proteins required for virulence and rigidity of Plasmodium falciparum-infected human erythrocytes. Cell *134*, 48-61.

Markus, M.B. (1976). Possible support for the sporozoite hypothesis of relapse and latency in malaria. Trans R Soc Trop Med Hyg *70*, 535.

Marti, M., Good, R.T., Rug, M., Knuepfer, E., and Cowman, A.F. (2004). Targeting malaria virulence and remodeling proteins to the host erythrocyte. Science *306*, 1930-1933.

Marti, M., and Spielmann, T. (2013). Protein export in malaria parasites: many membranes to cross. Curr Opin Microbiol *16*, 445-451.

Matz, J.M., Ingmundson, A., Costa Nunes, J., Stenzel, W., Matuschewski, K., and Kooij, T.W. (2015). In Vivo Function of PTEX88 in Malaria Parasite Sequestration and Virulence. Eukaryot Cell *14*, 528-534.

Matz, J.M., Matuschewski, K., and Kooij, T.W. (2013). Two putative protein export regulators promote Plasmodium blood stage development in vivo. Mol Biochem Parasitol *191*, 44-52.

McKenzie, F.E., Jeffery, G.M., and Collins, W.E. (2001). Plasmodium malariae blood-stage dynamics. J Parasitol *87*, 626-637.

McKenzie, F.E., Jeffery, G.M., and Collins, W.E. (2002). Plasmodium vivax blood-stage dynamics. J Parasitol *88*, 521-535.

McQueen, P.G., and McKenzie, F.E. (2004). Age-structured red blood cell susceptibility and the dynamics of malaria infections. Proc Natl Acad Sci U S A *101*, 9161-9166.

Mendis, K., Sina, B.J., Marchesini, P., and Carter, R. (2001). The neglected burden of

Plasmodium vivax malaria. Am J Trop Med Hyg *64*, 97-106.

Mesén-Ramírez, P., Reinsch, F., Blancke Soares, A, Bergmann, B., Ullrich, A., Tenzer, S., and Spielmann*, T. (2016). Stable translocation intermediates jam global protein export in *Plasmodium falciparum* parasites and link the PTEX component EXP2 with translocation activity. PLoS Pathog.

Meyer, C.G., Marks, F., and May, J. (2004). Editorial: Gin tonic revisited. Trop Med Int Health *9*, 1239-1240.

Miller, J.F., Dower, W.J., and Tompkins, L.S. (1988). High-voltage electroporation of bacteria: genetic transformation of Campylobacter jejuni with plasmid DNA. Proc Natl Acad Sci U S A *85*, 856-860.

Miller, L.H., Baruch, D.I., Marsh, K., and Doumbo, O.K. (2002). The pathogenic basis of malaria. Nature *415*, 673-679.

Moore, L.R., Fujioka, H., Williams, P.S., Chalmers, J.J., Grimberg, B., Zimmerman, P.A., and Zborowski, M. (2006). Hemoglobin degradation in malaria-infected erythrocytes determined from live cell magnetophoresis. FASEB J *20*, 747-749.

Mueller, I., Zimmerman, P.A., and Reeder, J.C. (2007). Plasmodium malariae and Plasmodium ovale--the "bashful" malaria parasites. Trends Parasitol *23*, 278-283.

Mundwiler-Pachlatko, E., and Beck, H.P. (2013). Maurer's clefts, the enigma of Plasmodium falciparum. Proc Natl Acad Sci U S A *110*, 19987-19994.

Naing, C., Whittaker, M.A., Nyunt Wai, V., and Mak, J.W. (2014). Is Plasmodium vivax malaria a severe malaria?: a systematic review and meta-analysis. PLoS Negl Trop Dis *8*, e3071.

Niang, M., Bei, A.K., Madnani, K.G., Pelly, S., Dankwa, S., Kanjee, U., Gunalan, K., Amaladoss, A., Yeo, K.P., Bob, N.S*., et al.* (2014). STEVOR is a Plasmodium falciparum erythrocyte binding protein that mediates merozoite invasion and rosetting. Cell Host Microbe *16*, 81-93.

Niang, M., Yan Yam, X., and Preiser, P.R. (2009). The Plasmodium falciparum STEVOR multigene family mediates antigenic variation of the infected erythrocyte. PLoS Pathog *5*, e1000307.

Oh, S.S., Voigt, S., Fisher, D., Yi, S.J., LeRoy, P.J., Derick, L.H., Liu, S., and Chishti, A.H. (2000). Plasmodium falciparum erythrocyte membrane protein 1 is anchored to the actin-spectrin junction and knob-associated histidine-rich protein in the erythrocyte skeleton. Mol Biochem Parasitol *108*, 237-247.

Pace, C.N., and Scholtz, J.M. (1998). A helix propensity scale based on experimental studies of peptides and proteins. Biophys J *75*, 422-427.

Pachlatko, E., Rusch, S., Muller, A., Hemphill, A., Tilley, L., Hanssen, E., and Beck, H.P. (2010). MAHRP2, an exported protein of Plasmodium falciparum, is an essential component of Maurer's cleft tethers. Mol Microbiol *77*, 1136-1152.

Pauling, L., Corey, R.B., and Branson, H.R. (1951). The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. Proc Natl Acad Sci U S A *37*, 205-211.

Petersen, W., Kulzer, S., Engels, S., Zhang, Q., Ingmundson, A., Rug, M., Maier, A.G., and Przyborski, J.M. (2016). J-dot targeting of an exported HSP40 in Plasmodium falciparum-infected erythrocytes. Int J Parasitol.

Prajapati, S.K., Culleton, R., and Singh, O.P. (2014). Protein trafficking in Plasmodium falciparum-infected red cells and impact of the expansion of exported protein families. Parasitology, 1-11.

Przyborski, J.M., Miller, S.K., Pfahler, J.M., Henrich, P.P., Rohrbach, P., Crabb, B.S., and Lanzer, M. (2005). Protein transport and trafficking in Plasmodium falciparum-infected erythrocytes. EMBO J *24*, 2306-2317.

Qwarnstrom, E.E., Page, R.C., Gillis, S., and Dower, S.K. (1988). Binding, internalization, and intracellular localization of interleukin-1 beta in human diploid fibroblasts. J Biol Chem *263*, 8261-8269.

Rauscher, S., Baud, S., Miao, M., Keeley, F.W., and Pomes, R. (2006). Proline and glycine control protein self-organization into elastomeric or amyloid fibrils. Structure *14*, 1667-1676.

Riglar, D.T., Richard, D., Wilson, D.W., Boyle, M.J., Dekiwadia, C., Turnbull, L., Angrisano, F., Marapana, D.S., Rogers, K.L., Whitchurch, C.B.*, et al.* (2011). Super-resolution dissection of coordinated events during malaria parasite invasion of the human erythrocyte. Cell Host Microbe *9*, 9-20.

Rose, G.D., Gierasch, L.M., and Smith, J.A. (1985). Turns in peptides and proteins. Adv Protein Chem *37*, 1-109.

Rothman, J.E. (1994). Mechanisms of intracellular protein transport. Nature *372*, 55-63.

Rowe, J.A., Claessens, A., Corrigan, R.A., and Arman, M. (2009). Adhesion of Plasmodium falciparum-infected erythrocytes to human cells: molecular mechanisms and therapeutic implications. Expert Rev Mol Med *11*, e16.

Rowe, J.A., and Kyes, S.A. (2004). The role of Plasmodium falciparum var genes in malaria in pregnancy. Mol Microbiol *53*, 1011-1019.

Rug, M., Prescott, S.W., Fernandez, K.M., Cooke, B.M., and Cowman, A.F. (2006). The role of KAHRP domains in knob formation and cytoadherence of P falciparum-infected human erythrocytes. Blood *108*, 370-378.

Russo, I., Babbitt, S., Muralidharan, V., Butler, T., Oksman, A., and Goldberg, D.E. (2010). Plasmepsin V licenses Plasmodium proteins for export into the host erythrocyte. Nature *463*, 632-636.

Saliba, K.J., Horner, H.A., and Kirk, K. (1998). Transport and metabolism of the essential vitamin pantothenic acid in human erythrocytes infected with the malaria parasite Plasmodium falciparum. J Biol Chem *273*, 10190-10195.

Sam-Yellowe, T.Y. (1996). Rhoptry organelles of the apicomplexa: Their role in host cell invasion and intracellular survival. Parasitol Today *12*, 308-316.

Sam-Yellowe, T.Y. (1996). Rhoptry organelles of the apicomplexa: Their role in host cell invasion and intracellular survival. Parasitol Today *12*, 308-316.

Sargeant, T.J., Marti, M., Caler, E., Carlton, J.M., Simpson, K., Speed, T.P., and Cowman, A.F. (2006). Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. Genome Biol *7*, R12.

Saridaki, T., Frohlich, K.S., Braun-Breton, C., and Lanzer, M. (2009). Export of PfSBP1 to the Plasmodium falciparum Maurer's clefts. Traffic *10*, 137-152.

Saunders, R., and Deane, C.M. (2010). Synonymous codon usage influences the local protein structure observed. Nucleic Acids Res *38*, 6719-6728.

Schellman, J.A., and Schellman, C.G. (1997). Kaj Ulrik Linderstrom-Lang (1896-1959). Protein Sci *6*, 1092-1100.

Schlagenhauf, P., Adamcova, M., Regep, L., Schaerer, M.T., and Rhein, H.G. (2010). The position of mefloquine as a 21st century malaria chemoprophylaxis. Malar J *9*, 357.

Scholtz, J.M., and Baldwin, R.L. (1992). The mechanism of alpha-helix formation by peptides. Annu Rev Biophys Biomol Struct *21*, 95-118.

Schulze, J., Kwiatkowski, M., Borner, J., Schluter, H., Bruchhaus, I., Burmester, T., Spielmann, T., and Pick, C. (2015). The Plasmodium falciparum exportome contains non-canonical PEXEL/HT proteins. Mol Microbiol *97*, 301-314.

Shigekawa, K., and Dower, W.J. (1988). Electroporation of eukaryotes and prokaryotes: a

general approach to the introduction of macromolecules into cells. Biotechniques *6*, 742-751.

Shute, P.G. (1946). Latency and long-term relapses in benign tertian malaria. Trans R Soc Trop Med Hyg *40*, 189-200.

Sikder, A.R., and Zomaya, A.Y. (2005). An overview of protein-folding techniques: issues and perspectives. Int J Bioinform Res Appl *1*, 121-143.

Singh, B., and Daneshvar, C. (2013). Human infections and detection of Plasmodium knowlesi. Clin Microbiol Rev *26*, 165-184.

Singh, S., Alam, M.M., Pal-Bhowmick, I., Brzostowski, J.A., and Chitnis, C.E. (2010). Distinct external signals trigger sequential release of apical organelles during erythrocyte invasion by malaria parasites. PLoS Pathog *6*, e1000746.

Sleebs, B.E., Gazdik, M., O'Neill, M.T., Rajasekaran, P., Lopaticki, S., Lackovic, K., Lowes, K., Smith, B.J., Cowman, A.F., and Boddey, J.A. (2014). Transition state mimetics of the Plasmodium export element are potent inhibitors of Plasmepsin V from P. falciparum and P. vivax. J Med Chem *57*, 7644-7662.

Sleebs, B.E., Lopaticki, S., Marapana, D.S., O'Neill, M.T., Rajasekaran, P., Gazdik, M., Gunther, S., Whitehead, L.W., Lowes, K.N., Barfod, L.*, et al.* (2014). Inhibition of Plasmepsin V activity demonstrates its essential role in protein export, PfEMP1 display, and survival of malaria parasites. PLoS Biol *12*, e1001897.

Snow, R.W., Guerra, C.A., Noor, A.M., Myint, H.Y., and Hay, S.I. (2005). The global distribution of clinical episodes of Plasmodium falciparum malaria. Nature *434*, 214-217.

Spielmann, T., and Gilberger, T.W. (2010). Protein export in malaria parasites: do multiple export motifs add up to multiple export pathways? Trends Parasitol *26*, 6-10.

Spielmann, T., and Gilberger, T.W. (2015). Critical Steps in Protein Export of Plasmodium falciparum Blood Stages. Trends Parasitol *31*, 514-525.

Spielmann, T., Hawthorne, P.L., Dixon, M.W., Hannemann, M., Klotz, K., Kemp, D.J., Klonis, N., Tilley, L., Trenholme, K.R., and Gardiner, D.L. (2006). A cluster of ring stage-specific genes linked to a locus implicated in cytoadherence in Plasmodium falciparum codes for PEXEL-negative and PEXEL-positive proteins exported into the host cell. Mol Biol Cell *17*, 3613-3624.

Spillman, N.J., Beck, J.R., and Goldberg, D.E. (2015). Protein export into malaria parasite-infected erythrocytes: mechanisms and functional consequences. Annu Rev

Biochem *84*, 813-841.

Spycher, C., Klonis, N., Spielmann, T., Kump, E., Steiger, S., Tilley, L., and Beck, H.P. (2003). MAHRP-1, a novel Plasmodium falciparum histidine-rich protein, binds ferriprotoporphyrin IX and localizes to the Maurer's clefts. J Biol Chem *278*, 35373-35383.

Spycher, C., Rug, M., Klonis, N., Ferguson, D.J., Cowman, A.F., Beck, H.P., and Tilley, L. (2006). Genesis of and trafficking to the Maurer's clefts of Plasmodium falciparum-infected erythrocytes. Mol Cell Biol *26*, 4074-4085.

Spycher, C., Rug, M., Pachlatko, E., Hanssen, E., Ferguson, D., Cowman, A.F., Tilley, L., and Beck, H.P. (2008). The Maurer's cleft protein MAHRP1 is essential for trafficking of PfEMP1 to the surface of Plasmodium falciparum-infected erythrocytes. Mol Microbiol *68*, 1300-1314.

Staines, H.M., Alkhalil, A., Allen, R.J., De Jonge, H.R., Derbyshire, E., Egee, S., Ginsburg, H., Hill, D.A., Huber, S.M., Kirk, K*., et al.* (2007). Electrophysiological studies of malaria parasite-infected erythrocytes: current status. Int J Parasitol *37*, 475-482.

Staines, H.M., Powell, T., Thomas, S.L., and Ellory, J.C. (2004). Plasmodium falciparum-induced channels. Int J Parasitol *34*, 665-673.

Sturm, A., Amino, R., van de Sand, C., Regen, T., Retzlaff, S., Rennenberg, A., Krueger, A., Pollok, J.M., Menard, R., and Heussler, V.T. (2006). Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids. Science *313*, 1287-1290.

Stya, M., Dower, S.K., Prickett, K.S., Gillis, S., and Conlon, P.J. (1988). Development and characterization of two neutralizing monoclonal antibodies to human interleukin-1 alpha. J Biol Response Mod *7*, 162-172.

Sullivan, D.J., Jr., Gluzman, I.Y., and Goldberg, D.E. (1996). Plasmodium hemozoin formation mediated by histidine-rich proteins. Science *271*, 219-222.

Sutherland, C. (2007). A challenge for the development of malaria vaccines: polymorphic target antigens. PLoS Med *4*, e116.

Tamez, P.A., Bhattacharjee, S., van Ooij, C., Hiller, N.L., Llinas, M., Balu, B., Adams, J.H., and Haldar, K. (2008). An erythrocyte vesicle protein exported by the malaria parasite promotes tubovesicular lipid import from the host cell surface. PLoS Pathog *4*, e1000118.

Tangpukdee, N., Duangdee, C., Wilairatana, P., and Krudsood, S. (2009). Malaria

diagnosis: a brief review. Korean J Parasitol *47*, 93-102.

Tarr, S.J., Cryar, A., Thalassinos, K., Haldar, K., and Osborne, A.R. (2013). The C-terminal portion of the cleaved HT motif is necessary and sufficient to mediate export of proteins from the malaria parasite into its host cell. Mol Microbiol *87*, 835-850.

Templeton, T.J. (2009). The varieties of gene amplification, diversification and hypervariability in the human malaria parasite, Plasmodium falciparum. Mol Biochem Parasitol *166*, 109-116.

Thevenet, P., Shen, Y., Maupetit, J., Guyon, F., Derreumaux, P., and Tuffery, P. (2012). PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. Nucleic Acids Res *40*, W288-293.

Tilley, L., Sougrat, R., Lithgow, T., and Hanssen, E. (2008). The twists and turns of Maurer's cleft trafficking in P. falciparum-infected erythrocytes. Traffic *9*, 187-197.

Trager, W., and Jensen, J.B. (1976). Human malaria parasites in continuous culture. Science *193*, 673-675.

Trager, W., Rudzinska, M.A., and Bradbury, P.C. (1966). The fine structure of Plasmodium falciparum and its host erythrocytes in natural malarial infections in man. Bull World Health Organ *35*, 883-885.

Trampuz, A., Jereb, M., Muzlovic, I., and Prabhu, R.M. (2003). Clinical review: Severe malaria. Crit Care *7*, 315-323.

Tu, Y. (2011). The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. Nat Med *17*, 1217-1220.

Udagama, P.V., Atkinson, C.T., Peiris, J.S., David, P.H., Mendis, K.N., and Aikawa, M. (1988). Immunoelectron microscopy of Schuffner's dots in Plasmodium vivax-infected human erythrocytes. Am J Pathol *131*, 48-52.

Urdal, D.L., Call, S.M., Jackson, J.L., and Dower, S.K. (1988). Affinity purification and chemical analysis of the interleukin-1 receptor. J Biol Chem *263*, 2870-2877.

van Ooij, C., Tamez, P., Bhattacharjee, S., Hiller, N.L., Harrison, T., Liolios, K., Kooij, T., Ramesar, J., Balu, B., Adams, J.*, et al.* (2008). The malaria secretome: from algorithms to essential function in blood stage infection. PLoS Pathog *4*, e1000084.

Waller, R.F., Reed, M.B., Cowman, A.F., and McFadden, G.I. (2000). Protein trafficking to the plastid of Plasmodium falciparum is via the secretory pathway. EMBO J *19*,

1794-1802.

Wellems, T.E., Hayton, K., and Fairhurst, R.M. (2009). The impact of malaria parasitism: from corpuscles to communities. J Clin Invest *119*, 2496-2505.

Wellems, T.E., and Plowe, C.V. (2001). Chloroquine-resistant malaria. J Infect Dis *184*, 770-776.

White, N.J. (2004). Antimalarial drug resistance. J Clin Invest *113*, 1084-1092.

Wickert, H., Gottler, W., Krohne, G., and Lanzer, M. (2004). Maurer's cleft organization in the cytoplasm of plasmodium falciparum-infected erythrocytes: new insights from three-dimensional reconstruction of serial ultrathin sections. Eur J Cell Biol *83*, 567-582.

Wickert, H., Wissing, F., Andrews, K.T., Stich, A., Krohne, G., and Lanzer, M. (2003). Evidence for trafficking of PfEMP1 to the surface of P. falciparum-infected erythrocytes via a complex membrane network. Eur J Cell Biol *82*, 271-284.

Wickham, M.E., Rug, M., Ralph, S.A., Klonis, N., McFadden, G.I., Tilley, L., and Cowman, A.F. (2001). Trafficking and assembly of the cytoadherence complex in Plasmodium falciparum-infected human erythrocytes. EMBO J *20*, 5636-5649.

Winzeler, E.A., and Manary, M.J. (2014). Drug resistance genomics of the antimalarial drug artemisinin. Genome Biol *15*, 544.

Winzeler, E.A., and Manary, M.J. (2014). Drug resistance genomics of the antimalarial drug artemisinin. Genome Biol *15*, 544.

Wright, G.J., and Rayner, J.C. (2014). Plasmodium falciparum erythrocyte invasion: combining function with immune evasion. PLoS Pathog *10*, e1003943.

Publications

De Niz, M., **Ullrich*, A**., Heiber*, A, Blancke Soares, A., Pick, C., Lyck, R., Keller, D., Kaiser, G., Prado, M., Flemming, S., del Portillo, H., Janse, C.J., T Heussler*, V., and Spielmann*, T. (2016). The machinery underlying malaria parasite virulence is conserved between rodent and human malaria parasites. Nat Commun.

Mesén-Ramírez, P., Reinsch, F., Blancke Soares, A, Bergmann, B., **Ullrich, A.**, Tenzer, S., and Spielmann*, T. (2016). Stable translocation intermediates jam global protein export in *Plasmodium falciparum* parasites and link the PTEX component EXP2 with translocation activity. PLOS Path.

Appendix

N-termini classified based on their export promoting capacity.

Bioinformatics analysis. N-termini investigated until 2013 are included.

Exported proteins	Sequence
MAHRP1$^{1-20}$	AEQAAVQPES
MAHRP2$^{1-20}$	QPCPYDVYNQ
REX1$^{1-38}$	ADYSSNEEET
SBP1$^{1-26}$	CSAARAFDFD
STEVOR$^{44-63}$	AQTQIHNPHY
STEVOR $^{44-63}$Q45A	AATQIHNPHY
STEVOR $^{44-63}$scr_Q3	AQNIEQIPIE
GBP 130$^{87-106}$	AEGEDTCARK
GBP130$^{87-106}$ E88A	AAGEDTCARK
GBP130 $^{87-106}$scr_E3	AETDTCAKKG
ETRAMP5$^{25-44}$D27A	AQLDMGSVHN
REX2$^{1-20}$	KMYLAEIFSS
PFF0090w	TDHLLDFNMY
PF07_0007	SQPQKQQNEE
PF11_0505	EAEKKEEKQE
PF08_0003	NLEQFKNINK
PFL2515c	SDPWADYDPN
PFI1755$_{c1-61}$	SEPVVEEQDL
PfEMP3$_{1-82}$	AQVLGNTRLS
PFI1755c$_{1-61}$:GFP_EP2A	SAPVVEEQDL
KAHRP$_{1-69}$	AQKQHEKKKK
KAHRP$_{1-69}$:GFP_QP2A	AAKQHEKKKK
PfEMP3$_{1-82}$:GFP_QP2A	AQVLGNTRLS
SS:CapProt:PfEMP3$_{63-82}$	AAVLGNTRLS
PFI1755c$_{1-61}$	SAAAAEEQDL
PFI1755c$_{1-61}$:GFP_mutant3	SAPVVEEQDA
PFI1755c$_{1-61}$:GFP_mutant5	SEPVVAAAAL
RESAwt	YGE
RESA_E6A	YGA
RESA_del.4	GE
RESA_del.5	YE
RESA_G5A	YAE

Not exported proteins	Sequence
truncated mTRAP (R)	SALYEHMNTK
REX1$^{60-79}$	NGTMLLNSIK
REX2$^{1-20}$scrambled	SSKGLALKFM
REX2$^{1-20}$scrambled+E7	SSKGLELKFM
STEVOR $^{44-63}$scrambled	YTNIEQIPIE
GBP130$^{87-106}$scrambled	AKTDTCAKKG
PiAvr3$^{48-67}$	KNEENEETSE
ETRAMP5$^{25-44}$	DQLDMGSVHN
SERA7$^{23-42}$	QEKPPPDSTV
SERA7$^{23-42}$Q25A	AEKPPPDSTV
PFF0090w$_{scr}$	LSFTHLHMGL
PF11_0505$_{scr}$	KEEETLVEKK
PFI1755c$_{1-61}$:GFP_SP1D	DEPVVEEQDL
PFI1755c$_{1-61}$_SP1Y_EP2G	YGPVVEEQDL
PFI1755c$_{1-61}$_SP1N_EP2G	NGPVVEEQDL
PFI1755c$_{1-61}$:GFP_ EP2R	SRPVVEEQDL
KAHRP$_{1-69}$:GFP_AP1D	DQKQHEKKKK
SS:CapProt:PfEMP363-82:GFP_AP1Y_QP2G	YGVLGNTRLS
PfEMP3$_{1-64}$:GFP	AQATGAGTAA
PFI1755c$_{1-61}$:GFP_mutant4	SAPVVAAAAL
PfEMP3$_{1-82}$:GFP_AP1D	DQVLGNTRLS
KAHRP SS + NTS-GFP	.DKIGQQVHD
KAHRP + NTS N_C-GFP	.QKQH
SERA5 + PEXEL –GFP	AQQGSTGASP
Ste$^{44-53}$$_{scr}$	YTNIEQIPI
SERA7Q23A$^{23-32}$	AEKPPPDST
neutral N-terminus	AGAGAGAGAG
SSG	SSGSSGSSGS

28 N-termini included

Exported proteins	Sequence
KAHRP + NTS N-GFP	AQKQHEHHHH
PfEMP3 wt	AQVLGNTRLS
KAHRP$_{1-70}$C7A	AQKQHEHAAA
KAHRP$_{1-70}$C10A	AQAQAEAAAA
KAHRP$_{1-70}$C12A	AQAAAAAAAA
GBP130 C10A	AEGAAAAAAA
SteQ45A_Q47A	AATAIHNPHY
SERA7Q23A$^{23-32}$_A25-28	AEAAAADSTV
AE-SERA7$^{33-40}$_A33-36	AEAAAAGDER
AE-SERA7$^{33-40}$_A36-39	AEGANAAAAR
AE-SERA7$^{33-40}$	AEGANPGDER

43 N-termini included

Mixed phenotype	Sequence
REX2$^{1-20}$scrambled+E3	SEKGLALKFM
PFI1755c$_{1-61}$:GFP_mutant5	SAPVVSTSTL
PF07_0007$_{scr}$	ANNTQTEAQP
SERA7Q23A$^{23-32}$_A27-30	AEKPAAAATV
REX2$^{1-10}$	KMYLAEIFS
AAS	AASAASAASA

6 N-termini included

%	export	no export	mixed
A	23,60	8,56	28,81
V	4,06	5,82	5,08
L	4,31	7,88	6,78
I	2,03	3,42	1,69
F	1,27	1,71	3,39
P	3,81	5,82	5,08
M	0,76	2,74	3,39
S	4,82	9,25	13,56
T	3,55	5,82	8,47
C	1,27	0,34	0,00
W	0,25	0,00	0,00
Y	3,30	2,40	1,69
N	4,82	3,77	3,39
Q	8,38	6,16	3,39
D	5,08	5,14	0,00

E	11,68	9,93	6,78
K	5,84	8,56	6,78
R	2,54	1,03	0,00
H	4,31	2,40	0,00
G	4,06	8,22	1,69

Presence of the respective amino acids in the N-termini of exported/not exported/partially exported proteins.

Exported proteins										
Position	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
A	58,1	23,3	16,3	23,3	20,9	16,3	14	18,6	14	16,3
V	0	0	9,3	9,3	9,3	6,98	9,3	2,33	0	2,33
L	0	2,33	2,33	18,6	9,3	0	0	0	9,3	9,3
I	0	0	0	2,33	6,98	0	4,65	2,33	2,33	0
F	0	0	0	0	2,33	0	4,65	2,33	2,33	0
P	0	2,33	14	2,33	4,65	9,3	0	16,3	9,3	0
M	0	2,33	0	0	2,33	0	0	0	2,33	0
S	16,3	2,33	0	2,33	2,33	0	2,33	2,33	2,33	14
T	2,33	0	14	6,98	2,33	4,65	9,3	0	2,33	2,33
C	2,33	0	2,33	0	0	2,33	4,65	0	0	0
W	0	0	0	2,33	0	0	0	0	0	0
Y	9,3	0	4,65	0	2,33	0	2,33	2,33	4,65	16,3
N	2,33	0	2,33	0	4,65	11,6	14	11,6	4,65	4,65
Q	2,33	27,9	2,33	20,9	0	4,65	4,65	9,3	2,33	2,33
D	0	6,98	0	4,65	4,65	6,98	2,33	9,3	9,3	2,33
E	2,33	27,9	9,3	4,65	2,33	25,6	14	2,33	11,6	6,98
K	2,33	0	9,3	2,33	4,65	2,33	4,65	9,3	6,98	11,6
R	0	0	0	0	2,33	0	0	9,3	4,65	6,98
H	0	0	2,33	0	9,3	6,98	4,65	2,33	11,6	2,33
G	2,33	4,65	11,6	0	9,3	2,33	4,65	0	0	2,33

Frequency of amino acids in the exact given position within the N-termini of known proteins, showing export.

Not exported proteins										
Position	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
A	23,1	7,14	7,14	0	7,14	11,1	11,1	7,41	11,1	3,7
V	0	0	7,14	17,9	17,9	0	3,7	11,1	0	7,41
L	3,85	0	7,14	10,7	10,7	11,1	7,41	0	7,41	22,2
I	0	0	0	7,14	0	0	7,41	0	7,41	0
F	0	0	3,57	0	0	0	0	3,7	7,41	0
P	0	0	17,9	7,14	7,14	7,41	0	3,7	0	3,7
M	0	3,57	0	3,57	3,57	0	3,7	3,7	0	7,41
S	23,1	14,3	0	3,57	7,14	0	7,41	14,8	11,1	18,5
T	0	3,57	7,14	7,14	7,14	3,7	7,41	7,41	11,1	0
C	0	0	0	0	0	3,7	0	0	0	0
W	0	0	0	0	0	0	0	0	0	0
Y	11,5	0	3,57	3,57	0	0	0	0	0	0
N	7,69	3,57	3,57	0	3,57	7,41	3,7	3,7	0	3,7
Q	3,85	25	3,57	10,7	0	7,41	3,7	14,8	0	0
D	15,4	3,57	0	7,14	0	0	7,41	0	14,8	3,7
E	0	14,3	7,14	7,14	7,14	29,6	18,5	3,7	0	7,41
K	11,5	3,57	28,6	0	0	0	7,41	14,8	14,8	14,8
R	0	3,57	0	0	0	0	0	7,41	0	0
H	0	0	0	0	14,3	3,7	3,7	0	7,41	0
G	0	17,9	3,57	14,3	14,3	11,1	7,41	3,7	7,41	7,41

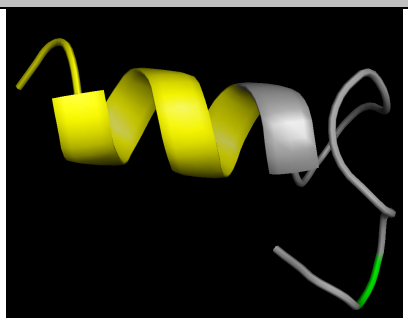
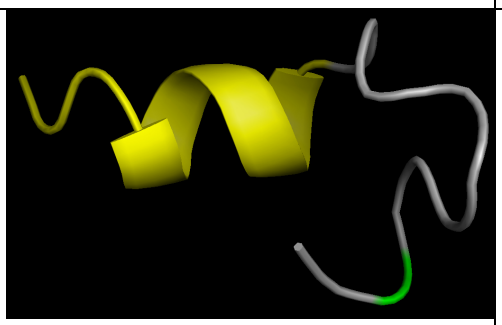
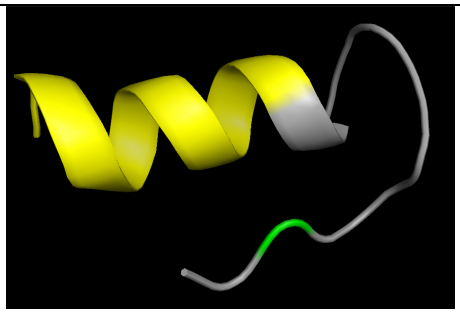
Frequency of amino acids in the exact given position within the N-termini of known proteins, showing no export.

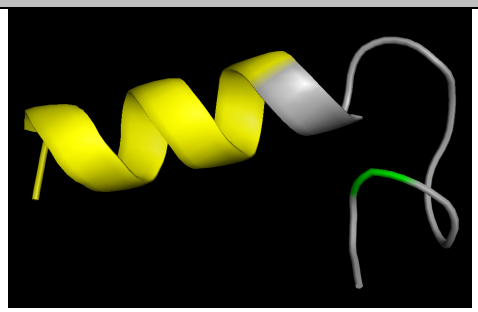
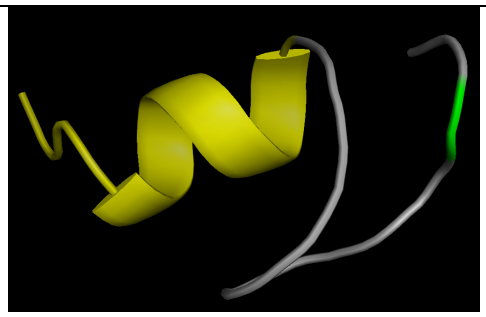
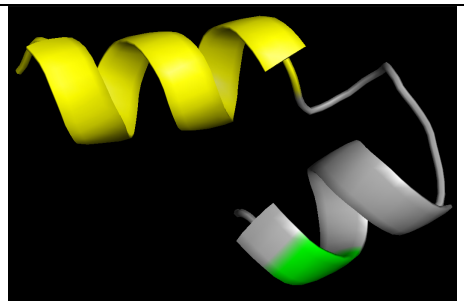
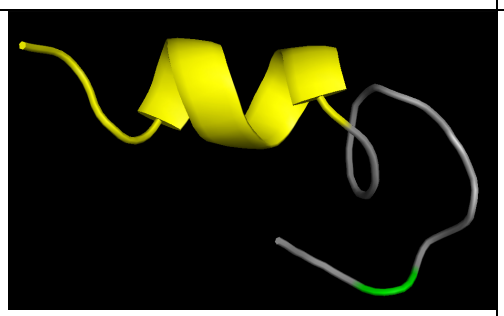
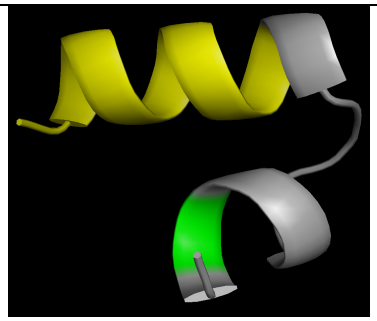
Proteins showing a mixed phenotype										
Position	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
A	50	33,3	0	16,7	50	33,3	33,3	50	0	20
V	0	0	0	16,7	16,7	0	0	0	0	20
L	0	0	0	16,7	16,7	0	16,7	0	0	20
I	0	0	0	0	0	0	16,7	0	0	0
F	0	0	0	0	0	0	0	16,7	16,7	0
P	0	0	16,7	16,7	0	0	0	0	0	20
M	0	16,7	0	0	0	0	0	0	0	20
S	33,3	0	16,7	0	0	33,3	0	16,7	33,3	0
T	0	0	0	16,7	0	16,7	16,7	0	33,3	0
C	0	0	0	0	0	0	0	0	0	0
W	0	0	0	0	0	0	0	0	0	0
Y	0	0	16,7	0	0	0	0	0	0	0
N	0	16,7	16,7	0	0	0	0	0	0	0
Q	0	0	0	0	16,7	0	0	0	16,7	0
D	0	0	0	0	0	0	0	0	0	0
E	0	33,3	0	0	0	16,7	16,7	0	0	0
K	16,7	0	33,3	0	0	0	0	16,7	0	0
R	0	0	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	0	0	0
G	0	0	0	16,7	0	0	0	0	0	0

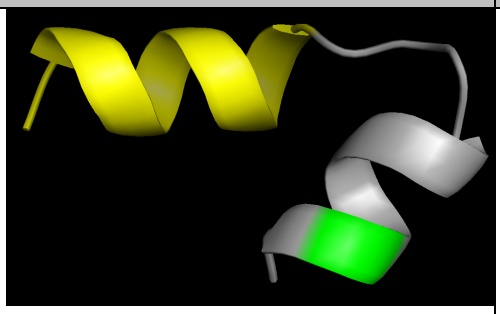
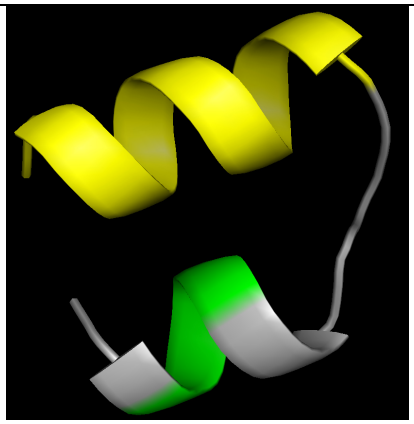
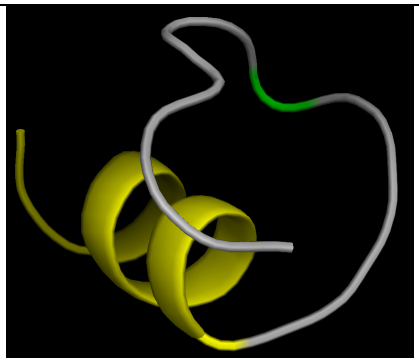
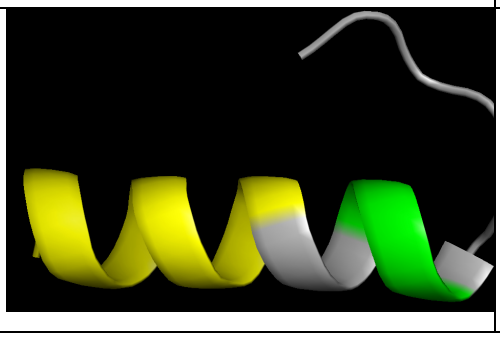
Frequency of amino acids in the exact given position within the N-termini of known proteins, showing a mixed phenotype. As only six N-termini are included, the statistic is not reliable.

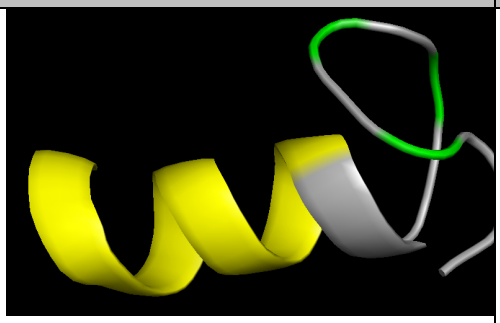
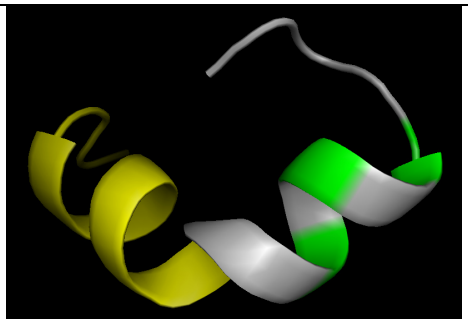
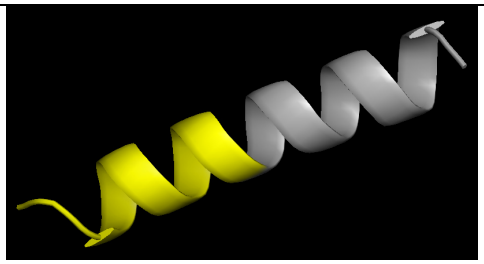
Prediction of the secondary structures of the investigated N-termini (using PEP-FOLD, Thevenet et al., 2012).

The name and sequence (left; amino acids are shown in one letter code) of the respective protein are followed by the predicted secondary structure (PEP-FOLD, Thevenet et al., 2012). The neutral N-termini are represented in grey, the inserted amino acids in green and the appended mTRAP reporter (first 10 amino acids) in yellow. On the right the secondary structure forming propensity of the inserted amino acids (Creighton, 1992) and the export promoting capacity are listed. ✓: high secondary structure forming propensity/high export rate, ~: mixed phenotype, ×: low secondary structure forming propensity/not exported.

Name	Sequence	Secondary structure prediction	Sec. structure forming propensity	expor-ted?
n_Q2	MA**Q**AGAGAGAG		✓	✓
n_D2	MA**D**AGAGAGAG		×	×
n_L2	MA**L**AGAGAGAG		✓	✓

Name	Sequence	Secondary structure prediction	Sec. structure forming propensity	expor-ted?
n_M2	MA**M**AGAGAGAG		✓	~
n_T2	MA**T**AGAGAGAG		✓	✓
n_S2	MA**S**AGAGAGAG		x	~
n_N2	MA**N**AGAGAGAG		x	~/ x
n_R2	MA**R**AGAGAGAG		x	~/ x

Name	Sequence	Secondary structure prediction	Sec. structure forming propensity	expor-ted?
n_A2	MA**A**AGAGAGAG		✓	✗
n_E3$^{A\text{-}flank}$	MA**AEA**GAGAG			~
n_E6	MAGAGA**E**AGAG		✓	✗
n_E6-8	MAGAGA**EEE**AG			~

Name	Sequence	Secondary structure prediction	Sec. structure forming propensity	expor-ted?
n_E3_E5_E7	MAG**E**G**E**G**E**GAG			~
n_E4_E6_E8	MAGA**E**A**E**A**E**AG			~
AAS	MAASAASAASA			✓

Acknowledgements

First of all I'd like to thank Dr. Tobias Spielmann for giving me the opportunity to conduct my PhD in his laboratory. Thank you for your supervision, your many stunning ideas, your great effort in always finding solutions, your permanent addressability and for amusing me in numerous moments, Tobi.
Further I would like to thank Prof. Dr. Thorsten Burmester for reviewing my thesis. I know you have little time, due to that an even bigger thanks for taking over this job. Prof. Dr. Iris Bruchhaus and Dr. Christina Deschermeier – thank you a lot for escorting me as my Co-Supervisors. You gave me beneficial feedback and I know that you have been there, when I asked for help.
Further I would like to thank Prof. Dr. Volker Heussler and Mariana for the great collaboration. The work with you and my time in Bern was amazing.

Very, very special props to the best lab group ever. It was amazing to have you Spielmänner around each day: Alex, Arlett, Bärbelchen, Horsti, Ferdi, Flo, Jakobus, Leo, Maria, Melissa, Nick, Sandra (it was a pleasure to have you as my master student), Svenni and Paolo. As this is technically only the half of my amazing group, my thanks goes likewise to the Gilbergers: Arne, Benthe, Boris, Dorle, Jaanis, Loulou, Maya, Micha, Sarah, Steffen, Tati and Tim. The work related input, the useful ideas and the many lovely discussions and exhilarating moments, but especially the solidarity and sympathy was exceptional. You often made my day but in particular you made my time here something very special.

A deep thanks to my proofreaders, who are – lucky me – also my friends. I appreciate your time and thank you a lot for helping me bringing this thing on the road. I am just happy to have you and the rest of my lovely ones around. I know you all supported me every time, no matter if near or far. A special thanks to Kathi. Thanks for absolutely everything, especially for always being there. And here always is really always.
And last but definitely not least möchte ich meiner Familie danken. Mama und Papa, ohne euch wäre das nicht möglich gewesen. Ich danke euch so sehr für alles – die vielen guten Gespräche, die Unterstützung, das Verständnis, das zu Hause. Genauso danke ich Oma, Opa, Chrissi und dem Rest meiner wunderbar verrückten Sippschaft.