Evaluation of the protozoan parasite *L. tarentolae* as a eukaryotic expression host in biomedical research

Inaugural-Dissertation

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

Submitted to the Department of Biology, Chemistry and Pharmacy

of Freie Universität Berlin

by

Stephan Klatt

born in Braunschweig, Germany

April 2013

I hereby declare that all experiments and writing contained within this thesis were conducted by myself, all references used are cited accordingly and any personal assistance has been acknowledged by name.

All experiments for this thesis were conducted from October 2008 to April 2013 in the group of Dr. Zoltán Konthur at the Department of Vertebrate Genomics under the supervision of Prof. Dr. Hans Lehrach at the Max Planck Institute for Molecular Genetics.

1. Reviewer: Prof. Dr. Hans Lehrach
Max Planck Institute for Molecular Genetics
Department of Vertebrate Genomics
Ihnestraße 63-73
14195 Berlin, Germany

2. Reviewer: Prof. Dr. Burghardt Wittig
Freie Universität Berlin
Molecular Biology and Bioinformatics
Arnimallee 22
14195 Berlin, Germany

Date of defence: August 7th 2013

Acknowledgement

I would like to thank Prof. Dr. Hans Lehrach, Director at the Max Planck Institute for Molecular Genetics, for giving me the great opportunity to conduct my dissertation in his department. Furthermore, I thank Prof. Dr. Burghardt Wittig from Freie Universität Berlin for his kindness to serve as an academic advisor of my thesis.

First and foremost, I would like to express my sincere gratitude to my direct supervisor Dr. Zoltán Konthur for his continuous support, great tutoring, advice and fruitful discussions as well as for the opportunity to travel, meet and collaborate with international scientists. Additionally, I would like to thank Dr. Sylvia Krobitsch for scientific conversations and all the members of the AG Konthur, current and previous, for their helpful hands and scientific exchange. In particular, Carola Stoschek for all her patience and support, Dr. Florian Rubelt, Yuliya Georgieva and Dejan Gagoski for the good working atmosphere, support and friendship.

My sincere thanks are also owed to our collaboration partners, Dr. Daniela Hartl and Dr. Daniel Kolarich. Furthermore, I thank Dr. Reinhard Breitling for valuable advice.

Moreover, I would like to thank my parents Rotraud and Friedhelm and my friends (for their support and patience). In particular, I want to thank Sina Nowak for proofreading parts of my thesis.

Last but not the least I would like to express my great gratitude to Carola, for her deep understanding, support and love.

Table of contents

Acknowledgement III

Table of contents IV

List of abbreviations VI

1 Introduction 9

1.1 General information about Leishmania species 11

1.1.1 The genus Leishmania 11

1.1.2 Life cycle and geographical distribution of Leishmania 12

1.1.3 Reproduction: Clonality versus sexuality 15

1.1.4 Infectivity and disease pattern of leishmaniasis 15

1.1.5 Immune defence and evolutionary acquired mimicry 16

1.1.6 Treatment and potential drug targets of Leishmania parasites 16

1.2 Leishmania tarentolae as biomedical expression host 17

1.2.1 Similarities and differences to human-pathogenic species 18

1.2.2 Advantages of the protozoan protein expression system 19

1.2.3 Cell Culture Conditions and the amastigote-like form 19

1.2.4 Posttranslational protein modifications 20

1.2.5 Gene expression and expression vector variants 22

2 Aim 25

3 Manuscripts 26

3.1 Manuscript I 28

3.2 Manuscript II 39

3.2 Manuscript III 54

3.3 Manuscript IV 63

4 Discussion 73

5 Current state of research and outlook 81

6 Summary 83

7 Zusammenfassung 85

8 References 88

List of abbreviations

2-D PAGE	Two-dimensional poly-acrylamide gel electrophoresis
AA	Amino-acid
AD	Alzheimer's disease
APP	Amyloid precursor protein
Aprt	Adenine-phosphorybosyl transferase gene
BirA	Biotin ligase
bp	Base pair
CAT	Chloramphenicol acetyl transferase
CHO	Chinese Hamster Ovary
CL	Cutaneous Leishmaniasis
CMP	Cytidine monophosphate
CO_2	Carbon dioxide
CPB	Cysteine protease B
CRP	C-reactive protein
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
eGFP	Enhanced green fluorescent protein
EPO	Erythropoietin
FDFT1	Farnesyl-diphosphate farnesyltransferase
GPI	Glycosylphosphatidylinositol
His_6-tag	Hexahistidine-tag
Hsp	Heat shock protein

IFN-γ	Interferon-gamma
Ig	Immunoglobulin
JBS	Jena Bioscience
L. tarentolae	*Leishmania tarentolae*
LEXSY BHI	Leishmania Expression System, Brain-Heart-infusion based medium
LmjF	*Leishmania major* strain Friedlin
LPG	Lipophosphoglycan
LtaP	*Leishmania tarentolae* strain Parrot-TarII
ML	Mucocutaneous Leishmaniasis
MLEE	Multilocus enzyme electrophoresis
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
mTKIN	Cysteine-free mutant of kinesin
N	Asparagine
NW	New World
Odc	Ornithin decarboxylase locus
OW	Old World
P	Proline
p53	Cellular tumour protein p53
PC4	Serine protease proprotein convertase 4
PNGase F	Peptide -N-Glycosidase F
PSA	Promastigote surface antigen protein
PTM	Posttranslational protein modification

Rab GTPase	Ras related protein guanosine triphosphate enzyme
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
S	Serine
S1	Security level 1
S2	Security level 2
sAP1	Secreted acid phosphatase 1 gene
sAPPalpha	Soluble amyloid precursor protein alpha
scFv	Single chain fragment variable
SITS	Species-independent translational sequences
SOD1	Cu/Zn Superoxide dismutase
SORLA	Sorting protein-related receptor containing LDLR class A repeats
SP	Secretory signal peptide
Ssu	18S ribosomal RNA locus
T	Threonine
TEM	Transmission electron microscopy
TET repressor	Tetracycline-controlled transcriptional activation repressor
tPA	Tissue plasminogen activator
TS	Trans-Sialidase
UTR	Untranslated region
VL	Visceral Leishmaniasis
WHO	World Health Organization

1 Introduction

The successful use of recombinant proteins as therapeutic agents in human medicine revolutionized biotechnological research in the 1980s [1]. Today, more than 25 years later, recombinant proteins are frequently used, for example, as a treatment of cancer, anaemia and infertility [2], or they form the basis of many diagnostics. The human genome consists of around 23.000 protein-coding genes [3, 4]. The number of different proteins, however, is significantly higher, since alternative mRNA splicing events and posttranslational protein modifications (PTMs) have led to many different proteins emerging from one single gene. For instance, proteins form structural cell elements, transport diverse macromolecules, regulate cellular homeostasis, tag external antigens for primary immune response or are secreted into the extracellular environment.

The recombinant therapeutic proteins which are most frequently approved are monoclonal antibodies, proteins with enzymatic activity or generally glycosylated proteins [5-7]. For their production, genetically engineered bacterial organisms, fungi, or mammalian cells are used, which are generally able to correctly fold the protein and synthetize it into a biologically active form [2]. Common pitfalls of many heterologous expression systems are the lack of correct PTM. Here, protein glycosylation plays a major role as more than 50% of all human proteins are glycosylated [8]. Importantly, many proteins need the glycosylation for their biological activity [9, 10]. Prokaryotic systems are not able to glycosylate human recombinant proteins. Yeasts, insect cells, or herbal hosts can glycosylate, but their glycosylation profile is different to the human counterpart, which can e.g. negatively influence protein functionality or the half-life [11]. The expression of human recombinant proteins in mammalian cells, such as CHO-cells (Chinese Hamster Ovary), achieves the highest homogeneity in respect to PTMs [12]. As the need of recombinant proteins for therapeutic or diagnostic purposes is continuously increasing, there is a great demand for expression systems, which can on the one hand offer mammalian-like PTMs and on the other hand result in high protein yield.

In general, the expression system of the eukaryotic parasite *Leishmania tarentolae* is such a system [13]. *L. tarentolae* represents a competitive alternative to the mammalian cell lines for the following reasons: First, the parasite is not pathogenic to humans [14], it is relatively easy to cultivate [15, 16], and, it offers satisfying recombinant protein yields (0.1-30mg/L) [17]. Additionally, the expression system can easily be extended to an industrial production scale.

Second, with the expression of the human hormone erythropoietin (EPO), high glycosylation pattern similarity to the human-derived counterpart could be shown [18], making *L. tarentolae* generally attractive for the expression of human glycoproteins. Third, the parasite has the ability of masking itself through the uptake and surface presentation of host-derived molecules, enabling reproduction and survival inside the insect and saurian host. This evolutionary acquired mimicry can be biotechnologically utilized (subchapter 1.1.5).

To date, *L. tarentolae* has been used as vaccine candidate against human leishmaniasis [19, 20], to screen for antileishmanial drugs [21], to study RNA editing [22], gene amplification [23, 24] as well as for the expression of eukaryotic recombinant proteins. Overall, this broad spectrum of applications also suggests *L. tarentolae* to be a highly qualified expression system. Especially recombinant protein production is a growing application area [17, 25]. Worldwide, more than 150 laboratories in academia and industry are working with *L. tarentolae* as a protein expression host (Dr. Breitling, personal communication). This number is likely to increase in the future, since in 2012, the genome of *L. tarentolae* (Parrot-TarII strain) was successfully sequenced [26].

Utilizing *L. tarentolae* as a biotechnological host for human recombinant protein expression requires the understanding of its life cycle and metabolic adaptations. Unfortunately, most research does not address the biology of *L. tarentolae* or other saurian Leishmania species, but human-pathogenic ones like *L. major* or *L. donovani*. For this, the whole genus of Leishmania is taken into consideration. In chapter 1.1, general information about the genus Leishmania including its life cycle, geographical distribution, reproduction, infectivity, disease pattern, immune defence, treatment and potential drug targets are given. On the other hand, chapter 1.2 addresses the advantages of *L. tarentolae* as a biomedical expression host, its cell culture conditions, studied PTMs with expressed protein examples, gene expression and established expression vector variants.

To sum up, the understanding of the genus Leishmania and examples of successfully expressed recombinant proteins provide the background of my PhD thesis. Taking this knowledge into consideration and expression of selected human recombinant proteins with a biomedical background should answer the main goal of my thesis if *L. tarentolae* is capable of producing human proteins of therapeutic standard or rather of high homogeneity to the appropriate counterpart.

1.1 General information about Leishmania species

1.1.1 The genus Leishmania

The genus Leishmania belongs to the class of Kinetoplastida [27] and to the order of Trypanosomatida, which diverged very early in the eukaryotic lineage [28]. Since its first description by W.B. Leishman [29], C. Donovan and J.H. Wright [30] in 1903/04, around thirty different species are assigned today infecting different vertebrates such as reptiles, dogs or man. Today, the reference technique for the identification of a Leishmania species is multilocus enzyme electrophoresis (MLEE, [31, 32]).

The genus of Leishmania can be divided into the three subgenera Leishmania, Viannia and Sauroleishmania [33]. Furthermore, species within the subgenera are grouped to complexes according to biological and biochemical characteristics [34]. To the subgenus Leishmania belong 15 species grouped in 4 complexes. The subgenus Viannia consists of 9 species, where not all species can be grouped. Species of the subgenus Sauroleishmania are not grouped at all. Around 10 species, including *L. tarentolae*, belong to this subgenus [35], but only three are well studied. Most Sauroleishmania species only infect different reptile species like lizards [35]. According to the MLEE method, a group of five Leishmania species cannot be assigned to any of the subgenera (e.g. *L. hertigi* or *L. colombiensis* [33]). A simplified presentation of the genus Leishmania is given in *Figure 1*, which is not a phylogenetic tree.

However, classification of Leishmania species into a taxonomy has never been absolutely certain, neither in the past nor today. Recent studies on the analysis of the mitochondrial cytochrome b gene [36] or the heat-shock protein (hsp) 70 gene subfamily [37, 38] have reopened the debate on the classification of the phylogenetic tree of Leishmania species, including the grouping into complexes [33].

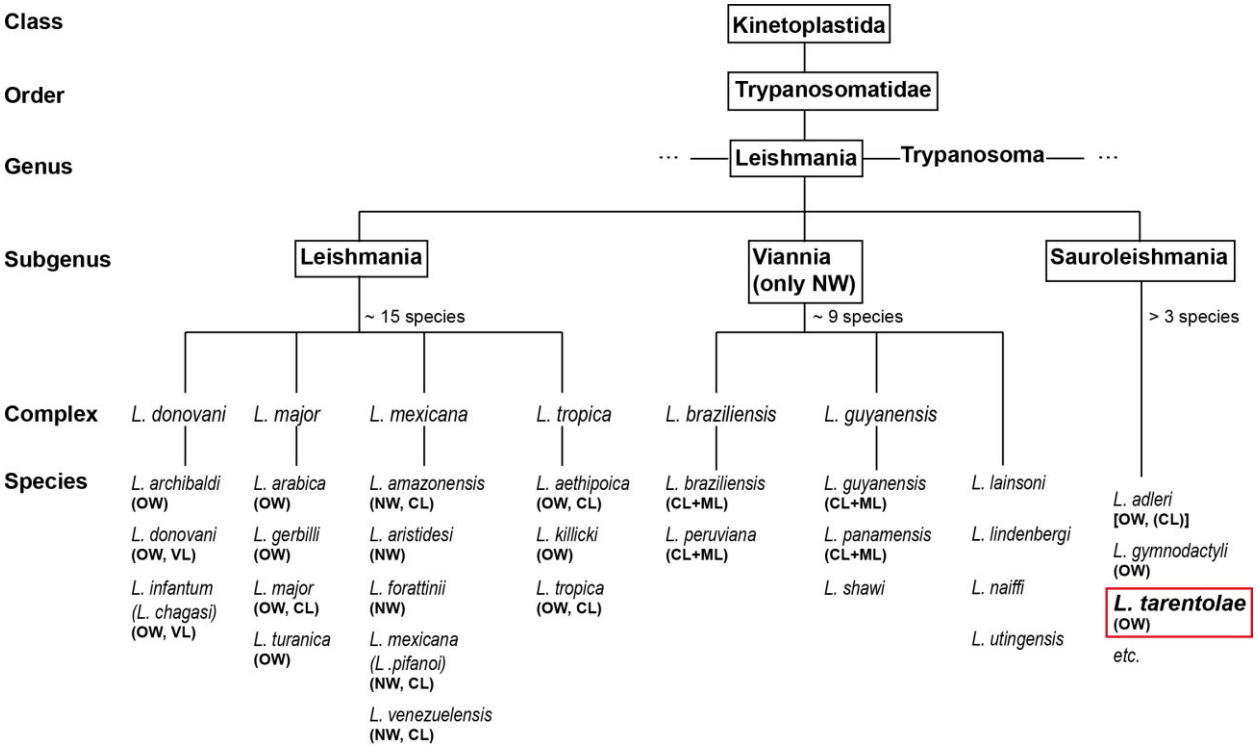

Figure 1: Simplified presentation of the genus Leishmania (no phylogenetic tree). The genus Leishmania consists of the three subgenera Leishmania, Viannia and Sauroleishmania. Furthermore, the subgenera can be divided into complexes. Leishmania species have their natural distribution in the Old World (OW) and New World (NW). In humans, they can cause cutaneous leishmaniasis (CL), mucocutaneous leishmaniasis (ML) or visceral leishmaniasis (VL). L. tarentolae is highlighted in red.

1.1.2 Life cycle and geographical distribution of Leishmania

The life cycle of Leishmania species [39] alternates between promastigotes in insect and amastigotes in vertebrate hosts [40]. Their shape ranges from round-like to lance like structure with a length of 4-12µm and a width of 0.5-3µm. Furthermore, they possess a relatively large nucleus and one single mitochondrion, where mitochondrial DNA is found locally condensed and named kinetoplast. A simplified overview of the two main morphological forms is given in *Figure 2 A+B*.

In more detail, Phlebotomine sand flies of the genera Phlebotomus, Lutzomyia and Sergentomyia [35, 41] are infected by taking up a blood meal, where assimilated amastigotes change into procyclic, motile and flagellated promastigotes colonizing the midgut and/or the

hindgut of the insect vector [34]. Providing that the parasite migrates back to the pharynx and the proboscis of the sand fly, they are retransmitted during a blood meal to different vertebrate hosts as non-dividing metacyclic promastigotes. Here, they proliferate as non-motile and 'rudimentary' flagellated amastigotes [42] inside specific cells of the immune system. Finally, infected immune cells burst and the released amastigotes are phagocytized by new immune cells or they are taken up by a sand fly restarting the life cycle. For *L. tarentolae*, the amastigote form is still under debate. However, an amastigote-like form was described (see subchapter 1.2.3). A simplified life cycle of *L. tarentolae* (including promastigotes and amastigote-like forms) is illustrated in *Figure 2C*.

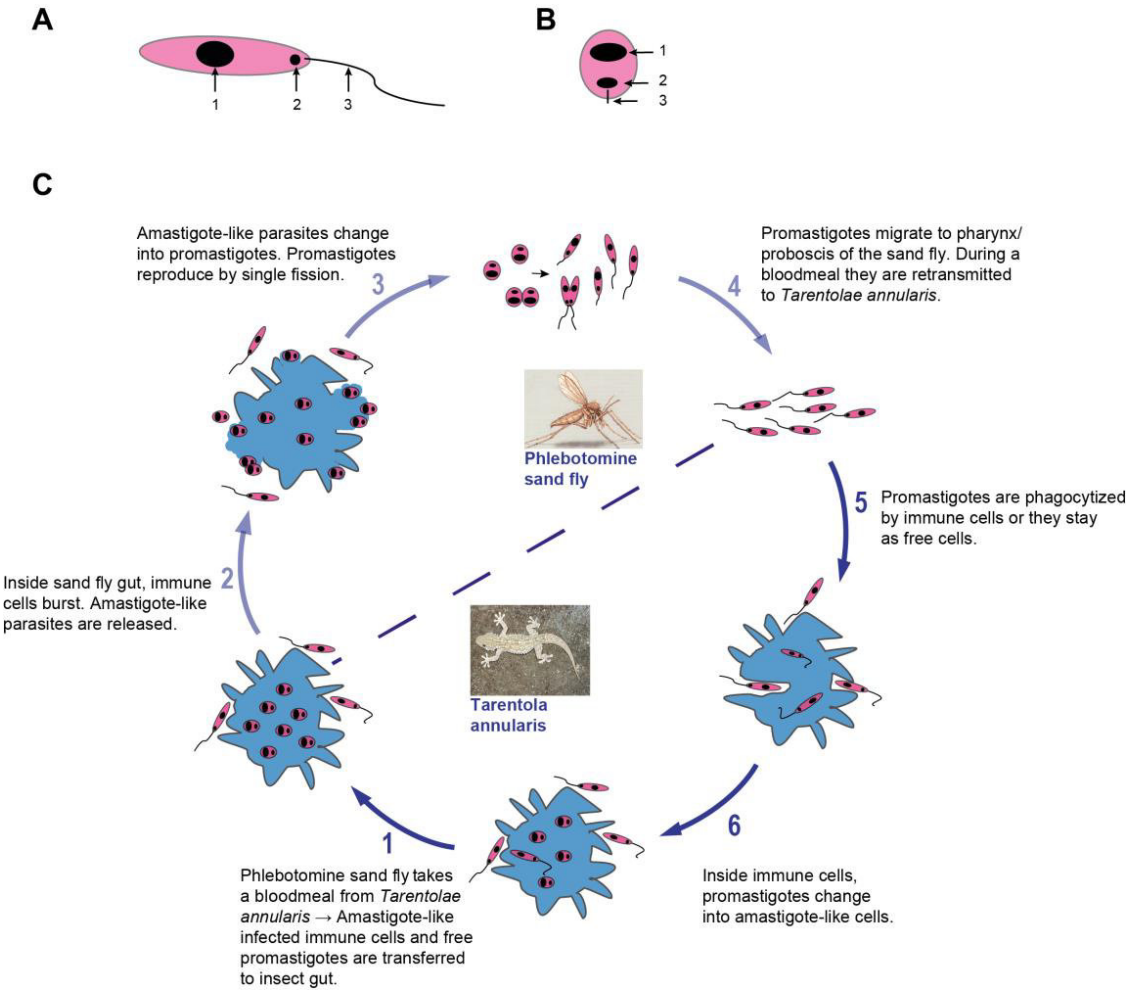

Figure 2: Simplified life cycle of L. tarentolae with main morphological forms. (A) Promastigote and (B) amastigote-like form with the nucleus (1), kinetoplast (2) and flagella (3). (C) Alternating life cycle between sand flies (image from WHO) and the gecko T. annularis (image from JBS).

The geographical distribution of Leishmania is closely linked to its insect vector. Phlebotomine sand flies as well as human leishmaniasis can be found nearly all over the world (*Figure 3*, blue highlighted). Sand flies of the genus Phlebotomus and Sergentomyia have their natural habitat in Southern Europe, Africa and Asia (Middle East, India) [43]. This region is summarized as the Old World (OW). In the New World (NW; Middle- and South-America), the genus of Lutzomyia has its natural habitat [43]. The natural habitat of *L. tarentolae* is located in southern Europe and North Africa (*Figure 3*, red highlighted). Here, the parasite is often found in lizards like *Tarentola annularis/mauritanica*. For a long time, Australia was said to be free of Leishmania. This presumption had to be disproved in 2004, where cutaneous leishmaniasis was described in red kangaroos [44]. In contrast to all other endemic regions, day-feeding midges (Forcipomyia, subgenus Lasiohelea) instead of Phlebotomine sand flies were found as new vector for Leishmania [45]. This newly described insect vector and leishmaniasis in Australia is disregarded here.

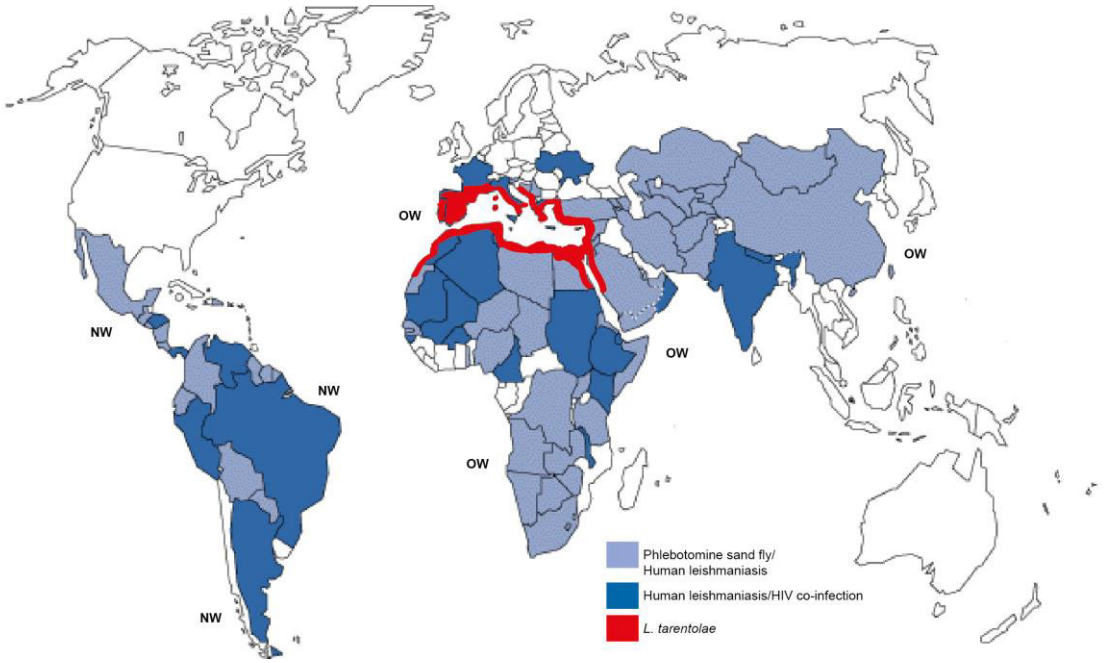

Figure 3: Phlebotomine sand flies/Human leishmaniasis (light blue), HIV-coinfection (dark blue) and geographical distribution of L. tarentolae (red). OW: Old World; NW: New World. Image modified from WHO.

1.1.3 Reproduction: Clonality versus sexuality

Even today the mode of reproduction of Leishmania parasites is still under debate. Two major hypotheses are postulated: Clonality versus sexuality [46]. Underpinned by several independent studies and tests for heterozygosity, recombinant or identical genotypes etc., the preferred hypothesis is clonal reproduction by single or multiple fission [47-49]. Although there has been little evidence for sexual recombination of Leishmania parasites in the past [50, 51], some new studies verify its existence. For example, genetic recombination was experimentally verified and a sexual cycle of Leishmania inside the insect vector consistent with meiotic processes was proposed [52]. Furthermore, strong homozygosities were shown for a Leishmania population, being incompatible with a predominantly clonal mode of reproduction at an ecological time scale [53]. Summarized, reproduction of Leishmania parasites (at least of *L. braziliensis*) works most likely as follows: Clonal reproduction of Leishmania is found in the vertebrate host as well as in the insect vector. In contrast, frequent recombination events between different individuals are limited to the insect vector [46]. Further support comes from the close relative *Trypanosoma brucei*, where this model has been verified [54, 55].

1.1.4 Infectivity and disease pattern of leishmaniasis

Leishmania cause a wide spectrum of diseases and chronic infections in different vertebrate hosts [27]. The disease is called leishmaniasis. According to the World Health Organization (WHO) (http://www.who.int/leishmaniasis), around 12 million people are suffering from it with 1.5-2 million new infections per year (mortality rate 3.5%). The disease pattern differs between species and is depending on other factors, such as the immunity of the infected host [56]. There are three distinguishable disease patterns: Depending on the Leishmania species, the disease can cause locally restricted and in most cases self-curing cutaneous lesions. This disease pattern is called cutaneous leishmaniasis (CL) [57-59]. In contrast, when facial mucosal tissue is infected, the disease is called mucocutaneous leishmaniasis (ML) [58, 59]. Visceral leishmaniasis (VL) [60, 61] is the most dangerous form, where the parasites proliferate in organs like the liver, spleen or the bone marrow. In many cases, this form of leishmaniasis is fatal. Whether the clinical manifestation will be CL, ML or VL strongly depends on the species capabilities [56], such as temperature sensitivity [62], immunoevasion or tissue tropism [63].

Transmission of leishmaniasis to men is not restricted to infected sand flies, but can also be transmitted through the contact with open wounds or mucous membranes of infected men or

domestic animals like dogs. In endemic regions, dogs and rodents are the main vertebrate reservoir of Leishmania parasites. Luckily, the vaccine CaniLeish against canine leishmaniasis was permitted in the European Union in 2011 [64, 65], appreciably reducing infectiousness.

1.1.5 Immune defence and evolutionary acquired mimicry

In general, sand fly transmitted metacyclic promastigotes are immediately recognized and phagocytized mainly by macrophages [66-68] and less frequently by dendritic cells, neutrophils and fibroblasts [69]. Alternatively, promastigotes of some species can actively enter these immune cells [56]. Inside immune cells, parasites pass into acidic phagolysosomes (lysosome-like compartment) or into parasitophorus vacuoles. Subsequent elimination of the invaders cannot be completed due to evolutionary acquired mimicry. In other words: Phagocytized Leishmania parasites undergo morphological and metabolic adaptations ensuring their survival [66, 69, 70]. For example, they extremely shorten their flagella [42] or they temporarily pause phagolysosome biogenesis of the host by transferring the surface-anchored glycolipid lipophosphoglycan (LPG) to the membrane of macrophages [71-74] to prepare differentiation into the amastigote form. In the end, they lack prominent surface glycocalyx of glycosylphosphatidylinositol (GPI) proteins or other GPI-anchored macromolecules [75, 76]. The arising acidic pH (4.5-5.5) in combination with the body temperature (33-37°C) is the signal-giving impulse for differentiation [77]. As another example, the parasite shifts its main source of metabolic energy from glucose to fatty acid oxidation, or uses glycerol and amino acids as precursors for the synthesis of sugars due to its exogenous lack [78, 79]. However, the uptake and cell surface presentation of sialoglycans has been reported [80, 81]. These terminal sugar residues are for example involved in cell-cell interactions or regulations of the hosts' immune response [82, 83]. Presenting host-derived sialoglycans on the parasites glycocalyx could partially explain the mimicry or pathogenicity, but it remains to be investigated whether sialic acids play a role for intracellular survival [81]. Summarized, Leishmania parasites can successfully survive inside the vertebrate host by metabolic adaptations and evolutionary acquired mimicry.

1.1.6 Treatment and potential drug targets of Leishmania parasites

The easiest way to diagnose leishmaniasis is probing potential material (from skin lesion, peripheral blood or the liver) and analysing it via microscopic techniques. Amastigotes can be

easily detected by probe staining with e.g. Giemsa [56]. Common drugs to treat a Leishmania infection are pentavalent antimonials [59, 84], which inhibit the fatty acid oxidative and glycolytic pathways of amastigotes [85]. Increasing resistances gave rise to the development of new drugs like amphotericin B and its lipid formulations, oral miltefosine or intravenous paromomycin [84, 85]. Alternatively, thermotherapy (e.g. hot water bath or infrared light) can be used to treat CL [62].

Leishmania species possess several different cell organelles and metabolites which do not exist or exist in an altered form in vertebrates. Potential drug targets are for example glycosomes, which are related to peroxisomes and house the first steps of glycolysis [86, 87]. Other targets are the acidocalcisomes, special acidic calcium-storage organelles [88, 89], or the flagella with the flagellar pocket, the region of protein secretion or endocytosis [90, 91]. Leishmania only possess one single mitochondrion with a DNA-dense region, the kinetoplast. This kinetoplast and its replication involve many special proteins offering potential drug targets [92, 93]. Inhibition of the GPI-biosynthesis is another drug developmental approach [94]. Many glycoproteins of the glycocalyx are membrane-linked via such a GPI-anchor.

1.2 *Leishmania tarentolae* as biomedical expression host

As described above, collected knowledge and the fundamental understanding of the biology of Leishmania species is essential for the development of new drugs and vaccines against leishmaniasis. Unfortunately, research on human-pathogenic species is complex and expensive due to the risk of human infection and the necessity of adequate laboratory equipment. As a good alternative, non-human-pathogenic species like *L. tarentolae* can be studied, offering a very similar biology and metabolism and being only security level 1 (S1). *L. tarentolae* was originally isolated from the lizard *Tarentola annularis* in 1921 [14], and is the best studied saurian-pathogenic species today (most publications in PubMed). Other studied saurian species are for example *L. adleri* [95] or *L. gymnodactyli* [96]. Like human-pathogenic species, *L. tarentolae* has been used as vaccine candidate against leishmaniasis [19, 20]. Furthermore, the parasite is used to study RNA editing [22] and gene amplification [23, 24]. However, during the last decade research is more and more concentrating on its use as an expression host for recombinant eukaryotic proteins. A few studies could show high PTM homogeneity to human counterparts, verifying the great ability of *L. tarentolae* for biomedical applications [17, 25]. In 2012, the genome of *L. tarentolae* (Parrot-TarII strain) was successfully sequenced [26],

increasing the number of sequenced Leishmania species to six [97, 98]. Genomes of all sequenced Leishmania species are available at the online database TriTrypDB (www.tritrypdb.org). The availability of the genome will improve the understanding of the subgenus Sauroleishmania and enable to find similarities and differences to human-pathogenic species.

1.2.1 Similarities and differences to human-pathogenic species

Compared to the genomes of *L. major* or *L. infantum,* the genome of *L. tarentolae* consists of the same amount of chromosomes but the genome is smaller in size (Table 1). However, >90% of the gene contents are shared with human-pathogenic species. Hence, the important question is: What are the differences of the remaining ~10%? The analyses of the genome predict 95 coding sequences to be unique to *L. tarentolae*. Furthermore, 250 genes found in human-pathogenic species completely lack in the lizard parasite. Many of these 250 genes are preferentially expressed in the amastigote stage of human-pathogenic species. Although promastigotes of *L. tarentolae* are able to enter human phagocytic cells and differentiate into amastigote-like forms, there is no clear evidence for their replication within macrophages [19-21]. *L. tarentolae* possesses several virulence factors like cysteine protease B (CPB), lipophosphoglycan LPG3 and the leishmanolysin GP63 [99], but it lacks the major virulence factor A2. A2 has been proved to play a major role in parasite virulence and visceralization capability [100]. Additionally, development of amastigotes inside lizards is still debated [101, 102]. Underrepresented genes and their products affect antioxidant defence or parts of the secretory pathway, e.g. vesicular-mediated protein transport. In detail, several adaptins are missing, which are involved in the formation of clathrin-associated adaptor protein complexes necessary for the formation of transport vesicles within the secretory pathway [103]. Other missing genes are coding for glycosyltransferases (see genome of *L. major*, genes LmjF35.5250 and LmjF29.2110) or the endoplasmic reticulum chaperone calreticulin (LmjF31.2600) ensuring the folding and quality control of nascent secretory proteins [104]. In contrast, two gene families are enriched in *L. tarentolae* promastigotes- the already mentioned leishmanolysin (GP63), a surface glyco-metalloprotease [105, 106], and promastigote surface antigen proteins (PSA).

Summarized, the gene content difference could partially explain why *L. tarentolae* is not able to survive in human macrophages, why they are mostly found as free parasites in the insect vector

and why their amastigote form is still under debate. They seem to be better adapted to the promastigote (insect) stage.

Table 1: Genome size of L. tarentolae compared to other Leishmania species

Species	Number of chromosomes	Genome size (bp)
L. tarentolae	36	30.44 million
L. major	36	32.82 million
L. infantum	36	32.13 million

1.2.2 Advantages of the protozoan protein expression system

There is a wide range of protein expression systems, each having its niche and fulfilling favoured requirements. For example, prokaryotic systems like the one of *E. coli* are world standard and used for many different applications [107, 108]. In relation to the expression of eukaryotic proteins they frequently entail drawbacks, such as improper protein folding, formation of inclusion bodies, or the lack of PTMs [109-111]. In contrast, eukaryotic systems offer PTMs, and in general, more proteins are expressed in their soluble form. However, protein yield is usually lower in eukaryotic expression systems as compared to prokaryotic ones [1, 112]. Thus, a heterologous protein expression system should offer both- PTM's and high yield. The ability of the eukaryotic parasite *L. tarentolae* to express complex eukaryotic proteins with highly diverse PTMs in combination with its easy, bacteria-like handling makes the parasite a promising expression system for eukaryotic recombinant and biomedically relevant proteins [15, 16, 113, 114].

1.2.3 Cell Culture Conditions and the amastigote-like form

In cell culture as well as in the insect host, *L. tarentolae* is found in its promastigote form. In general, the parasite has to be cultivated in static or agitated suspension culture under aerobic conditions in the dark at 26°C. This is the optimal evaluated growth temperature and also occurring in the sand fly host. In contrast, the body temperature in geckos is quite variable [115]. For growth, different complex and synthetic media are available for selection [114]. In this study, the complex medium LEXSY BHI developed by JBS (Cat-No. ML-411) was chosen. It is serum-

free and has to be enriched with the iron-containing porphyrin Hemin (porcine Hemin), which is essential for promastigote growth and survival. After gene transfection by electroporation, single transfectants can be normally picked from Agar-plates after 5-10 days. Consequently, up to 2-3 weeks pass by before the potential expressed recombinant protein can be analysed. Compared to other expression hosts, this takes considerably time and is a disadvantage. For a detailed description of *L. tarentolae* cultivation and clonal selection please go to the method section of Manuscript I.

The occurrence of *L. tarentolae* amastigotes is still under debate. Past experiments tried to trigger the differentiation of promastigotes to amastigotes from other Leishmania species by mimicking phagolysosomal conditions [116, 117]. In detail, the temperature was raised to 37°C, the pH was acidified to 5.5 or 4.5 and in some cases, parasites were additionally cultivated in a CO_2 incubator with 5-7% CO_2. Analyses of typical expressed amastigote proteins and gene expression profile studies revealed slight differences between isolated and induced amastigotes [118, 119]. Based on those differences, induced amastigotes are referred to as amastigote-like form today [119]. Induction of amastigote-like forms of *L. tarentolae* has also been successfully performed (for strain LEM-125, not for the sequenced strain TarII) [21].

1.2.4 Posttranslational protein modifications

A major advantage of the protein expression system of *L. tarentolae* is its principal capacity to fold and posttranslationally modify eukaryotic proteins in a mammalian-characteristic manner and to solubly express them, enabling medical and biotechnological applications. In general, many PTMs are important for the biological activity, stability and immunological property of the protein.

Among the plethora of PTMs, protein glycosylation must be emphasised. Secretory and membrane proteins of higher eukaryotes are highly glycosylated, being modified within the secretory pathway [120]. Altogether, more than 50% of all human proteins are glycosylated [8]. Leishmania are also specifically rich in glycoproteins [121]. Glycosylation is a quality control mechanism for the folding status of proteins [122]. Compared to its non-glycosylated counterparts, glycosylated proteins show a higher stability [123]. Glycans protect their proteins against enzymatic degradation and through the presentation of specific sugar epitopes they can regulate the interaction with cell receptors. In contrast, cytoplasmic or nuclear proteins are rarely glycosylated [124, 125]. Furthermore, terminal sugar residues like sialic acids

(N-acetylneuraminic acids) directly regulate the immune response and they directly influence the serum half-life of glycoproteins e.g. immunoglobulins (Ig) [11]. The general glycosylation characteristics of *L. tarentolae* are very close to the one of mammals (Figure 4) [18], but are limited due to the absence of enzymes for the sialic acid biosynthesis pathway [126].

Surprisingly, terminal sialic acids (α2,3- and α2,6-linked) were found on endogenous membrane-glycoproteins of *L. major* and *L. donovani* [80, 81, 127]. To date, the mechanism behind the uptake of free sialic acid glycans is unknown. However, a CMP-sialic acid transporter in the Golgi-apparatus membrane was reported, transporting sialic acids from the cytoplasm into the Golgi [80, 81]. General uptake of sialic acids is probably occurring through endocytosis.

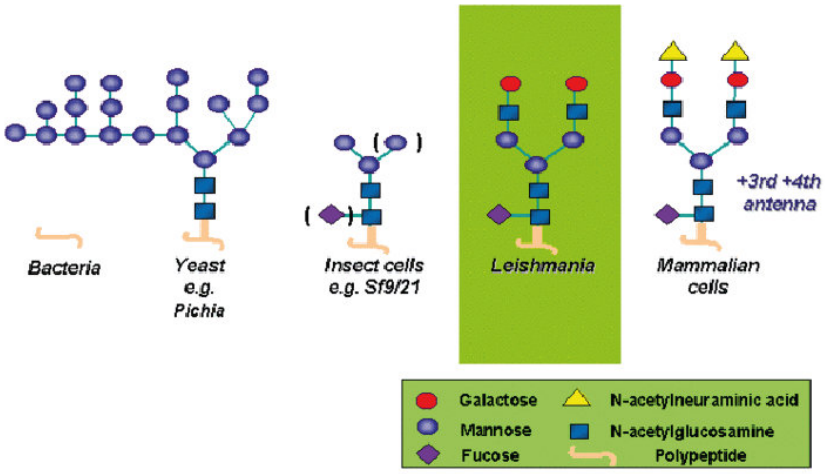

Figure 4: N-glycosylation pattern of L. tarentolae compared to different organisms. Compared to bacteria, yeasts and insect cells, L. tarentolae is offering the highest homogeneity to the mammalian counterpart, only missing N-acetylneuraminic acids. *Image from JBS, LEXSY Essence.*

Beside glycosylation, phosphorylation is another important PTM. Studies of mammalian cells suggest that as many as one-third of all cellular proteins are covalently modified by protein phosphorylation [128]. Protein phosphorylation is often a reversible process, and many important signaling pathways and cellular processes are regulated by phosphorylation/dephosphorylation events [129]. Further PTMs are for example disulphide bonds [130], acetylations [131] or the rare protein biotinylation [132].

Here are some examples of analysed posttranslationally modified and medically relevant proteins expressed with *L. tarentolae*: The glycoprotein hormone EPO expressed with *L. tarentolae*, was shown to be biologically active, natively processed at the N-terminus, and N-glycosylated [18]. In detail, a mammalian-type biantennary fully galactosylated, core-α-1,6-fucosylated N-glycan was described. EPO is necessary for erythrocyte production and often used in hemic diseases. With the glycoprotein heterotrimeric laminin LM-322 expressed in *L. tarentolae*, similar cell adhesion activity compared to mammalian-derived LM-322 could be successfully shown, indicating its proper folding and assembly with the formation of disulphide bridges [9, 133]. Although not described for *L. tarentolae*, phosphorylation of the human recombinant cellular tumour protein p53 was successfully demonstrated in an avirulent strain of *L. donovani* [134], indicating that this outcome should be repeatable with *L. tarentolae* due to their close relationship. Unfortunately, not many publications address PTMs of recombinant proteins expressed in *L. tarentolae* and other Leishmania species. In terms of further expressed glycoproteins, a few publications show their functionality without analysing the glycosylation pattern [10, 133]. The absolute majority is analysing endogenous proteins and metabolism changes during transition from pro- to amastigotes and infection.

1.2.5 Gene expression and expression vector variants

The expression system of *L. tarentolae* is the only commercialized protozoan protein expression system, distributed by JBS and known as LEXSY (http://www.jenabioscience.com/) [135]. Developed constitutive [18] and inducible-integrative [15] expression vectors enable the production of intracellular or secretory recombinant proteins. By homologous recombination (double cross-over), the expression cassette is either inserted into the chromosomal 18S ribosomal RNA locus (ssu; constitutive system) or into the ornithin decarboxylase locus (odc; inducible-integrative system) of *L. tarentolae*. Ssu is a repetitive locus of the *L. tarentolae* genome with high rates of transcription by the host RNA polymerase I [18, 136], whereas the odc locus is transcribed by RNA polymerase II. Gene expression regulation by RNA polymerase I is a unique attribute of *L. tarentolae* and all trypanosomatids. In all other organisms, gene expression is regulated by RNA polymerase II [137]. Furthermore, transcription in *L. tarentolae* is polycistronic, there are no introns and, therefore, no *cis*-splicing reactions [28]. Posttranscriptional processing of pre-mRNA is effected by *trans*-splicing reactions [138] and polyadenylation within intergenic regions. Protein expression regulation may be influenced by

the structure of these intergenic regions [28, 137]. An overview of the transcription process is given in *Figure 5*.

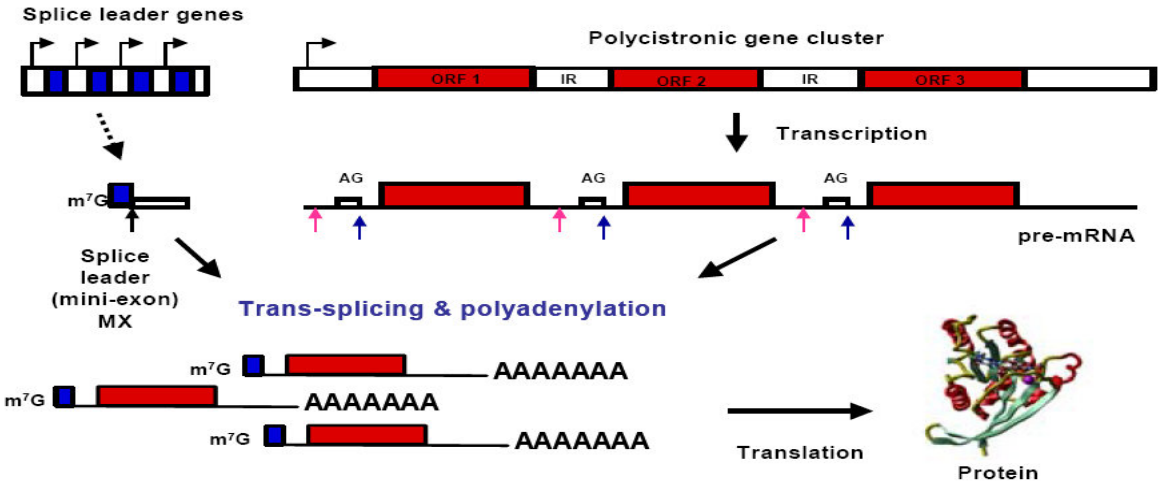

Figure 5: Transcription in L. tarentolae
Transcription in *L. tarentolae* is polycistronic, there are no introns and, therefore, no *cis*-splicing reactions. Posttranscriptionally processing of pre-mRNA is effected by *trans*-splicing reactions and polyadenylation within intergenic regions. *Image from JBS, LEXSY Essence.*

Over the years, further expression vector variants were developed, enabling protein expression in an inducible-episomal [139] or cell-free system [140-142]. In the inducible systems, protein expression is activated through addition of the enhancer tetracycline, being analogue to the well-known bacterial T7 RNA Polymerase/TET repressor system [15]. In the inducible-episomal system, plasmids are maintained extrachromosomally as self-replicating episomes. Finally, proteins can be produced in *L. tarentolae* lysates using the cell-free approach. Species-independent translational sequences (SITS) mediate cell-free protein synthesis bypassing the early translation initiation factors. SITS in combination with targeted suppression of endogenous mRNA of *L. tarentolae* are necessary factors to create a cell-free system.

In my work, only constitutive expression vectors (pLEXSY_sat2 and pLEXSY_hyg2) are used, as they were established first and are promising for successfully protein expression. They offer an expression cassette, which is flanked by ssu coding regions to secure genomic integration. Furthermore, vectors contain three untranslated regions (UTRs), a multi-cloning site for the target gene, an antibiotic resistance gene and a part only functional for clonal selection and vector replication in *E. coli*. UTRs have a major influence on gene regulation [136]. For this,

UTRs consist of parts of the *L. tarentolae* calmodulin cluster (camCBA), of regions from the dihydrofolate reductase–thymidylate synthase locus (*L. major*) and/or of the adenine-phosphorybosyl transferase gen (aprt) [18]. Furthermore, an existing secretory signal peptide (SP) of the secreted acid phosphatase 1 of *L. mexicana* (sAP1) allows the choice between cytoplasmic and secretory recombinant protein expression.

For functional validation, enhanced green fluorescent protein (eGFP) was expressed in all expression vector variants. Further constitutively expressed proteins are for example the Ca^{2+}-dependent serine protease proprotein convertase 4 (PC4) [143], the chloramphenicol acetyl transferase (CAT) [144] or the successfully crystallized Cu/Zn Superoxide dismutase (SOD1) [145], also being expressed in the inducible-episomal system [139]. Furthermore, several human Rab GTPases tagged with eGFP were expressed in the inducible-episomal [139] as well as in the cell-free system [140, 141]. More expressed proteins of different origins and expression vector variants can be found in the review of Basile G. et al [17].

2 Aim

Compared to other expression systems, *L. tarentolae* offers some great advantages. Attributes like its easy bacteria-like handling [15, 16], its non-pathogenicity to humans [14], and evolutionary acquired mimicry enable diverse applications and are great advantages of the parasite. Additionally, many studies report on the abilities of *L. tarentolae* to posttranslationally modify recombinant proteins in a mammalian-like manner, making it more attractive as compared to e.g. yeast or insect cell systems [9, 10, 18]. Unfortunately, not even a handful of these studies is confirming above mentioned characteristic. Therefore, the following aim was set:

Evaluation of the potential of L. tarentolae to express human recombinant proteins of biomedical interest

The main goal of my work is to evaluate the potential of the *L. tarentolae* expression system for biomedical research. This includes the suitability of *L. tarentolae* to express recombinant proteins being potentially human therapeutic agents, e.g. antibody fragments. Therefore, following experiments should be conducted to fulfill this aim or rather to confirm this promising ability: First, new expression vectors should be designed to form the experimental basis. Here, the constitutive expression system of JBS was used. Second, to increase the expression-secretion efficiency of secretory proteins, modification of a SP should be performed and subsequently tested. Third, to confirm mammalian-like PTMs of recombinant proteins expressed with *L. tarentolae*, a highly glycosylated human protein should be selected. Fourth, an in-vivo biotinylation system should be established to enable one-step purification of intracellular/cytoplasmic recombinant proteins preferably resulting in a pure product.

3 Manuscripts

Manuscript I (SP optimization; antibody-fragment expression-secretion study):

Parental constitutive expression vectors of JBS include the SP of sAP1 of *Leishmania mexicana*. For successful protein secretion of high yield, the design of the SP is most important. The certified online tool SignalP (http://www.cbs.dtu.dk/services/SignalP-3.0/) can predict the presence and location of SPs and their cleavage sites in polypeptides and formed the initial point of the study [146, 147]. In detail, the effect of minor amino-acid changes on the cleavage-site and on the cleavage probability of the SP was determined. Several different SPs were designed in-silico, the four most promising candidates were cloned and their functionality was verified by expression-secretion studies with human antibody fragments (scFv). In the end, the protein yield of scFv's could be increased close to an order of magnitude. These results pave the way for further expression-secretion studies in *L. tarentolae*, for example for the expression of full length Ig's or potential highly glycosylated proteins.

Manuscript II (Mammalian glycoprotein expression-secretion; mass spectrometry analyses):

Before a high yield of eukaryotic expressed proteins gets profitable, other factors like solubility and the posttranslational modification of the protein must be fulfilled as well. Glycosylation is one of the most common PTM to occur in eukaryotic protein synthesis, and Leishmania are specifically rich in glycoproteins [121]. For this, the highly glycosylated human protein sAPPalpha was chosen to verify mammalian-like glycosylation characteristics of *L. tarentolae*. Furthermore, the gene for sAPPalpha was cloned into the SP-optimized vector (see Manuscript I) to secure high protein yield. sAPPalpha is a cleavage product of the transmembrane glycoprotein amyloid precursor protein (APP), the etiologic agent in Alzheimer's disease (AD). Detailed analysis of the glycosylation pattern of sAPPalpha identified among eukaryotic species a highly conserved core pentasaccharide. Additionally, for the first time O-glycosylation of a recombinant protein expressed in *L. tarentolae* has been demonstrated. Summarized, human sAPPalpha produced in *L. tarentolae* was N- and O-glycosylated on similar sites as described for mammalian-expressed sAPPalpha and showed similar biological activity. This demonstrates that *L. tarentolae* is a very suitable and simple to handle expression system for mammalian glycoproteins.

Manuscript III (CDK5 regulation by sAPPalpha):

In this study, high-resolution two-dimensional polyacrylamide gel electrophoresis was used to analyse proteome changes downstream of sAPPalpha in neurons. As outcome, regulation of the expression and activity of CDK5 (AD kinase) by sAPPalpha was demonstrated. Furthermore, functional evidence was given that the sAPPalpha receptor SORLA (sorting protein-related receptor containing LDLR class A repeats) is essential for sAPPalpha functionality.

Manuscript IV (Development of an in-vivo biotinylation system; cytoplasmic protein expression):

After improving secretory protein expression, production of intracellular/cytoplasmic proteins was set as the next goal. For this, the powerful *in-vivo* biotinylation system was established as a molecular tool box. As a starting point, strain P10-BirA as well as eGFP Co-transfectants were used, which were already established during my diploma thesis. The *in-vivo* biotinylation system allows site-specific biotinylation of a desired target protein. The only prerequisite is the co-expression of the mammalian codon-optimized biotin ligase (hBirA, from *E. coli*) [148] together with a target protein carrying the specific recognition sequence for the enzyme [149]. The biotin ligase expressing strain was analyzed in detail and tested with the co-expression of eGFP (enhanced green fluorescent protein), p53 (cellular tumor antigen p53), FDFT1 (Farnesyl-diphosphate farnesyltransferase) and mTKIN (kinesin). Fluorescence microscopy studies verified fluorescent eGFP. p53 is a known autophagy-inducing protein. Its functionality could be verified with transmission electron microscopy analyses showing autophago(lyso)somes.. Additionally, the biotin-tag as well as the poly-histidine tag were functional, not disturbing protein functionality.

3.1 Manuscript I

Secretory signal peptide modification for optimized antibody-fragment expression-secretion in *Leishmania tarentolae*

Klatt S, Konthur Z.

Microbial Cell Factories, 2012 July 25; 11:97. PMID: 22830363

DOI: http://dx.doi.org/10.1186/1475-2859-11-97

My contribution to this manuscript:

The experimental part was unexceptionally performed by myself. Dr. Zoltán Konthur and I have designed all experiments of the study and written the manuscript.

Klatt and Konthur *Microbial Cell Factories* 2012, **11**:97
http://www.microbialcellfactories.com/content/11/1/97

MICROBIAL CELL FACTORIES

RESEARCH Open Access

Secretory signal peptide modification for optimized antibody-fragment expression-secretion in *Leishmania tarentolae*

Stephan Klatt[1,2] and Zoltán Konthur[1*]

Abstract

Background: Secretory signal peptides (SPs) are well-known sequence motifs targeting proteins for translocation across the endoplasmic reticulum membrane. After passing through the secretory pathway, most proteins are secreted to the environment. Here, we describe the modification of an expression vector containing the SP from secreted acid phosphatase 1 (SAP1) of *Leishmania mexicana* for optimized protein expression-secretion in the eukaryotic parasite *Leishmania tarentolae* with regard to recombinant antibody fragments. For experimental design the online tool SignalP was used, which predicts the presence and location of SPs and their cleavage sites in polypeptides. To evaluate the signal peptide cleavage site as well as changes of expression, SPs were N-terminally linked to single-chain Fragment variables (scFv's). The ability of *L. tarentolae* to express complex eukaryotic proteins with highly diverse post-translational modifications and its easy bacteria-like handling, makes the parasite a promising expression system for secretory proteins.

Results: We generated four vectors with different SP-sequence modifications based on in-silico analyses with SignalP in respect to cleavage probability and location, named pLTEX-2 to pLTEX-5. To evaluate their functionality, we cloned four individual scFv-fragments into the vectors and transfected all 16 constructs into *L. tarentolae*. Independently from the expressed scFv, pLTEX-5 derived constructs showed the highest expression rate, followed by pLTEX-4 and pLTEX-2, whereas only low amounts of protein could be obtained from pLTEX-3 clones, indicating dysfunction of the SP. Next, we analysed the SP cleavage sites by Edman degradation. For pLTEX-2, -4, and -5 derived scFv's, the results corresponded to *in-silico* predictions, whereas pLTEX-3 derived scFv's contained one additional amino-acid (AA).

Conclusions: The obtained results demonstrate the importance of SP-sequence optimization for efficient expression-secretion of scFv's. We could successfully demonstrate that minor modifications in the AA-sequence in the c-region of the natural SP from SAP1, based on *in-silico* predictions following the (-3, -1) rule, resulted in different expression-secretion rates of the protein of interest. The yield of scFv production could be improved close to one order of magnitude. Therefore, SP-sequence optimization is a viable option to increase the overall yield of recombinant protein production.

Keywords: Secretory signal peptide (SP), SignalP, single-chain Fragment variable (scFv), *Leishmania tarentolae*

* Correspondence: Konthur@molgen.mpg.de
[1]Max Planck Institute for Molecular Genetics, Ihnestraße 63-73, 14195 Berlin, Germany
Full list of author information is available at the end of the article

Background

Leishmania tarentolae is a eukaryotic flagellated unicellular parasite with a broad range of applications [1-3]. An increasing field of interest is its use as a host for recombinant protein expression [4-8]. It allows complex eukaryotic protein expression at high levels compared to other eukaryotes, and has the ability to post-translationally modify proteins. Furthermore, its easy bacteria-like handling makes the parasite a promising expression system for eukaryotic proteins. All these characteristics suggest Leishmania to be a good choice for the expression-secretion of recombinant antibody fragments.

A single-chain Fragment variable (scFv) is the smallest functional entity of a monoclonal antibody consisting of a single-polypeptide. It is composed of the variable regions of the heavy (V_H) and the light (V_L) chain of immunoglobulins, which are connected with a flexible amino-acid (AA) linker of varying length [9].

Antibodies, like many other proteins, are naturally secreted. For targeted protein transport, special sequence motifs are usually necessary [10-12]. Secretory signal peptides (SP) function as sorting signals. In general, they are located at the N-terminus of proteins and their length ranges between 15–30 AAs [13]. During translocation across the endoplasmic reticulum membrane, the SP is usually cleaved off and the protein is entering the secretory pathway [11,13,14]. Changes of 2–4 AAs of the SP-sequence can result in new cleavage sites and in changed expression-secretion efficiency in e.g. lactic acid bacteria [15]. Secretory leader sequence optimization has been widely applied in other organisms as well, such as *E. coli* [16], different yeast strains [17,18], or in insect cells using the baculoviral expression system [19].

In this study, we explored the use of *L. tarentolae* as host for protein expression-secretion of four human recombinant scFv's derived from a semi-synthetic single-framework phage display antibody library [20,21]. To accommodate scFv's with efficient cleavage sites, we followed a two step strategy. First, we set out to *in-silico* model SP-sequences in combination with suitable restriction sites for cloning using SignalP [22]. Second, we designed appropriate vector constructs to evaluate resulting SP-sequences *in-vivo* for optimized scFv expression-secretion in *L. tarentolae*. As a starting point we utilised the commercial *L. tarentolae* protein expression-secretion vector pLEXSY-sat2 (Jena Bioscience), containing the SP-sequence of secreted acid phosphatase 1 (SAP1 [UniProt:Q25332]) of *L. mexicana* [23]. This SP-sequence has been successfully applied in a similar or the same vector for the expression-secretion of e.g human erythropoietin [4], C-reactive protein [24] and heterotrimeric laminins [25].

Results and discussion

In-silico analysis of natural secretory signal peptides

A number of online tools are available for predicting signal peptides and corresponding cleavage sites in protein constructs based on their AA sequence [13]. In two comparative studies, the online program SignalP was identified to be the method of choice [26,27]. Thus, we applied this online tool for cleavage site prediction using an algorithm based on Hidden Markov Models (HMM) on the SP-sequence in vector pLEXSY-sat2 in combination with human scFv sequences.

First, we analysed the cleavage site for natural human IgG expression-secretion with its natural IgG V_H leader peptide [UniProt:Q9Y298] [28] in plasma cells using SignalP 3.0 [29] (**MDWTWRILFLVAAATGTHA**_scFv), which resulted in a 100% cleavage prediction at the first AA of the scFv (Figure 1A). In parallel, we compared the cleavage site of scFv constructs with the *Erwinia carotovora* pectate lyase 2 (pelB) leader peptide [UniProt: P0C1C1] [30] (**MKYLLPTAAAGLLLLAAQPAMA**_scFv; Figure 1B). The SP of pelB is frequently used for scFv expression-secretion in *Escherichia coli* and results in the same cleavage site as during natural human IgG expression-secretion. Next, we analysed whether a simple transfer of a scFv into *L. tarentolae* expression-secretion vector pLEXSY-sat2 would yield the same results. Endogenously, the SP of SAP1 (**MASRLVRVLAAAMLVAAAVSVDA**_SAP1) is cleaved off completely. *In-silico*, SignalP 3.0 assigns a probability of ~95% for this cleavage event (Figure 1C). If scFv's were theoretically to be cloned directly in fusion with the SAP1 SP-sequence neglecting the necessity for a cloning site, the predicted cleavage probability is reduced to ~83% (Figure 1D). In pLEXSY-sat2, the scFv could be cloned downstream of the SP-sequence into the multiple cloning site II (MCS_II, underlined in the following) by using the available restriction site XbaI. This would result in the AA-sequence **MASRLVRVLAAAMLVAAAVSVDAGASLD**_scFv. The *in-silico* cleavage prediction of this construct worsens. No unique cleavage site was assigned, rather multiple sites were predicted with similar probabilities ranging only around 10-30% (Figure 1E). Further, none of the cleavage sites were predicted to yield the same cleavage product as *E. coli* or human cells, i.e. targeting the first AA of the scFv (glutamic acid (E); **E**VQLLES. . .), which is defined as position 1. This is attributed to the fact that the AA composition upstream position 1 (−1 to −3) has an influence on the cleavage site [31]. Additionally, at least one amino-acid is not cleaved off extending the scFv-sequence N-terminally, which could potentially influence protein functionality. For all these reasons we conclude that the use of the pLEXSY-sat2 vector in its current state is not sensible for expression-secretion of scFv's in *L. tarentolae*. Hence, the vector needs specific

Klatt and Konthur *Microbial Cell Factories* 2012, **11**:97
http://www.microbialcellfactories.com/content/11/1/97

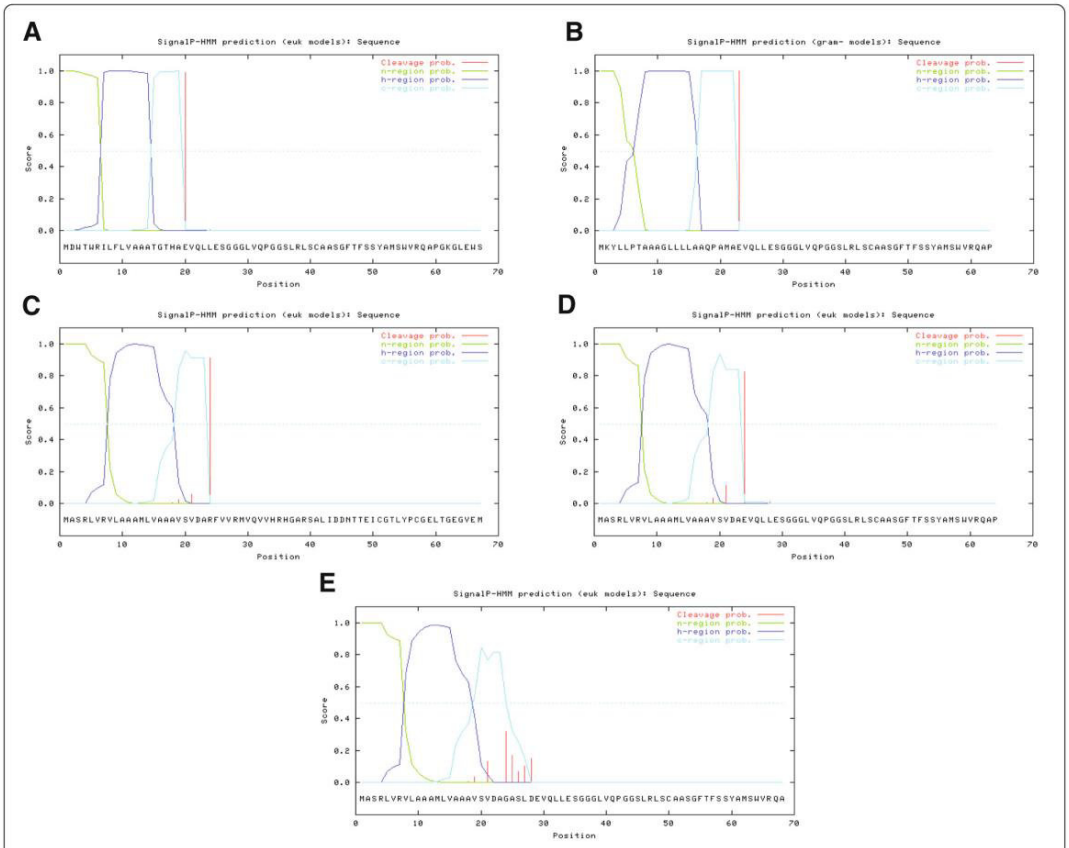

Figure 1 SignalP 3.0 analysis of different SP-sequences. Graphical output of SignalP-HMM, showing the posterior probabilities for n-, h- and c-region and cleavage sites. (**A**) human IgG VH leader peptide with scFv for expression in human cells (**B**) *Erwinia carotovora* pelB SP with human scFv for *E. coli* expression; (**C**) *Leishmania mexicana* secreted acid phosphatase 1 (SAP1); (**D**) the SAP1 SP-sequence with human scFv neglecting any cloning strategy; (**E**). the SAP1 SP-sequence in pLEXSY-sat2 with scFv cloned via available XbaI RE-site.

modification in regard to cleavage site as well as cleavage probability.

In-silico modification of secretory signal peptide SAP1 for scFv expression

Modification of the SP-sequence to enhance cleavage site probability at the first AA of scFv's was carried out *in-silico* following two rationales: First, to optimize AA composition defined by the multi cloning site (MCS II) accommodating unique restriction enzyme (RE) sites for scFv cloning. Second, to find a site with the highest possible cleavage probability and lowest number of alternatives at the first AA residue of the scFv. Ideally, fewer cleavage sites predicted from SignalP and a higher probability of only one successful cleavage should lead to better expression-secretion results of the protein.

The scFv's in our experiments were derived from the phage display vector pIT2, which contains a unique NcoI RE-site upstream of position 1. To enable simple cloning, the MCS II should contain a NcoI RE-site or, alternatively, a RE-site with compatible overhang, such as PciI. Hence, in an iterative process we first set out to modify only the MCSII by introducing changes in the vector pLEXSY-sat2 on the nucleic acid level to accommodate compatible RE-sites for cloning and analysed the results *in-silico* applying SignalP. This approach resulted in the design of pLTEX-2, described in detail below. In a second step, we included the modification of the last AA positions of the SAP1 SP-sequence and conducted the same analyses to obtain the desired characteristics of SP-sequence cleavage at position 1. This approach led to the design of vectors pLTEX-3 to pLTEX-5 (Figure 2).

Klatt and Konthur *Microbial Cell Factories* 2012, **11**:97
http://www.microbialcellfactories.com/content/11/1/97

The detailed rational for all four vectors is described in detail below.

pLTEX-2 was designed on the merit to maintain the original SP-sequence of SAP1, only exchanging the AA encoded by MCS II to accommodate a PciI restriction site. The SP-sequence will be cleaved at the identical position as in SAP1 and, hence, the scFv is extended N-terminally by four AAs. pLTEX-2 has a predicted major cleavage site with a probability of ~81% at the glycine (G) residue in position −4 relative to the scFv's start and two minor cleavage sites at positions −7 and −3 with probabilities of ~10% or below (Figure 2A).

The constructs pLTEX-3 to pLTEX-5 were designed to contain a SP-sequence with a high cleavage probability directly at position 1. To achieve this, the natural SAP1 SP-sequence was shortened by the two AA residues aspartic acid (D) and alanine (A), corresponding to positions −6 and −5 in pLTEX-2. At the same time we introduced variations in the AA residue at position −3 by incorporating asparagine (N), threonine (T) and alanine (A), respectively.

pLTEX-3 contains an asparagine residue at position -3, a predicted major cleavage site with a probability of ~61% at position 1 as well as two additional cleavage sites upstream with lower probabilities (~21% and ~10%; Figure 2B).

pLTEX-4 and pLTEX-5 contain a threonine or alanine residue at position -3 and their predicted major cleavage sites with a probability of close to 95% is at position 1 directly. Both constructs have additional predicted minor cleavage sites with probabilities of below ~5% (Figure 2C,D). The additional modification in pLTEX-4 and pLTEX-5 at the second N-terminal position of the SP-sequence (A to T) was introduced to abolish a NcoI RE-site in the vector and is not considered essential for the functionality of the SP [32].

Generation of expression-secretion vectors pLTEX-1 to pLTEX-5

Novel vectors were designed based on the above considerations for AA composition of SP-sequences. First, the general transfection strategy in *L. tarentolae* was assessed.

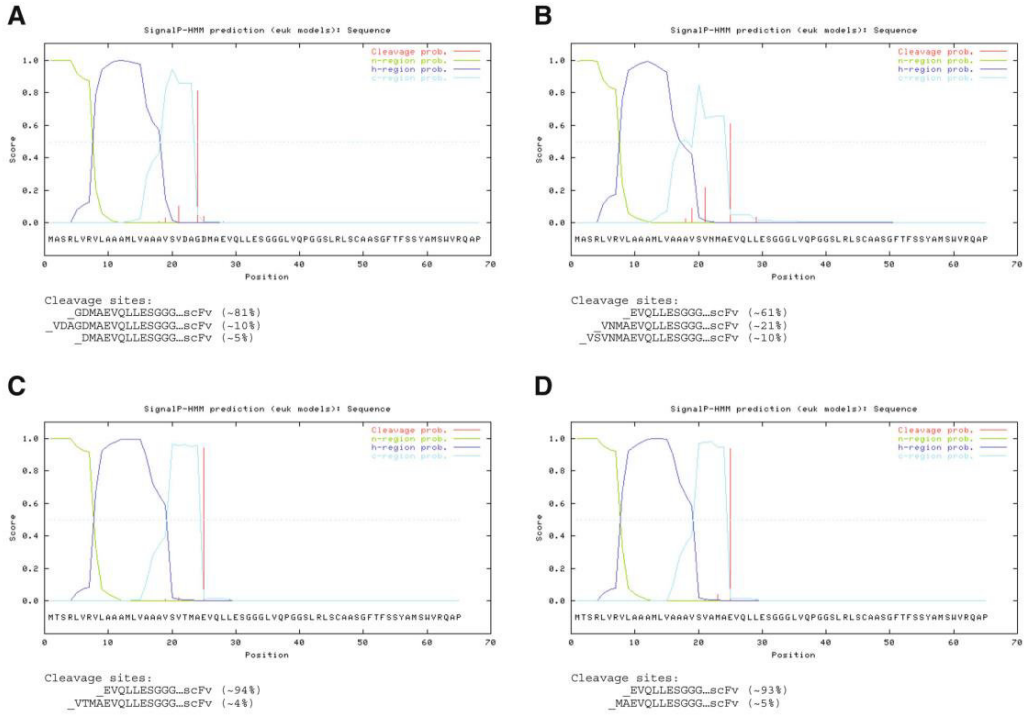

Figure 2 SignalP 3.0 analysis of novel pLTEX vectors with a scFv. Graphical output of SignalP-HMM, showing the posterior probabilities for n-, h- and c-region and cleavage sites of pLTEX vectors with scFv for (**A**) pLTEX-2 (**B**) pLTEX-3, (**C**) pLTEX-4 and (**D**) pLTEX-5. Alternative cleavage sites and respective probabilities are displayed.

Klatt and Konthur *Microbial Cell Factories* 2012, **11**:97
http://www.microbialcellfactories.com/content/11/1/97

To obtain stable expression of recombinant protein a linearised construct needs to be transfected to allow chromosomal integration into a predefined locus by homologous recombination [4]. The original pLEXSY-sat2 vector (Jena Bioscience) is linearised by RE SwaI (Figure 3A). However, all scFv's used in this study contain a conserved SwaI RE-site. Therefore we replaced SwaI with EcoRV. At the same time we reduced the pBluescript II KS (–) backbone of the vector by 1.1 kb to enhance transformation efficiency in *E. coli* generating vector pLTEX-1 (Figure 3A). Note, that pLTEX-1 in its current state is not suitable for scFv expression according to our *in-silico* analysis alike pLEXSY-sat2, since multiple cleavage sites with poor cleavage probabilities are anticipated. Next, the cloning strategy for scFv was determined. MCS I and MCS II were analysed for the availability of compatible restriction sites and modified to allow cloning of scFv's via NcoI and NotI. During the same process, modified SP-sequences were introduced using specific primer-sets creating expression-secretion vectors pLTEX-2 to pLTEX-5 (Figure 3B).

In pLTEX-2 and pLTEX-3, a PciI RE-site was introduced into MCS II, since PciI produces a compatible overhang to NcoI. In vectors pLTEX-4 and pLTEX-5, a NcoI RE-site was introduced in MCS II and at the same time another abolished in MCS I.

Analysis of vector-dependent scFv expression rates *in vivo*

To test the expression-secretion efficiency of the new vectors pLTEX-2 to pLTEX-5, four independent scFv's were cloned into each vector and transfected into *L. tarentolae*. Stable integration of the expression cassette into

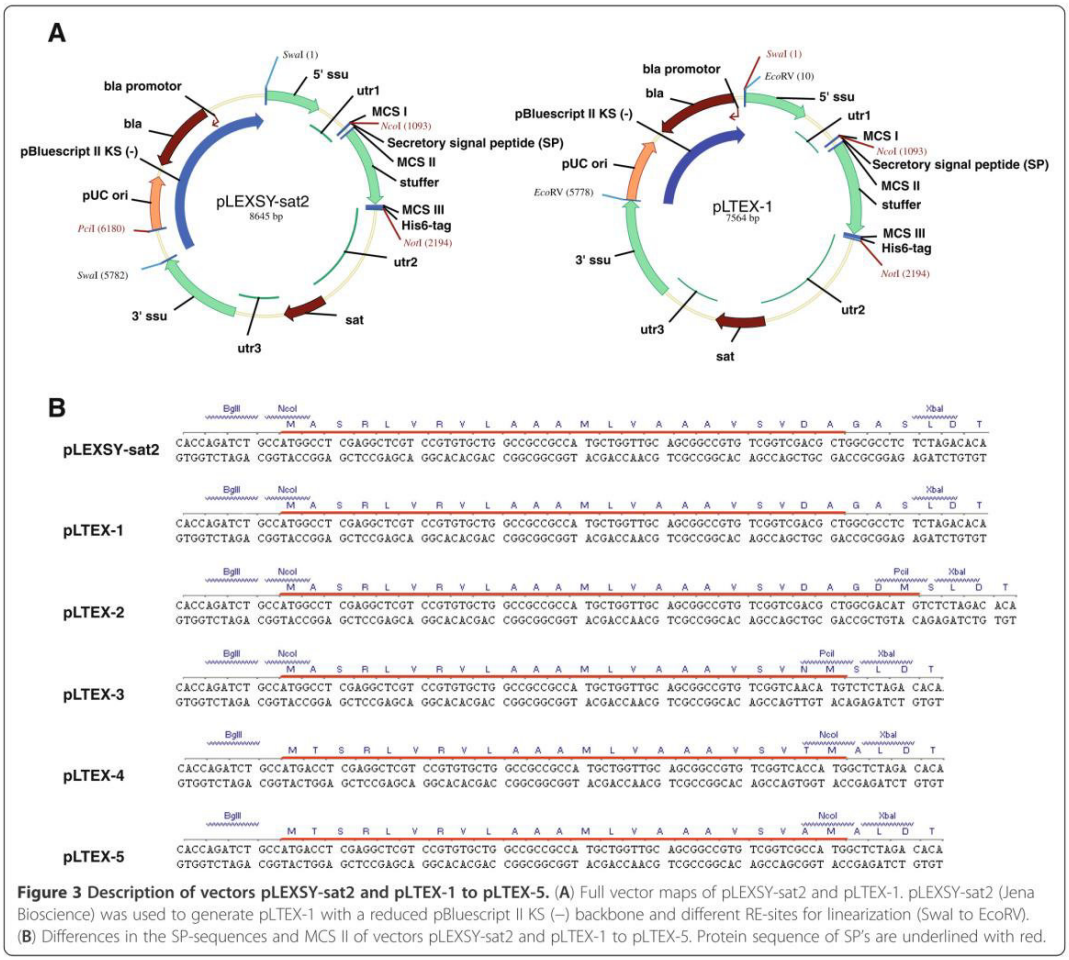

Figure 3 Description of vectors pLEXSY-sat2 and pLTEX-1 to pLTEX-5. (**A**) Full vector maps of pLEXSY-sat2 and pLTEX-1. pLEXSY-sat2 (Jena Bioscience) was used to generate pLTEX-1 with a reduced pBluescript II KS (–) backbone and different RE-sites for linearization (SwaI to EcoRV). (**B**) Differences in the SP-sequences and MCS II of vectors pLEXSY-sat2 and pLTEX-1 to pLTEX-5. Protein sequence of SP's are underlined with red.

the host chromosome was analysed by genomic PCR and could be verified for all 16 constructs (data not shown).

In all pLTEX vectors, protein expression occurs constitutively and was tested for up to four clones per construct. For each construct, the clones showed similar expression levels according to SDS-PAGE analysis and no difference in expression-secretion between single-clones and multi-clones (collection of unknown number of single-clones) from the same vector was observed (data not shown).

First, we analysed the culture supernatant of all 16 different scFv constructs by Western blot for the presence of scFv and observed concentration differences (Figure 4A). Independent of which scFv was cloned, most protein could be detected in supernatants of clones expressed in pLTEX-5 followed by pLTEX-4, pLTEX-2 and pLTEX-3. The nature of SP-sequence and its corresponding cleavage site can influence the expression-secretion efficiency [15] and has a direct influence on the expression level of scFv's.

Next, we purified scFv from the culture supernatant of all 16 clones by immobilised metal-ion chromatography (IMAC) and analysed the data by Coomassie-dyed SDS-PAGE (Figure 4B). As expected, the same order of expression efficiency was observed and the intensity of the protein band corresponding to the scFv (~27 kDa) was quantified densitometrically (Figure 4C). The highest expression rate for scFv's was recorded from clones in vector pLTEX-5, with a median expression of 3.83 mg/L. Clones from vectors pLTEX-4 and pLTEX-2 had a median expression rate of 1.05 mg/L and 0.53 mg/L, respectively. The lowest median expression rate was observed in clones from vector pLTEX-3 (0.04 mg/L).

The observed expression-secretion efficiencies are in good agreement with our knowledge on SP-sequences, formulated as the signal peptidase recognition site or (−3, −1) rule [33] based on the analysis of 78 different natural eukaryotic signal sequences [31]. The preference of a small neutral AA in position −1 and exclusion of a certain large polar, aromatic and charged AA in position −3 can be explained by the nature of the signal peptidase I, which is mainly responsible for SP cleavage. The following ranking was deduced according a 'best fit' analysis for AA and signal peptidase I combination for position −1: A > G,S > C,T > Q; for position −3: A > V > C,S,T > L,I > G [33].

In concordance with this, all our vectors follow this general rule. In detail, our choice for alanine in position −1 respectively to the cleavage site in all pLTEX vectors is optimal. However, the AAs in position −3 are different. Here, pLTEX-5 and pLTEX-4 have the preferred small neutral AA (A, T) whereas pLTEX-2 has a hydrophobic AA (V). The SP-sequence of pLTEX-3 has a large polar AA (N) in position −3 resulting in poor expression, possibly due to steric hindrance within the active site of signal peptidase I resulting in low cleavage efficiency. The expression level is in good agreement with the order of the AA observed in position −3 with alanine giving the best expression-secretion rate.

Finally, we analysed the AA-composition of the N-terminus of all 16 obtained scFv's by Edman degradation [34]. In pLTEX-2 the SP is cleaved off at the predicted glycine residue four AA downstream of the first AA (E) of the scFv. For vectors pLTEX-4 and pLTEX-5 the N-terminus also follows the prediction. The SP is cleaved off at the predicted site, which is glutamic acid (E). Additionally, a second cleavage site could be detected at position −2, extending the scFv's by two AAs; methionine

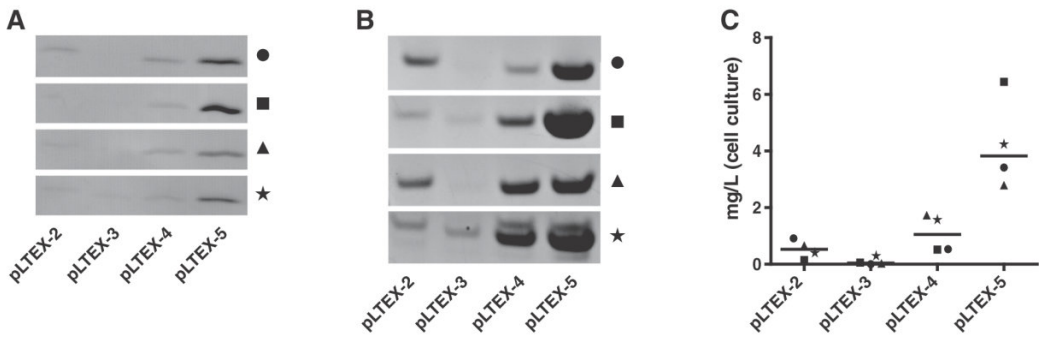

Figure 4 Expression-secretion analysis of 16 scFv's in vectors pLTEX-2 to pLTEX-5. (**A**) Western blot analysis of scFv's in culture supernatant. Detection of scFv's with Protein L-HRP (1:5000). (**B**) SDS-PAGE analysis of Ni-NTA purified scFv's from 100 ml cultures. Proteins are separated on 4-12% Bis-Tris gels (Invitrogen). (**C**) Protein yield from 100 ml culture volumes grown for 72 h of IMAC-purified scFv's determined in duplicate by densitometrical analysis on SDS-PAGE. Black circles, squares, triangles and stars represent scFv-1 to −4, respectively.

and alanine (MA). Notably, the AA-composition in position −5 and position −3 also fulfils the signal peptidase recognition site rule, but with low efficiency. The observed frequency of this cleavage isotype has become visible after digesting large quantities of the protein (data not shown).

Surprisingly, the expected *in-silico* cleavage site at the first AA of the scFv could not be confirmed in the samples derived from vector pLTEX-3. The SP was cleaved one AA downstream at an alanine residue extending the scFv's. We propose that the offset of signal peptidase I occurs as a consequence of the unfavourable AA composition downstream of position 1, resulting in inefficient cleavage explaining the low expression-secretion yield.

Conclusions

To our knowledge, this study represents the first in detail analysis of a signal peptidase recognition site in *L. tarentolae*. We could successfully demonstrate that minor modifications in the AA-sequence of the natural SP from SAP1 based on *in-silico* predictions, following the (-3, -1) rule resulted in different expression-secretion rates of the protein of interest. Using these general principals the yield of scFv production could be improved close to one order of magnitude.

Methods

Generation of expression vectors pLTEX-1 to pLTEX-5

All plasmids and primers used in this study are listed in Tables 1 and 2, respectively. Vector pLTEX-1 (Figure 3A) was generated in three steps: First, the pBluescript II KS (−) backbone (ori for *E. coli*) from pLEXSY-sat2 (Jena Bioscience) was reduced from 2.86 kb to 1.82 kb by PCR amplification using primer 1 and 2. The primers added EcoRV restriction sites to both ends of amplicon 1. Secondly, the chromosomal integration cassette was PCR amplified using primers 3 and 4 (5783 bp) from pLEXSY-sat2. Again, the primers added EcoRV restriction site to both ends creating amplicon 2. Next, amplicon 1 and 2 were digested with EcoRV, ligated, transformed into electrocompetent DH10B cells (Invitrogen) and plated on Ampicillin-containing agar-plates for clonal selection. The obtained vector was named pLTEX-1P (P for precursor). In a final step, the expression cassette was exchanged in pLTEX-1P with a parental fragment from vector pLEXSY-sat2. For this, both vectors were digested with

Table 1 Vectors and constructs used in this study

Name	Characteristics	Reference
pLEXSY-sat2	Expression vector pLEXSY-sat2 with sat gene (streptothricine acetyltransferase) allowing selection of recombinant LEXSY strains with antibiotic NTC; and Amp for DH10B	Jena Bioscience (Cat.-No. EGE-234)
pLTEX-1	Based on pLEXSY-sat2; with reduced pBlueScript backbone and changed RE-site for vector linearization prior transfection (SwaI to EcoRV)	this study
pLTEX-2	based on pLTEX-1; changed SP sequence	this study
pLTEX-3	based on pLTEX-1; changed SP sequence	this study
pLTEX-4	based on pLTEX-1; changed SP sequence	this study
pLTEX-5	based on pLTEX-1; changed SP sequence	this study
pLTEX-2_scFv-1	Vector for expression-secretion of scFv-1	this study
pLTEX-3_scFv-1	Vector for expression-secretion of scFv-1	this study
pLTEX-4_scFv-1	Vector for expression-secretion of scFv-1	this study
pLTEX-5_scFv-1	Vector for expression-secretion of scFv-1	this study
pLTEX-2_scFv-2	Vector for expression-secretion of scFv-2	this study
pLTEX-3_scFv-2	Vector for expression-secretion of scFv-2	this study
pLTEX-4_scFv-2	Vector for expression-secretion of scFv-2	this study
pLTEX-5_scFv-2	Vector for expression-secretion of scFv-2	this study
pLTEX-2_scFv-3	Vector for expression-secretion of scFv-3	this study
pLTEX-3_scFv-3	Vector for expression-secretion of scFv-3	this study
pLTEX-4_scFv-3	Vector for expression-secretion of scFv-3	this study
pLTEX-5_scFv-3	Vector for expression-secretion of scFv-3	this study
pLTEX-2_scFv-4	Vector for expression-secretion of scFv-4	this study
pLTEX-3_scFv-4	Vector for expression-secretion of scFv-4	this study
pLTEX-4_scFv-4	Vector for expression-secretion of scFv-4	this study
pLTEX-5_scFv-4	Vector for expression-secretion of scFv-4	this study

Table 2 Oligonucleotides used in this study

Name	Sequence (5' to 3')
(1) pre-pBR322-ori-EcoRV	CGGATATCTGAGCAAAAGGCCAGCAAAA
(2) pre-blaP-SwaI-EcoRV	CGGATATCCAATTTAAATGCGGAACCCCTATTTGTTTATT
(3) 5_EcoRV_5_ssu	TTGGATATCTTGGCGAAACGCC
(4) EcoRV_3_ssu	GATTTAGATATCGGTGAACTTTCGGG
(5) Pci_linker_sat4	GCTCTAGAGACATGTCGCCAGCGTCGACCG
(6) Pci_linker_sat5	GCTCTAGAGACATGTTGACCGACACGGCCGCTG
(7) BglII_linker_sat6	GAAGATCTGCCATGACCTCGAGGCTCGTC
(8) LEXSY_A264	CATCTATAGAGAAGTACACGTAAAAG
(9) NcoI_linker_sat6	GCTCTAGAGCCATGGTGACCGACACGGCCGCTG
(10) NcoI_linker_sat7	GCTCTAGAGCCATGGCGACCGACACGGCCGCTG
(11) LMB3	CAG GAA ACA GCT ATG AC
(12) LEXSY_scFv_KpnI	GGTGGGTACCCCGTTTGATTTCCACCTTGGTC

BsrGI and BamHI, respective fragments were gel-purified, ligated and transformed into DH10B cells to generate the expression vector pLTEX-1. Vectors pLTEX-2 to –5 are based on pLTEX-1 and only differ in the SP-sequence. Vector pLTEX-2 was generated from pLTEX-1 and amplicon 3, which was a PCR fragment from pLEXSY-sat2 using primers 3 and 5. Both, PCR fragment and pLTEX-1 were digested with BamHI and XbaI, ligated together and transformed into DH10B cells to generate pLTEX-2. Vectors pLTEX-1 and pLTEX-2 only differ in MCS II that now contains a RE-site for PciI. Vector pLTEX-3 was generated as described for pLTEX-2, except using primer 6 instead of 5. For the generation of pLTEX-4, another precursor (pLTEX-4P) was necessary. First, a PCR amplicon using primers 7 (adding cutting site BglII) and 8 was generated on pLEXSY-sat2 as template. The PCR-fragment and pLTEX-1 were then digested with BglII and KpnI, ligated together and transformed into DH10B to generate pLTEX-4P. In a second step, a PCR amplicon using primers 3 and 9 (adding cutting site NcoI) was generated on pLTEX-4P. This PCR-fragment and pLTEX-1 were digested with BamHI and XbaI, ligated and transformed into DH10B cells to generate pLTEX-4. Vector pLTEX-5 was constructed starting from pLTEX-4P by generating a PCR amplicon using primers 3 and 10 (adding cutting site NcoI). Both, PCR-fragment and pLTEX-1 were digested with BamHI and XbaI, ligated and transformed into DH10B cells to generate pLTEX-5. All new vectors were sequence-verified by DNA sequencing applying the Sanger method.

Cloning of antibody fragments

Individual scFv's were PCR-amplified from the original clones in vector pIT2 using primer 11 and 12 (adding cutting site KpnI). The PCR-fragments (729 bp) were digested with NcoI and KpnI for cloning, ligated into cut vector backbones and transformed into DH10B. Vectors pLTEX-2 and pLTEX-3 were digested with PciI and KpnI, while vectors pLTEX-4 and pLTEX-5 were digested with NcoI + KpnI. Transformants were picked from Amp^R agar-plates and sequence verified. Altogether, 16 different scFv-containing vector constructs were designed (Table 1).

E. coli growth conditions and cell transformation

E. coli strain 'ElectroMAX^{Tm} DH10B^{Tm} (Invitrogen; Cat.-No. 18290–015) was grown at 37°C with 160-180 rpm in liquid 2YT growth medium [16 g/l bacto-tryptone (Becton, Dickinson and Company), 10 g/l bacto-yeast extract (Becton, Dickinson and Company), 5 g/l NaCl (Calbiochem); resuspended in Millipore water] supplemented with ampicillin (100 μg/ml; Roth), if necessary. Electro-competent cells were transformed with 1 μl (50-200 ng/μl) purified plasmid-DNA (QIAprep Spin Miniprep Kit; Qiagen) using the 'Micropulser Electroporator' (BioRad) at 1.8 kV, 5 msec. For clonal selection, agar-plates were prepared by adding 15 g/L bacto-agar-agar (Becton, Dickinson and Company) to 2YT growth medium, plating cells and incubating them over night at 37°C.

L. tarentolae growth conditions and stable transfection

L. tarentolae strain P10 (Jena Bioscience) was grown at 26°C under standard air-supply in the dark in LEXSY BHI liquid medium (Jena Bioscience) with porcine Hemin, penicillin/streptomycin (Pen/Strep) and Nourseothricin (NTC; 100 μg/ml), if necessary (all additives: Jena Bioscience). Strain P10 was passaged twice a week by inoculation of new growth medium in flat flaks (static) with a dilution of 1:30. Electroporation of cells was carried out in a prechilled 2 mm electroporation cuvette using a Gene Pulser II with pulse controller and capacitance extender plus (Biorad). For transfection, 5-10 μl purified and linearized plasmid-DNA (50-

150 ng/μl) was added to 390-395 μl densely grown cells (over night as agitated suspension culture (125 rpm); OD_{600} >1.5). Cells were pulsed once at 0.45 kV and 450 μF to get 5-6 msec pulsing time with ~0.45 kV and ~20Ω, chilled shortly and cultured stationary over night at 26°C. On the next day, the majority of the culture was dispensed onto LEXSY BHI agar plates, while the rest was cultivated in flat flaks with new selection pressure (NTC) to get multi-clone cultures. After 6–10 days, single clones were picked and grown in liquid culture by continuously increasing the volume from 0.1 ml to 10 ml in LEXSY BHI growth medium. Finally, transfection success was verified for integration of the expression cassette into the host chromosome by analytical PCR as described [4].

Production of scFv's expressed in *Leishmania tarentolae*

Positive transfectants were used for scFv-production in 100 ml of LEXSY BHI medium containing Hemin, Pen/Strep and NTC. For this, medium was inoculated 1:30 with densely grown cultures and cultivated in agitated suspension culture (125 rpm) for 2–3 days at 26°C until OD_{600} 2.5-4.5 was reached. Cells were harvested by centrifugation for 18 min, at 1811 g and 4°C. Pellet and supernatant were separated. 10 ml/l PMSF (0.1molar; Phenylmethanesulfonyl fluoride, Sigma-Aldrich; inhibits proteases) and 1 ml Ni-NTA agarose (Invitrogen) were added to 100 ml supernatant and rotated for 2.5-3 h at 4°C. Ni-NTA was washed 2x with 5 ml washing buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole) and elution was carried out 4x with 1 ml elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole). All centrifugation steps were performed for 5 min, at 652 g and 4°C. Elution fractions were pooled and analysed on 4-12% Bis-Tris SDS-PAGE (Invitrogen) and western blot after electrophoretical transfer to nitrocellulose membrane. Detection of scFv's was carried out using Protein L-HRP (Pierce) diluted 1:5000 in 2% milk powder in PBS with 0.1% Tween 20 and subsequent incubation with CN/DAB substrate (Pierce). The protein concentration was determined by gel densitometric analysis of the scFv bands in Coomassie-dyed SDS-PAGE using ImageJ (Version 1.45 s). As reference, a calibration curve on the basis of a quantifiable protein in a marker (Precision Plus Unstained Ladder, 25 kDa band, 0.12 – 1.2 μg protein) or a scFv expressed in *E. coli* was established and used for calculating the protein concentration applying linear regression equation. N-terminal scFv's analysis by Edman degradation was undertaken by Proteome Factory (Berlin, Germany).

Abbreviations

AA: Amino-acid; HMM: Hidden Markov Model; LEXSY BHI: Leishmania expression system brain-heart fusion medium I; MCS: Multiple-cloning site; NTC: Nourseothricin; Pen/Strep: Penicillin/Streptomycin; RE: Restriction enzyme; SAP1: Secreted acid phosphatase 1; scFv: Single-chain fragment variable; SP: Secretory signal peptide.

Competing interests

Both authors declare that they have no competing interests.

Authors' contributions

SK has performed the experiments. ZK has conceived the study. SK and ZK have designed the experiments and written the manuscript. Both authors read and approved the final manuscript.

Acknowledgments

The authors acknowledge support by the German Federal Ministry for Education and Research (BMBF) through Grant No. 0315448B and the Max Planck Society for the Advancement of Sciences. We thank Dr. Thomas Schirrmann for his advice in gel-densitometry and Dr. Jonathan Woodsmith for critical reading of the manuscript.

Author details

[1]Max Planck Institute for Molecular Genetics, Ihnestraße 63-73, 14195 Berlin, Germany. [2]Freie Universität Berlin, Faculty of Biology, Chemistry and Pharmacy, Takustraße 3, 14195 Berlin, Germany.

Received: 15 February 2012 Accepted: 1 July 2012
Published: 25 July 2012

References

1. Basak A, Shervani NJ, Mbikay M, Kolajova M: **Recombinant proprotein convertase 4 (PC4) from Leishmania tarentolae expression system: purification, biochemical study and inhibitor design.** *Protein Expr Purif* 2008, **60**(2):117–126.
2. Breton M, Zhao C, Ouellette M, Tremblay MJ, Papadopoulou B: **A recombinant non-pathogenic Leishmania vaccine expressing human immunodeficiency virus 1 (HIV-1) Gag elicits cell-mediated immunity in mice and decreases HIV-1 replication in human tonsillar tissue following exposure to HIV-1 infection.** *J Gen Virol* 2007, **88**(Pt 1):217–225.
3. Yakovich AJ, Ragone FL, Alfonzo JD, Sackett DL, Werbovetz KA: **Leishmania tarentolae: purification and characterization of tubulin and its suitability for antileishmanial drug screening.** *Exp Parasitol* 2006, **114**(4):289–296.
4. Breitling R, Klingner S, Callewaert N, Pietrucha R, Geyer A, Ehrlich G, Hartung R, Muller A, Contreras R, Beverley SM, *et al*: **Non-pathogenic trypanosomatid protozoa as a platform for protein research and production.** *Protein Expr Purif* 2002, **25**(2):209–218.
5. Fritsche C, Sitz M, Weiland N, Breitling R, Pohl HD: **Characterization of the growth behavior of Leishmania tarentolae: a new expression system for recombinant proteins.** *J Basic Microbiol* 2007, **47**(5):384–393.
6. Fritsche C, Sitz M, Wolf M, Pohl HD: **Development of a defined medium for heterologous expression in Leishmania tarentolae.** *J Basic Microbiol* 2008, **48**(6):488–495.
7. Kushnir S, Gase K, Breitling R, Alexandrov K: **Development of an inducible protein expression system based on the protozoan host Leishmania tarentolae.** *Protein Expr Purif* 2005, **42**(1):37–46.
8. Mureev S, Kushnir S, Kolesnikov AA, Breitling R, Alexandrov K: **Construction and analysis of Leishmania tarentolae transgenic strains free of selection markers.** *Mol Biochem Parasitol* 2007, **155**(2):71–83.
9. Bird RE, Hardman KD, Jacobson JW, Johnson S, Kaufman BM, Lee SM, Lee T, Pope SH, Riordan GS, Whitlow M: **Single-chain antigen-binding proteins.** *Science* 1988, **242**(4877):423–426.
10. Rapoport TA: **Transport of proteins across the endoplasmic reticulum membrane.** *Science* 1992, **258**(5084):931–936.
11. von Heijne G: **The signal peptide.** *J Membr Biol* 1990, **115**(3):195–201.
12. von Heijne G: **Life and death of a signal peptide.** *Nature* 1998, **396**(6707):111–113.
13. Emanuelsson O, Brunak S, von Heijne G, Nielsen H: **Locating proteins in the cell using TargetP, SignalP and related tools.** *Nat Protoc* 2007, **2**(4):953–971.
14. Sakaguchi M: **Eukaryotic protein secretion.** *Curr Opin Biotechnol* 1997, **8**(5):595–601.
15. Borrero J, Jimenez JJ, Gutiez L, Herranz C, Cintas LM, Hernandez PE: **Protein expression vector and secretion signal peptide optimization to drive the**

production, secretion, and functional expression of the bacteriocin enterocin A in lactic acid bacteria. *J Biotechnol* 2011, **156**(1):76–86.
16. Jonet MA, Mahadi NM, Murad AM, Rabu A, Bakar FD, Rahim RA, Low KO, Illias RM: **Optimization of a heterologous signal peptide by site-directed mutagenesis for improved secretion of recombinant proteins in Escherichia coli.** *J Mol Microbiol Biotechnol* 2012, **22**(1):48–58.
17. Liu FC, Chen HL, Chong KY, Hsu AL, Chen CM: **Production of recombinant porcine colipase secreted by Pichia pastoris and its application to improve dietary fat digestion and growth of postweaning piglets.** *Biotechnol Prog* 2008, **24**(6):1333–1341.
18. Rakestraw JA, Sazinsky SL, Piatesi A, Antipov E, Wittrup KD: **Directed evolution of a secretory leader for the improved expression of heterologous proteins and full-length antibodies in Saccharomyces cerevisiae.** *Biotechnol Bioeng* 2009, **103**(6):1192–1201.
19. Futatsumori-Sugai M, Tsumoto K: **Signal peptide design for improving recombinant protein secretion in the baculovirus expression vector system.** *Biochem Biophys Res Commun* 2009, **391**(1):931–935.
20. de Wildt RM, Mundy CR, Gorick BD, Tomlinson IM: **Antibody arrays for high-throughput screening of antibody-antigen interactions.** *Nat Biotechnol* 2000, **18**(9):989–994.
21. Holt LJ, Bussow K, Walter G, Tomlinson IM: **By-passing selection: direct screening for antibody-antigen interactions using protein arrays.** *Nucleic Acids Res* 2000, **28**(15):E72.
22. Nielsen H, Engelbrecht J, Brunak S, von Heijne G: **A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites.** *Int J Neural Syst* 1997, **8**(5–6):581–599.
23. Wiese M, Ilg T, Lottspeich F, Overath P: **Ser/Thr-rich repetitive motifs as targets for phosphoglycan modifications in Leishmania mexicana secreted acid phosphatase.** *EMBO J* 1995, **14**(6):1067–1074.
24. Dortay H, Schmockel SM, Fettke J, Mueller-Roeber B: **Expression of human c-reactive protein in different systems and its purification from Leishmania tarentolae.** *Protein Expr Purif* 2011, **78**(1):55–60.
25. Sugino M, Niimi T: **Expression of multisubunit proteins in Leishmania tarentolae.** *Methods Mol Biol* 2012, **824**:317–325.
26. Choo KH, Tan TW, Ranganathan S: **A comprehensive assessment of N-terminal signal peptides prediction methods.** *BMC Bioinforma* 2009, **10**(Suppl 15):S2.
27. Petersen TN, Brunak S, von Heijne G, Nielsen H: **SignalP 4.0: discriminating signal peptides from transmembrane regions.** *Nat Methods* 2011, **8**(10):785–786.
28. Jacquemin MG, Desqueper BG, Benhida A, Vander Elst L, Hoylaerts MF, Bakkus M, Thielemans K, Arnout J, Peerlinck K, Gilles JG, *et al*: **Mechanism and kinetics of factor VIII inactivation: study with an IgG4 monoclonal antibody derived from a hemophilia A patient with inhibitor.** *Blood* 1998, **92**(2):496–506.
29. Bendtsen JD, Nielsen H, von Heijne G, Brunak S: **Improved prediction of signal peptides: SignalP 3.0.** *J Mol Biol* 2004, **340**(4):783–795.
30. Lei SP, Lin HC, Wang SS, Callaway J, Wilcox G: **Characterization of the Erwinia carotovora pelB gene and its product pectate lyase.** *J Bacteriol* 1987, **169**(9):4379–4383.
31. von Heijne G: **Patterns of amino acids near signal-sequence cleavage sites.** *Eur J Biochem* 1983, **133**(1):17–21.
32. von Heijne G: **Analysis of the distribution of charged residues in the N-terminal region of signal sequences: implications for protein export in prokaryotic and eukaryotic cells.** *EMBO J* 1984, **3**(10):2315–2318.
33. von Heijne G: **How signal sequences maintain cleavage specificity.** *J Mol Biol* 1984, **173**(2):243–251.
34. Edman P: **A method for the determination of amino acid sequence in peptides.** *Arch Biochem* 1949, **22**(3):475.

doi:10.1186/1475-2859-11-97
Cite this article as: Klatt and Konthur: Secretory signal peptide modification for optimized antibody-fragment expression-secretion in *Leishmania tarentolae*. *Microbial Cell Factories* 2012 **11**:97.

3.2 Manuscript II

Production of Glycosylated Soluble Amyloid Precursor Protein Alpha (sAPPalpha) in *Leishmania tarentolae*

Klatt S, Rohe M, Alagesan K, Kolarich D, Konthur Z, Hartl D.

Journal of Proteome Research, 2013, 12 (1), pp 396–403. PMID: 23214446

My contribution to this manuscript:

My experimental contributions to this manuscript include vector design with signal peptide cleavage analysis, cloning of sAPPalpha into the Leishmania expression-secretion vector pLTEX-5 and optimized production and purification of sAPPalpha from *L. tarentolae*.

By doing initial glycoprotein studies with PNGase F, I confirmed the high glycosylation status of sAPPalpha. I contacted Dr. Daniel Kolarich from the MPIKG, who is an expert in glycoproteomics, and convinced him to do the glycan analyses of sAPPalpha. We designed the procedures together and in the end, we could verify, inter alia, O-glycosylation of sAPPalpha, a PTM that has never been described for a recombinant protein expressed in *L. tarentolae* before. Furthermore, I was involved in writing of the manuscript.

3.2 Manuscript III

Soluble alpha-APP (sAPPalpha) regulates CDK5 expression and activity in neurons

Hartl D, **Klatt S**, Roch M, Konthur Z, Klose J, Willnow TE, and Rohe M

PLoS One, 2013 Jun 11; 8(6); e65920. PMID: 23776568

DOI: http://dx.doi.org/10.1371/journal.pone.0065920

My contribution to this manuscript:

My experimental contributions to this manuscript include the expression of sAPPalpha in *L. tarentolae* and subsequent purification. Furthermore, I have participated in the analysis and interpretation of the results.

OPEN ACCESS Freely available online

PLOS ONE

Soluble Alpha-APP (sAPPalpha) Regulates CDK5 Expression and Activity in Neurons

Daniela Hartl[1]*, Stephan Klatt[2,3], Manfred Roch[1], Zoltan Konthur[2,4], Joachim Klose[1], Thomas E. Willnow[5], Michael Rohe[5]*

1 Institute for Medical Genetics and Human Genetics, Charité - University Medicine, Berlin, Germany, **2** Max Planck Institute for Molecular Genetics, Berlin, Germany, **3** Free University Berlin, Faculty of Biology, Chemistry and Pharmacy; Berlin, Germany, **4** Max Planck Institute of Colloids and Interfaces, Potsdam, Germany, **5** Max-Delbrueck-Center for Molecular Medicine, Berlin, Germany

Abstract

A growing body of evidence suggests a role for soluble alpha-amyloid precursor protein (sAPPalpha) in pathomechanisms of Alzheimer disease (AD). This cleavage product of APP was identified to have neurotrophic properties. However, it remained enigmatic what proteins, targeted by sAPPalpha, might be involved in such neuroprotective actions. Here, we used high-resolution two-dimensional polyacrylamide gel electrophoresis to analyze proteome changes downstream of sAPPalpha in neurons. We present evidence that sAPPalpha regulates expression and activity of CDK5, a kinase that plays an important role in AD pathology. We also identified the cytoprotective chaperone ORP150 to be induced by sAPPalpha as part of this protective response. Finally, we present functional evidence that the sAPPalpha receptor SORLA is essential to mediate such molecular functions of sAPPalpha in neurons.

Citation: Hartl D, Klatt S, Roch M, Konthur Z, Klose J, et al. (2013) Soluble Alpha-APP (sAPPalpha) Regulates CDK5 Expression and Activity in Neurons. PLoS ONE 8(6): e65920. doi:10.1371/journal.pone.0065920

Editor: Doris Kretzschmar, Oregon Health and Science University, United States of America

Received February 26, 2013; **Accepted** April 29, 2013; **Published** June 11, 2013

Funding: Studies were funded by grants from the American Health Assistance Foundation (to MR; www.brightfocus.org; project A2011601) and the German Research Foundation (to DH; www.dfg.de; project HA6155/2-2). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: daniela.hartl@charite.de (DH); michael.rohe@mdc-berlin.de (MR)

Introduction

Amyloid precursor protein (APP) is a major etiologic agent in Alzheimer disease (AD). Amyloidogenic processing of APP by beta- and gamma-secretases generates Amyloid-beta (Abeta), the main component of senile plaques. Accumulating evidence supports the notion that progression of AD is correlated with progressive accumulation of Abeta in the brain which can be caused by elevated production and aggregation or impaired clearance of the peptide [1]. More recently, a growing body of evidence also implicated another APP cleavage product, soluble alpha-APP (sAPPalpha), in AD pathology. This product is produced from APP by an alternative, non-amyloidogenic processing pathway and levels of sAPPalpha were shown to be reduced in the cerebrospinal fluid of humans with sporadic or familial AD [2–4].

Important evidence for a central function of sAPPalpha in the brain was provided by studies with mice deficient for APP and APLP2, an APP homolog with overlapping functions. Perinatal lethality as well as other phenotypical abnormalities of APP/APLP2 double knockout mice were rescued by a sAPPalpha transgene [5]. Furthermore, if infused into the brains of rodents, sAPPalpha enhanced synaptogenesis and memory formation [6,7]. Other studies reported enhanced survival of cells and neurite outgrowth after application of sAPPalpha to cultured cortical and hippocampal neurons [8–11] and other cell types [12–14]. Together, sAPPalpha was proposed to have neurotrophic and neuroprotective properties, possibly counteracting neurotoxic effects of Abeta. Accordingly, loss of sAPPalpha as observed in AD patients might contribute to disease pathology.

After cleavage of APP by alpha-secretase, sAPPalpha is released into the extracellular space in a process that was reported to be coupled to synaptic activity [15]. Only little is known about mechanisms controlling receptor-mediated uptake and downstream signalling of sAPPalpha in neurons. However, recent evidence suggested that sorting protein-related receptor containing LDLR class A repeats (SORLA; also known as LR11), an important AD risk factor, may act as sAPPalpha receptor [16,17].

Here, we have used high-resolution two-dimensional polyacrylamide gel electrophoresis to determine proteins altered in expression after sAPPalpha application to primary cortical mouse neurons. We show that sAPPalpha regulates expression and activity of cyclin-dependent kinase 5 (CDK5), a kinase that plays an important role in AD pathology and that was previously shown to be activated by Abeta. We also identified hypoxia up-regulated protein 1 (ORP150) as effector protein potentially mediating neuroprotective functions of sAPPalpha. Finally, we present functional evidence that the sAPPalpha receptor SORLA determines these molecular functions of sAPPalpha.

Materials and Methods

Ethics Statement

All experiments performed with mice were conducted according to the guidelines of the German Animal Welfare Law. The study

was approved by the State Office of Health and Social Affairs Berlin (approval number T0297/01).

Preparation and Treatment of Primary Neurons

Primary cortical neurons were prepared from newborn Balb/c mice of either sex at postnatal day 1. Cortices were dissociated in papain (1 hour at 37°C) and cultured on poly-D-lysine/collagen coated culture dishes. The neurons were cultured for 4–5 days in Neurobasal-A medium (Gibco) including B27 supplement (Sigma), and GlutaMAX (Invitrogen) as previously described [18].

Neurons were treated with human recombinant neuron-specific sAPPalpha (APP isoform 695) produced in *E. coli* (SIGMA; for proteome analysis) or *Leishmania tarentolae* (for Western blot analysis) prepared as described before [19]. Neurons were supplied with sAPPalpha (300 ng/ml) or medium only (control) for one hour or 48 hours, respectively, by replacing half of the culture medium with fresh medium. For proteome analyses, the cells were harvested in ice cold PBS and cell pellets were frozen immediately in liquid nitrogen. Six individual samples of each, treated and control cells were collected for each condition (n = 6). Generation of animals genetically deficient for *Sorl1*$^{-/-}$ has been described before [20].

For detection of sAPPalpha, human specific monoclonal antibody (clone 2B3, IBL Hamburg), which recognizes the C-terminus of human sAPPalpha (DAEFRHDSGYEVHHQK) was used. All other antisera have been obtained from Cell Signaling Technology. Western blotting was conducted according to standard procedures. Protein signals were measured with a CCD camera based chemiluminescence imaging system (Peqlab). Quantification of signal intensities was accomplished with ImageJ software and significance of changes was determined applying Mann-Whitney-U test ($p \leq 0.05$) using Prism (version 5.0c, GraphPad).

Protein Extraction and 2-D Electrophoresis

For proteome analyses, the cells were harvested in ice cold PBS. Six individual samples of each treated and control cells were collected (n = 6). Protein extracts were prepared from frozen cell pellets as described [21]. Briefly, samples together with sample buffer (50 mM TRIZMA Base (Sigma-Aldrich), 50 mM KCl and 20% w/v glycerol at pH 7.5) as well as proteinase and phosphatase inhibitors (Complete and PhosStop, Roche Diagnostics) were ground to fine powder in liquid nitrogen and sonicated on ice subsequently. Afterwards, DNAse (Benzonase, Merck), urea and thiourea (Biorad; 6.5 M and 2 M, respectively) were added. Protein extracts were then supplied with 70 mM dithiothreitol (Biorad), 2% v/w of ampholyte mixture (Servalyte pH 2–4, Serva) and stored at −80°C until separation by 2-D electrophoresis.

High-resolution, large-gel 2D-electrophoresis was described previously [22]. The gel format was 40 cm (isoelectric focusing)×30 cm (SDS-PAGE)×0.9 mm (gel width). Proteins were first separated according to their isoelectric points (isoelectric focussing, IEF) using the carrier-ampholyte technique [21,22]. 40 µg of protein was applied to the acidic end of IEF gels (40 cm) and a carrier ampholyte mixture was added to establish a pH gradient spanning a range from pH 3 to 10. For SDS-PAGE, IEF gels were cut in half and run as "acidic" and "basic" sides. Two-dimensional protein patterns were obtained by sliver staining of gels as described [22]. The 2-D images were scanned at 300 dpi and 16-bit gray scale and saved in Tiff format to avoid loss of quality due to compression.

Protein spot patterns were evaluated using Delta2D imaging software (version 4.0, Decodon). Percent volume of spot pixel intensities was used for quantitative analysis of protein expression by Delta2D as described before [21]. Paired Student's t-test was applied to determine statistical significance of alterations (significance threshold $p \leq 0.05$) as described before [21,23]. Only fold changes exceeding 20% were considered.

Mass Spectrometry

For protein identification, 1200 µg protein extract each was separated on a 2-D gel and stained with a MS-compatible silver staining protocol [22]. Protein spots of interest were excised from gels and subjected to in-gel tryptic digestion followed by HPLC separation as described [22]. Peptides were characterized by an ESI-tandem-MS/MS on a LCQ Deca XP ion trap instrument (Thermo Finnigan, Waltham, MA). Mass spectra were evaluated using MASCOT (version 2.1) automatically searching SwissProt database (version 51.8/513877 sequences). MS/MS ion search was performed with the following set of parameters: (i) taxonomy: Mus musculus, (ii) proteolytic enzyme: trypsin, (iii) maximum of accepted missed cleavages: 1, (iv) mass value: monoisotopic, (v) peptide mass tolerance 0.8 Da, (vi) fragment mass tolerance: 0.8 Da, (vii) fixed modifications: none and (viii) variable modifications: oxidation of methionine and acrylamide adducts (propionamide) on cysteine. Only proteins with scores corresponding to $p < 0.05$, with at least two independent peptides identified were considered. The cut-off score for individual peptides was equivalent to $p < 0.05$ for each peptide as calculated by MASCOT.

Results

Proteome Analysis of Murine Neurons Treated with sAPPalpha

Cumulative evidence suggests neurotrophic properties of sAPPalpha [6–11,24,25]. However, little is known about the underlying molecular mechanisms. We performed high-resolution 2-D electrophoresis (2-DE) combined with mass spectrometry to screen for proteins altered downstream of sAPPalpha signalling in neurons. This 2-DE based approach allowed us to simultaneously quantify expression levels of several modified forms of the same protein as these appear as separate spots on 2-D gels allowing separate quantification. This aspect deemed particularly important as neurotrophic factors often induce changes in post-translational modification of target proteins such as phosphorylation [19,26].

Primary cortical neurons of mice were treated with sAPPalpha and proteome alterations following this treatment were determined by statistic evaluation of protein intensity signals on silver-stained 2-D gels using Delta2D software. In detail, cortical neurons were prepared from newborn mice and differentiated in culture for 4 days. Six individual preparations of primary neurons were compared in every group (n = 6, biological replicates) and every treatment condition was compared to controls (neurons treated with medium only). Recombinant sAPPalpha was applied to the cell culture media at a concentration of 300 ng/ml. The applied sAPPalpha concentration was described before to induce ERK phosphorylation in neurons [19] and to increase levels of neuroprotective genes when applied to organotypic hippocampal slice cultures [25].

We treated neurons for one hour and 48 hours, respectively, to cover acute as well as long-term effects. Comparison of treated versus control neurons revealed that 99 protein spots were significantly altered after one hour of sAPPalpha treatment and

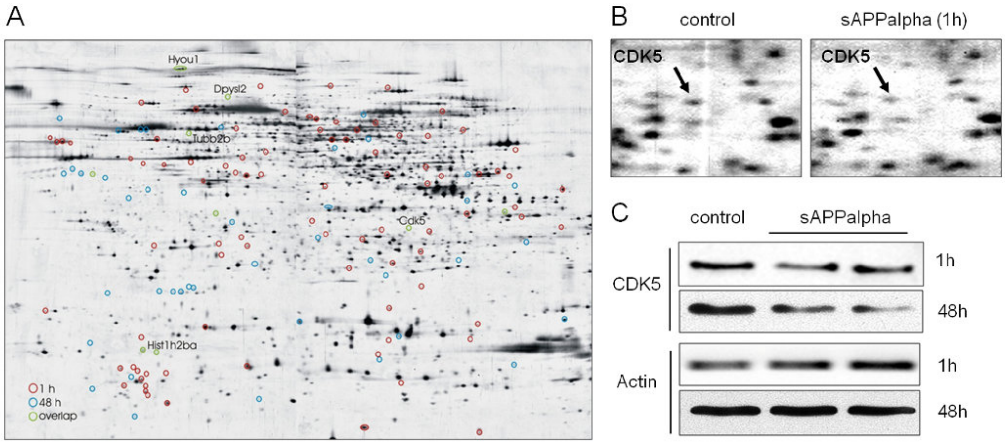

Figure 1. Protein expression changes in neurons treated with sAPPalpha. (**A**) Prototypic two-dimensional polyacrylamid gel of proteins extracted from primary cortical neurons stained with silver nitrate. Significant protein spot alterations in response to sAPPalpha treatment (300 ng/ml) for one hour (red circles), 48 hours (blue circles) or in both treatment conditions (green circles) are indicated (n = 6 biological replicates per genotype; paired Student's t-test $p<0.05$). Gene names of identified proteins altered in both treatment conditions are indicated at respective positions. (**B**) Magnification of gel region containing cyclin-dependent kinase 5 (CDK5). CDK5 is indicated in the non-treated (control) and treated (sAPPalpha) condition. (**C**) Western blots of CDK5 in primary cortical neurons treated with sAPPalpha for one hour or 48 hours, respectively. Treated neurons (sAPPalpha) show reduced expression levels of CDK5 as compared to control neurons (control). Actin served as loading control.
doi:10.1371/journal.pone.0065920.g001

47 protein spots were significantly altered after 48 hours of treatment. Nine proteins were altered under both treatment conditions (paired Student's t-test, $p \leq 0.05$, ratio cut-off $\geq 20\%$; Figure 1A).

Together, proteome analysis revealed a large number of acute protein alterations after sAPPalpha treatment. A lower number of proteins were altered under chronic conditions.

Reduced Expression of CDK5 and Altered Expression of CDK5 Associated Proteins in Neurons Treated with sAPPalpha

In order to identify significantly altered proteins, corresponding protein spots were excised from 2-D gels and identified by mass spectrometry (Table S1). We were most interested in proteins altered under both treatment conditions, as these proteins are most likely particularly relevant to sAPPalpha function. Among proteins altered under both treatment conditions, five proteins could be identified by mass spectrometry. These proteins were collapsin-response mediator protein 2 (CRMP2), histone H2B, hypoxia up-regulated protein 1 (ORP150), beta-tubulin and cyclin-dependent kinase 5 (CDK5). Expression of protein spots corresponding to CRMP2 and ORP150 was up-regulated and expression of spots corresponding to CDK5, histone H2B, and beta-tubulin was down-regulated in treated neurons.

Interestingly, CDK5 and CRMP2 were previously associated with AD pathology [27,28]. Moreover, CRMP2, histone H2B, and beta-tubulin interact with CDK5 [27,29,30]. CDK5 plays a critical role in AD pathomechanisms as CDK5 phosphorylates APP, tau, and BACE1, affecting both hallmarks of AD, amyloid and tau pathology [31]. We therefore decided to analyze CDK5 as sAPPalpha target in more detail.

Using Western blot analysis of primary cortical neurons, we confirmed down-regulation of CDK5 expression after acute as well as long-term treatment with sAPPalpha (Figure 1 C). Moreover, 24% of proteins altered after sAPPalpha treatment (Table S1) are known CDK5 associated proteins (13 out of 55 non-redundant identified proteins; Table 1).

Together, identification of altered proteins revealed that CDK5 might be the key downstream mediator of sAPPalpha signaling in neurons.

CDK5 as well as CDK5 Associated Proteins are not Regulated by sAPPalpha in Neurons Lacking SORLA

SORLA is documented to be an intracellular trafficking receptor regulating APP processing [32]. SORLA directly binds to APP within the d6 domain, which is also an integral part of sAPPalpha [17]. Recent evidence demonstrated that SORLA also binds and mediates internalization of sAPPalpha in neurons [16]. As SORLA is the best-known receptor for sAPPalpha in neurons, we analyzed whether the observed molecular effects of sAPPalpha were dependent on this receptor.

First, we analyzed whether SORLA mediates uptake of recombinant sAPPalpha in our experimental setup. Thus, we treated primary cortical neurons derived from either wild-type or *Sorl1* (the gene coding for SORLA)-deficient mice with sAPPalpha for one hour and quantified the amounts of intracellular recombinant sAPPalpha after treatment. In accordance with the previous study, we found a significant reduction of sAPPalpha signal intensity (about 50%) in *Sorl1*-deficient neurons as compared to wild-type neurons (Figure 2 A & B). This reduction is most likely due to reduced sAPPalpha uptake but it might also be the result of enhanced degradation.

We next asked, whether regulation of CDK5 expression by sAPPalpha was also influenced by SORLA and if other proteins associated with CDK5 function were altered after sAPPalpha treatment. Western blot analysis of primary cortical neurons treated with sAPPalpha revealed that CDK5 and the CDK5

Table 1. CDK5 associated proteins significantly altered in expression after treatment of neurons with sAPPalpha (1 h, 48 h indicate duration of treatment; ↑ up-regulation, ↓ down-regulation, ratio treated/control).

Treatment condition	Direction of regulation	Gene name	Protein name	Comments
1 h, 48 h	↑ (1.34); ↑ (1.37; 1.64); ↑ (1.52)	Crmp1; Dpysl2; Dpysl4	Dihydropyrimidinase-related protein/Collapsin-response mediator protein	Collapsin-response mediator proteins (CRMPs) are involved in apoptosis/ proliferation, cell migration, and differentiation. CRMP2 binds to microtubules and regulates axon outgrowth in neurons. This action is regulated by phosphorylation (via CDK5- and other kinases) at sites hyperphosphorylated in Alzheimer disease [29,35]. CRMPs are altered in expression after treatment of neurons with CDK5 inhibitor [38].
1 h, 48 h	↓ (0.52; 0.6; 0.61); ↑,↓ (0.83; 1.38)	Hist1h2ba; Hist1h4a	Histone H2B type 1-A; Histone H4	CDK5 phosphorylates a component of the histone deacetylase complex and thus regulates histone acetylation i.e. during neuronal cell death. CDK5 can also directly phosphorylates histones [31,51].
48 h	↓ (0.8)	Stxbp1	Syntaxin-binding protein 1 (Munc18-1)	CDK5 promotes Munc18-1 phosphorylation and calcium-dependent exocytosis [52].
48 h	↑ (1.23; 1.35)	Cbx3	Chromobox protein homolog 3	Also known as HP1gamma; repressor of E2F-dependent transcription which is regulated by CDK5 in the nucleus [53,54].
1 h, 48 h	↓ (0.71); ↓ (0.79; 0.81; 0.59)	Tubb2a; Tubb2b	Tubulin beta-2A chain; Tubulin beta-2B chain	CDK5 phosphorylates several tubulin associated proteins regulating tubulin dynamics [32]. CDK5 inhibition alters tubulin expression in neurons [39].
1 h	↑ (1.32); ↓ (0.69); ↑ (1.31)	Ina; Myh10; Npm1	Alpha-Internexin; Myosin-10; Nucleophosmin	CDK5 phosphorylation targets [55,56]. The intermediate-filament protein alpha-internexin was altered in neurons after CDK5 inhibition [38].

doi:10.1371/journal.pone.0033417.t001

adaptor protein p25 (but not the CDK5 adaptor p35) were significantly down-regulated in wild-type neurons after treatment with sAPPalpha for one hour (Mann-Whitney-U test, $p \leq 0.05$, $n = 6$). As p25 is generated from p35 by calpain-dependent cleavage, we also quantified p25/p35 ratios. Also, the ratio of p25/p35 was significantly down-regulated (Figure 3 A & B).

One of the CDK5 target proteins identified in the proteome screen to be significantly regulated upon sAPPalpha treatment was collapsin-response mediator protein 2 (CRMP2). CRMP2 is phosphorylated by CDK5 at Ser522 [33]. Analyzing the phosphorylation state of CRMP2 upon sAPPalpha addition revealed reduced CDK5 activity as phosphorylation of CRMP2 at Ser522 was significantly reduced in treated neurons (Mann-Whitney-U test, $p \leq 0.05$, $n = 6$).

In contrast, CDK5 was not significantly altered in *Sorl1*-deficient neurons after sAPPalpha treatment (Figure 3 A & B). The same finding was seen for expression of the CDK5-target phospho-CRMP2 (Ser522) and the CDK5 adaptor proteins p35 and p25 (Mann-Whitney-U test, $p \leq 0.05$, $n = 6$, Figure 3 A & B). Reduction of phospho-CRMP2 (Ser522) was also observed in non-treated *Sorl1*-deficient as compared to non-treated control neurons. This could be due to the fact that sAPPalpha levels are per se enhanced in *Sorl1*-deficient neurons due to altered APP processing [34].

Together, our results revealed that the expression of CDK5 and CDK5 associated proteins was altered after sAPPalpha treatment in neurons. Reduced phosphorylation of CRMP documented impaired CDK5 activity after sAPPalpha application. None of these alterations were detected in *Sorl1*-deficient neurons proving that SORLA is an essential sAPPalpha receptor.

Induction of Neuroprotective ORP150 by sAPPalpha

Up-regulation of ORP150 was previously shown to protect neurons from hypoxia and excitotoxicity [35,36]. Interestingly, expression of ORP150 was induced by sAPPalpha treatment under both treatment conditions (1 hour and 48 hours of treatment; Table S1). ORP150 might therefore be involved in sAPPalpha-mediated neuroprotection.

Western blot analysis of neurons treated with sAPPalpha confirmed significant up-regulation of ORP150 after sAPPalpha

Figure 2. Reduced uptake of sAPPalpha in SORLA-deficient neurons. Quantification of human, recombinant sAPPalpha in primary cortical neurons either non-treated (control) or treated with 300 ng/ml sAPPalpha for 1 h using Western blotting (**A**) and densitometric scanning of replicate blots (**B**). *Sorl1*$^{-/-}$ neurons show reduced levels sAPPalpha after one hour of treatment compared with wild-type cells (wt; $n = 8$, Mann-Whitney U test). Actin served as loading control.
doi:10.1371/journal.pone.0065920.g002

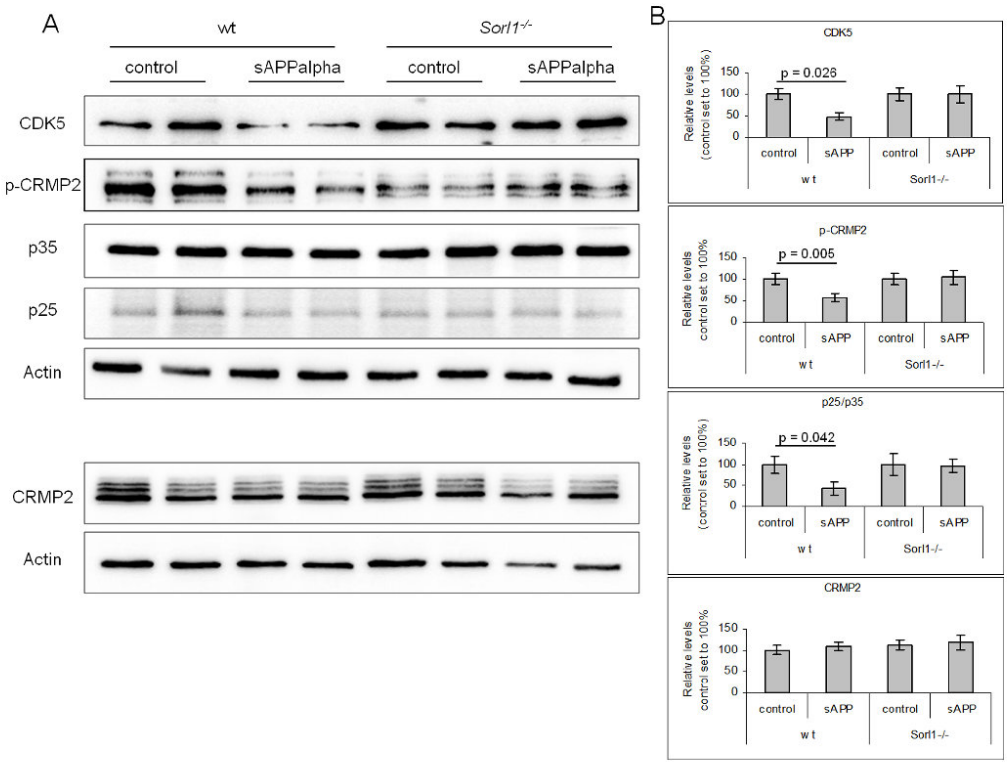

Figure 3. Reduction of CDK5 activity in neurons treated with sAPPalpha is SORLA-dependent. Quantification of CDK5, CDK5-target phospho-CRMP2, total CRMP2, and CDK5-adaptor proteins p35 and p25 in primary cortical neurons either non-treated (control) or treated with 300 ng/ml sAPPalpha for 1 h using Western blotting (**A**) and densitometric scanning of replicate blots (**B**). In wild- type neurons (wt), CDK5, p25 and phospho-CRMP2 (p-CRMP2) were significantly down-regulated after sAPPalpha treatment. *Sorl1* $^{-/-}$ neurons show no altered levels of the same proteins after one hour of sAPPalpha treatment (n = 6, Mann-Whitney U test). Actin served as loading control.
doi:10.1371/journal.pone.0065920.g003

treatment of neurons. Again, this effect was not observed in neurons lacking the sAPPalpha receptor SORLA (Mann-Whitney-U test, $p \leq 0.05$, n = 6; Figure 4 A & B). Together, ORP150, a potential effector protein of sAPPalpha in neurons, was induced after sAPPalpha treatment. This effect was SORLA-dependent.

Discussion

The APP cleavage product sAPPalpha is considered to be neurotrophic and neuroprotective [6–11,24,25]. Thus, loss of sAPPalpha activity in AD patients might contribute to disease pathology. However, little is known about molecular pathways underlying these effects. Using a 2-DE based proteomic approach, we uncovered that sAPPalpha significantly reduces expression and activity of CDK5 and influences expression of CDK5 target proteins in neurons. Interestingly, many of these CDK5 target proteins (cofilin, beta-tubulin, alpha-internexin and CRMP; table 1) were shown before to be altered in neurons treated with CDK5 inhibitors [37,38]. This finding further supports a role of sAPPalpha in regulating CDK5 signaling.

CDK5 is an atypical cyclin kinase as -in contrast to other cyclin kinases- it inhibits the cell cycle to keep neurons in their post-mitotic stage. Transfer of CDK5 out of the nucleus induces neurodegeneration [39]. CDK5 also regulates neuronal morphology via phosphorylation of central components of the cellular cytoskeleton, such as tubulin, and tau [40]. Because CDK5 phosphorylates tau, APP, and BACE1, a component of the beta-secretase, CDK5 is an important link between amyloid- and tau-pathology [31]. Regulation of CDK5 activity by sAPPalpha was suggested before, as overexpression of sAPPalpha inhibited glutamate-induced CDK5 activation in N2a cells [41]. In contrast to sAPPalpha, Abeta activates CDK5 (reviewed in [42]). Aberrant activation of CDK5 can lead to collapse of the synaptic cytoskeleton [43]. sAPPalpha potentially protects from this collapse, as it maintains synaptic integrity [44]. Together, regulation of CDK5 activity in neurons might be the most important determinant of how the two alternative APP cleavage products Abeta and sAPPalpha counteract each other.

We also identified expression changes of the CDK5 phosphorylation target CRMP2 in neurons treated with sAPPalpha. Interestingly, CRMP2 phosphorylation impairs

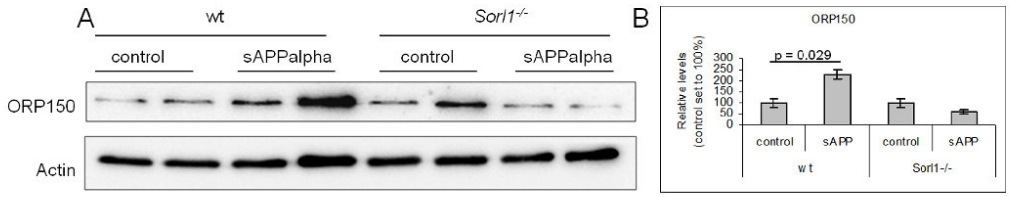

Figure 4. Induction of ORP150 in sAPPalpha treated neurons. Quantification of ORP150 in primary cortical neurons either non-treated (control) or treated with 300 ng/ml sAPPalpha for 1 h using Western blotting (**A**) and densitometric scanning of replicate blots (**B**). In wild-type neurons (wt), ORP150 is significantly up-regulated after sAPPalpha treatment. No alteration in ORP150 expression was observed in *Sorl1*$^{-/-}$ neurons after one hour of treatment (n = 6, Mann-Whitney U test). Actin served as loading control.
doi:10.1371/journal.pone.0065920.g004

neurite outgrowth, an effect that can be reversed by inhibition of CDK5 [45]. We observed reduced phosphorylation of CRMP2 at the CDK5-dependent phosphorylation site (Ser522) in neurons treated with sAPPalpha. We further identified down-regulation of p25 after sAPPalpha treatment of neurons. p25 is a CDK5 adaptor protein that accumulates in the brains of AD patients leading to abnormal CDK5 activation and hyperphosphorylation of CDK5 targets, such as tau and CRMP2 [27,42]. Together, down-regulation of p25 and phospho-CRMP2 demonstrate reduced activity of CDK5 in sAPPalpha treated neurons and further support the important role of sAPPalpha as neuroprotective factor.

One downstream effector protein providing neuroprotection might be ORP150, a protein induced in neurons after sAPPalpha treatment. ORP150 is an endoplasmic reticulum associated chaperone. Interestingly, induction of ORP150 protein has been demonstrated before in cultured neurons under hypoxic and excitotoxic stress conditions and in the brains of patients who died of epileptic seizures. However, overexpression of ORP150 protected neurons from hypoxia-induced cell death and mediated higher resistance of mice towards cerebral ischemia demonstrating that ORP150 protected cells from excitotoxicity [35,36]. Interestingly, sAPPalpha has been reported to protect neurons from excitotoxicity as well [41]. Based on our findings we now speculate that induction of ORP150 by sAPPalpha might mediate this effect.

Finally, we provide functional evidence concerning the importance of SORLA as sAPPalpha receptor mediating not only uptake of sAPPalpha, but also controlling regulation of downstream sAPPalpha targets. Direct binding of SORLA to sAPPalpha and uptake of sAPPalpha by SORLA was shown before [16,17]. We now document that SORLA is essential for the molecular function of sAPPalpha, as regulation of CDK5, phospho-CRMP2, p25 and ORP150 was no longer evident in neurons deficient for SORLA. As SORLA itself is considered an important genetic risk factor in AD [46] and low levels of this receptor have been documented in AD patients [47], it is likely that part of the neuroprotective properties ascribed to SORLA are related to its function as sAPPalpha receptor. Concerning its neuroprotective capabilities, so far SORLA was mainly recognized for its ability to regulate APP processing in the first place [20,34,48], but based in findings in this study SORLA might also serve as receptor of APP processing products controlling their respective cellular functions. That is at least the case for sAPPalpha.

Conclusion

Our study provides insight into the molecular pathways downstream of sAPPalpha in neurons. We identified CDK5 and ORP150 as potential mediators of sAPPalpha-dependent neuroprotection. Moreover, we demonstrate a central role for SORLA as the sAPPalpha receptor; loss of sAPPalpha function in AD is a plausible mechanism of loss of sAPPalpha function in AD [44]. Here, we provide a proteomic-based evidence of how APP processing regulates neuronal function. We further propose that proteins regulated by sAPPalpha might be potential drug targets in their own right.

Supporting Information

Table S1 Proteins identified as significantly altered in expression after 1 h or 48 h incubation with sAPPalpha. Proteins altered in both treatment conditions (1 h and 48 h) are highlighted in grey. (PDF)

Acknowledgments

We are indebted to A. Krajewski, Y. Kläre and M. Herrmann for expert technical assistance. We also thank O. Klein and G. Nebrich for mass spectrometry analyses.

Author Contributions

Conceived and designed the experiments: DH M. Rohe. Performed the experiments: DH M. Roch SK. Analyzed the data: DH SK M. Roch ZK JK TEW M. Rohe. Contributed reagents/materials/analysis tools: SK ZK. Wrote the paper: DH M. Rohe. Provided recombinant sAPPalpha:SK. All authors gave final approval of the version to be published.

References

1. Chasseigneaux S, Allinquant B (2012) Functions of Abeta, sAPPalpha and sAPPbeta : similarities and differences. J Neurochem 120 Suppl 1: 99–108.
2. Sennvik K, Fastbom J, Blomberg M, Wahlund LO, Winblad B, et al. (2000) Levels of alpha- and beta-secretase cleaved amyloid precursor protein in the cerebrospinal fluid of Alzheimer's disease patients. Neurosci Lett 278: 169–172.
3. Lannfelt L, Basun H, Wahlund LO, Rowe BA, Wagner SL (1995) Decreased alpha-secretase-cleaved amyloid precursor protein as a diagnostic marker for Alzheimer's disease. Nat Med 1: 829–832.
4. Olsson A, Hoglund K, Sjogren M, Andreasen N, Minthon L, et al. (2003) Measurement of alpha- and beta-secretase cleaved amyloid precursor protein in cerebrospinal fluid from Alzheimer patients. Exp Neurol 183: 74–80.
5. Ring S, Weyer SW, Kilian SB, Waldron E, Pietrzik CU, et al. (2007) The secreted beta-amyloid precursor protein ectodomain APPs alpha is sufficient to rescue the anatomical, behavioral, and electrophysiological abnormalities of APP-deficient mice. J Neurosci 27: 7817–7826.
6. Bell KF, Zheng L, Fahrenholz F, Cuello AC (2008) ADAM-10 over-expression increases cortical synaptogenesis. Neurobiol Aging 29: 554–565.

7. Meziane H, Dodart JC, Mathis C, Little S, Clemens J, et al. (1998) Memory-enhancing effects of secreted forms of the beta-amyloid precursor protein in normal and amnestic mice. Proc Natl Acad Sci U S A 95: 12683–12688.
8. Ohsawa I, Takamura C, Kohsaka S (1997) The amino-terminal region of amyloid precursor protein is responsible for neurite outgrowth in rat neocortical explant culture. Biochem Biophys Res Commun 236: 59–65.
9. Araki W, Kitaguchi N, Tokushima Y, Ishii K, Aratake H, et al. (1991) Trophic effect of beta-amyloid precursor protein on cerebral cortical neurons in culture. Biochem Biophys Res Commun 181: 265–271.
10. Qiu WQ, Ferreira A, Miller C, Koo EH, Selkoe DJ (1995) Cell-surface beta-amyloid precursor protein stimulates neurite outgrowth of hippocampal neurons in an isoform-dependent manner. J Neurosci 15: 2157–2167.
11. Mattson MP, Cheng B, Culwell AR, Esch FS, Lieberburg I, et al. (1993) Evidence for excitoprotective and intraneuronal calcium-regulating roles for secreted forms of the beta-amyloid precursor protein. Neuron 10: 243–254.
12. Saitoh T, Sundsmo M, Roch JM, Kimura N, Cole G, et al. (1989) Secreted form of amyloid beta protein precursor is involved in the growth regulation of fibroblasts. Cell 58: 615–622.
13. Jin LW, Ninomiya H, Roch JM, Schubert D, Masliah E, et al. (1994) Peptides containing the RERMS sequence of amyloid beta/A4 protein precursor bind cell surface and promote neurite extension. J Neurosci 14: 5461–5470.
14. Bhasin R, Van Nostrand WE, Saitoh T, Donets MA, Barnes EA, et al. (1991) Expression of active secreted forms of human amyloid beta-protein precursor by recombinant baculovirus-infected insect cells. Proc Natl Acad Sci U S A 88: 10307–10311.
15. Hoey SE, Williams RJ, Perkinton MS (2009) Synaptic NMDA receptor activation stimulates alpha-secretase amyloid precursor protein processing and inhibits amyloid-beta production. J Neurosci 29: 4442–4460.
16. Gustafsen C, Glerup S, Pallesen LT, Olsen D, Andersen OM, et al. (2013) Sortilin and SorLA Display Distinct Roles in Processing and Trafficking of Amyloid Precursor Protein. J Neurosci 33: 64–71.
17. Andersen OM, Schmidt V, Spoelgen R, Gliemann J, Behlke J, et al. (2006) Molecular dissection of the interaction between amyloid precursor protein and its neuronal trafficking receptor SorLA/LR11. Biochemistry 45: 2618–2628.
18. Rohe M, Synowitz M, Glass R, Paul SM, Nykjaer A, et al. (2009) Brain-derived neurotrophic factor reduces amyloidogenic processing through control of SORLA gene expression. J Neurosci 29: 15472–15478.
19. Klatt S, Rohe M, Alagesan K, Kolarich D, Konthur Z, et al. (2013) Production of Glycosylated Soluble Amyloid Precursor Protein Alpha (sAPPalpha) in Leishmania tarentolae. J Proteome Res 12: 396–403.
20. Andersen OM, Reiche J, Schmidt V, Gotthardt M, Spoelgen R, et al. (2005) Neuronal sorting protein-related receptor sorLA/LR11 regulates processing of the amyloid precursor protein. Proc Natl Acad Sci U S A 102: 13461–13466.
21. Rohe M, Nebrich G, Klein O, Mao L, Zabel C, et al. (2011) Kainate promotes alterations in neuronal RNA splicing machinery. J Proteome Res 10: 1459–1467.
22. Zabel C, Klose J (2009) High-resolution large-gel 2DE. Methods Mol Biol 519: 311–338.
23. Zabel C, Chamrad DC, Priller J, Woodman B, Meyer HE, et al. (2002) Alterations in the mouse and human proteome caused by Huntington's disease. Mol Cell Proteomics 1: 366–375.
24. Postina R, Schroeder A, Dewachter I, Bohl J, Schmitt U, et al. (2004) A disintegrin-metalloproteinase prevents amyloid plaque formation and hippocampal defects in an Alzheimer disease mouse model. J Clin Invest 113: 1456–1464.
25. Stein TD, Anders NJ, DeCarli C, Chan SL, Mattson MP, et al. (2004) Neutralization of transthyretin reverses the neuroprotective effects of secreted amyloid precursor protein (APP) in APPSW mice resulting in tau phosphorylation and loss of hippocampal neurons: support for the amyloid hypothesis. J Neurosci 24: 7707–7717.
26. Reichardt LF (2006) Neurotrophin-regulated signalling pathways. Philos Trans R Soc Lond B Biol Sci 361: 1545–1564.
27. Williamson R, van Aalten L, Mann DM, Platt B, Plattner F, et al. (2011) CRMP2 hyperphosphorylation is characteristic of Alzheimer's disease and not a feature common to other neurodegenerative diseases. J Alzheimers Dis 27: 615–625.
28. Crews L, Masliah E (2010) Molecular mechanisms of neurodegeneration in Alzheimer's disease. Hum Mol Genet 19: R12–20.
29. Kim D, Frank CL, Dobbin MM, Tsunemoto RK, Tu W, et al. (2008) Deregulation of HDAC1 by p25/Cdk5 in neurotoxicity. Neuron 60: 803–817.
30. Hou Z, Li Q, He L, Lim HY, Fu X, et al. (2007) Microtubule association of the neuronal p35 activator of Cdk5. J Biol Chem 282: 18666–18670.
31. Cheung ZH, Ip NY (2012) Cdk5: a multifaceted kinase in neurodegenerative diseases. Trends Cell Biol 22: 169–175.
32. Willnow TE, Carlo AS, Rohe M, Schmidt V (2010) SORLA/SORL1, a neuronal sorting receptor implicated in Alzheimer's disease. Rev Neurosci 21: 315–329.
33. Uchida Y, Ohshima T, Sasaki Y, Suzuki H, Yanai S, et al. (2005) Semaphorin3A signalling is mediated via sequential Cdk5 and GSK3beta phosphorylation of CRMP2: implication of common phosphorylating mechanism underlying axon guidance and Alzheimer's disease. Genes Cells 10: 165–179.
34. Rohe M, Carlo AS, Breyhan H, Sporbert A, Militz D, et al. (2008) Sortilin-related receptor with A-type repeats (SORLA) affects the amyloid precursor protein-dependent stimulation of ERK signaling and adult neurogenesis. J Biol Chem 283: 14826–14834.
35. Tamatani M, Matsuyama T, Yamaguchi A, Mitsuda N, Tsukamoto Y, et al. (2001) ORP150 protects against hypoxia/ischemia-induced neuronal death. Nat Med 7: 317–323.
36. Kitao Y, Ozawa K, Miyazaki M, Tamatani M, Kobayashi T, et al. (2001) Expression of the endoplasmic reticulum molecular chaperone (ORP150) rescues hippocampal neurons from glutamate toxicity. J Clin Invest 108: 1439–1450.
37. Gillardon F, Steinlein P, Burger E, Hildebrandt T, Gerner C (2005) Phosphoproteome and transcriptome analysis of the neuronal response to a CDK5 inhibitor. Proteomics 5: 1299–1307.
38. Gillardon F, Schrattenholz A, Sommer B (2005) Investigating the neuroprotective mechanism of action of a CDK5 inhibitor by phosphoproteome analysis. J Cell Biochem 95: 817–826.
39. Zhang J, Herrup K (2011) Nucleocytoplasmic Cdk5 is involved in neuronal cell cycle and death in post-mitotic neurons. Cell Cycle 10: 1208–1214.
40. Lalioti V, Pulido D, Sandoval IV (2010) Cdk5, the multifunctional surveyor. Cell Cycle 9: 284–311.
41. Han P, Dou F, Li F, Zhang X, Zhang YW, et al. (2005) Suppression of cyclin-dependent kinase 5 activation by amyloid precursor protein: a novel excitoprotective mechanism involving modulation of tau phosphorylation. J Neurosci 25: 11542–11552.
42. Shukla V, Skuntz S, Pant HC (2012) Deregulated cdk5 activity is involved in inducing Alzheimer's disease. Arch Med Res 43: 655–662.
43. Roselli F, Livrea P, Almeida OF (2011) CDK5 is essential for soluble amyloid beta-induced degradation of GKAP and remodeling of the synaptic actin cytoskeleton. PLoS One 6: e23097.
44. Tyan SH, Shih AY, Walsh JJ, Maruyama H, Sarsoza F, et al. (2012) Amyloid precursor protein (APP) regulates synaptic structure and function. Mol Cell Neurosci 51: 43–52.
45. Crews L, Ruf R, Patrick C, Dumaop W, Trejo-Morales M, et al. (2011) Phosphorylation of collapsin response mediator protein-2 disrupts neuronal maturation in a model of adult neurogenesis: Implications for neurodegenerative disorders. Mol Neurodegener 6: 67.
46. Rogaeva E, Meng Y, Lee JH, Gu Y, Kawarai T, et al. (2007) The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease. Nat Genet 39: 168–177.
47. Scherzer CR, Offe K, Gearing M, Rees HD, Fang G, et al. (2004) Loss of apolipoprotein E receptor LR11 in Alzheimer disease. Arch Neurol 61: 1200–1205.
48. Offe K, Dodson SE, Shoemaker JT, Fritz JJ, Gearing M, et al. (2006) The lipoprotein receptor LR11 regulates amyloid beta production and amyloid precursor protein traffic in endosomal compartments. J Neurosci 26: 1596–1603.
49. Sharma P, Steinbach PJ, Sharma M, Amin ND, Barchi JJ, Jr., et al. (1999) Identification of substrate binding site of cyclin-dependent kinase 5. J Biol Chem 274: 9600–9606.
50. Lilja L, Johansson JU, Gromada J, Mandic SA, Fried G, et al. (2004) Cyclin-dependent kinase 5 associated with p39 promotes Munc18–1 phosphorylation and Ca(2+)-dependent exocytosis. J Biol Chem 279: 29534–29541.
51. Sdek P, Zhao P, Wang Y, Huang CJ, Ko CY, et al. (2011) Rb and p130 control cell cycle gene silencing to maintain the postmitotic phenotype in cardiac myocytes. J Cell Biol 194: 407–423.
52. Futatsugi A, Utreras E, Rudrabhatla P, Jaffe H, Pant HC, et al. (2012) Cyclin-dependent kinase 5 regulates E2F transcription factor through phosphorylation of Rb protein in neurons. Cell Cycle 11: 1603–1610.
53. Tripathi BK, Zelenka PS (2009) Cdk5-dependent regulation of Rho activity, cytoskeletal contraction, and epithelial cell migration via suppression of Src and p190RhoGAP. Mol Cell Biol 29: 6488–6499.
54. Ramos-Echazabal G, Chinea G, Garcia-Fernandez R, Pons T (2012) In silico studies of potential phosphoresidues in the human nucleophosmin/B23: its kinases and related biological processes. J Cell Biochem 113: 2364–2374.

Supplementary table 1: Proteins identified as significantly altered in expression after 1 h or 48 h incubation with sAPPalpha. Proteins altered in both treatment conditions (1 h and 48 h) are highlighted in grey.

treatment condition	p-value	ratio (treated / control)	gene name	protein name	Unique peptides	Sequence coverage (%)
1h	0.048	1.240	Acaa1a	3-ketoacyl-CoA thiolase A, peroxisomal	8	19
1h	0.025	1.338	Aco2	Aconitate hydratase, mitochondrial	2	2
1h	0.025	1.391	Actb	Actin, cytoplasmic 1	6	22
1h	0.007	1.304	Actr1a	Alpha-centractin	6	18
1h	0.021	0.819	Arl3	ADP-ribosylation factor-like protein 3	3	19
1h	0.034	1.208	Atp5a1	ATP synthase subunit alpha, mitochondrial	20	33
1h	0.039	1.206	Atp5b	ATP synthase subunit beta, mitochondrial	19	31
1h	0.006	1.814	Atp5b	ATP synthase subunit beta, mitochondrial	4	13
1h	0.023	1.528	Atp6v1a	V-type proton ATPase catalytic subunit A	12	20
1h	0.012	0.779	Basp1	Brain acid soluble protein 1	8	58
48h	0.033	1.234	Cbx3	Chromobox protein homolog 3 (HP1-gamma)	2	7
48h	0.006	1.349	Cbx3	Chromobox protein homolog 3 (HP1-gamma)	2	7
1h	0.037	0.571	Cdk5	Cell division protein kinase 5	5	14
48h	0.050	0.751	Cdk5	Cell division protein kinase 5	5	14
48h	0.017	0.674	Cfl1	Cofilin-1	2	18
1h	0.004	1.537	Cops4	COP9 signalosome complex subunit 4	9	24
1h	0.010	1.369	Crmp1	Dihydropyrimidinase-related protein 1	2	5
1h	0.010	1.369	Dpysl2	Dihydropyrimidinase-related protein 2 (CRMP2)	2	5
48h	0.002	1.645	Dpysl2	Dihydropyrimidinase-related protein 2 (CRMP2)	5	12
1h	0.038	1.523	Dpysl4	Dihydropyrimidinase-related protein 4	2	5
1h	0.021	0.627	Eef1a1	Elongation factor Tu, mitochondrial	9	19
1h	0.042	2.060	Eef2	Elongation factor 2	9	10
1h	0.023	0.702	Endogl1	Nuclease EXOG, mitochondrial	4	10
1h	0.029	1.254	Erlin2	Erlin-2	9	22
1h	0.045	0.793	Fabp7	Fatty acid-binding protein, brain	3	31
48h	0.039	0.747	Fkbp2	FK506-binding protein 2	3	18
1h	0.039	1.328	Fscn1	Fascin	8	19
1h	0.033	0.828	Got2	Aspartate aminotransferase, mitochondrial	10	25
1h	0.035	0.525	Hist1h2ba	Histone H2B type 1-A	2	14
1h	0.005	0.598	Hist1h2ba	Histone H2B type 1-A	2	15
48h	0.026	0.613	Hist1h2ba	Histone H2B type 1-A	2	14
1h	0.043	0.826	Hist1h4a	Histone H4	2	19
1h	0.027	1.376	Hist1h4a	Histone H4	2	19
1h	0.023	1.643	Hmgcs1	Hydroxymethylglutaryl-CoA synthase, cytoplasmic	7	12
1h	0.036	0.825	Hnrnpa1	Heterogeneous nuclear ribonucleoprotein A1	8	29
1h	0.040	0.806	Hnrnpa1	Heterogeneous nuclear ribonucleoprotein A1	8	31
1h	0.023	0.721	Hnrnpa2b1	Heterogeneous nuclear ribonucleoproteins A2/B1	10	33
1h	0.008	1.250	Hsd17b8	Estradiol 17-beta-dehydrogenase 8	2	10
48h	0.025	1.440	Hyou1	Hypoxia up-regulated 1 (ORP150)	4	4
1h	0.043	1.246	Hyou1	Hypoxia up-regulated 1 (ORP150)	4	4
1h	0.011	1.253	Idh3a	Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial	10	29
1h	0.013	1.318	Ina	Alpha-Internexin	5	2
1h	0.007	0.775	Mtpn	Myotrophin	2	27
1h	0.046	0.693	Myh10	Myosin-10	3	2
1h	0.013	1.215	Ndufaf2	Mimitin, mitochondrial	8	51
1h	0.045	1.481	Ndufs3	NADH dehydrogenase [ubiquinone] iron-sulfur protein 3, mitochondrial	4	12
48h	0.016	0.801	Ndufs4	NADH dehydrogenase [ubiquinone] iron-sulfur protein 4, mitochondrial	4	25
1h	0.038	1.310	Npm1	Nucleophosmin	6	19
1h	0.002	1.263	Oxct1	Succinyl-CoA:3-ketoacid-coenzyme A transferase 1, mitochondrial	3	7
1h	0.050	0.816	Pcbp2	Poly(rC)-binding protein 2	2	8
1h	0.024	1.578	Pdcd6ip	Programmed cell death 6-interacting protein	2	3
1h	0.040	0.739	Pea15a	Astrocytic phosphoprotein PEA-15	2	22
1h	0.043	1.585	Ppia	Peptidyl-prolyl cis-trans isomerase A	5	37
1h	0.037	1.211	Psmd7	26S proteasome non-ATPase regulatory subunit 7	6	17
1h	0.045	0.821	Sh3bgrl	SH3 domain-binding glutamic acid-rich-like protein	5	44
1h	0.036	0.825	Slc25a22	Mitochondrial glutamate carrier 1	8	20
1h	0.029	1.692	Spr	Sepiapterin reductase	2	11
48h	0.048	0.803	Stxbp1	Syntaxin-binding protein 1	2	3
1h	0.022	0.787	Suclg1	Succinyl-CoA ligase [GDP-forming] subunit alpha, mitochondrial	3	11
1h	0.033	0.795	Tomm22	Mitochondrial import receptor subunit TOM22 homolog	2	8
1h	0.003	1.274	Trap1	Heat shock protein 75 kDa, mitochondrial	15	24
1h	0.048	0.708	Tubb2a	Tubulin beta-2A chain	2	6
1h	0.005	0.787	Tubb2b	Tubulin beta-2B chain	2	6
1h	0.025	0.818	Tubb2b	Tubulin beta-2B chain	5	12
48h	0.000	0.594	Tubb2b	Tubulin beta-2B chain	2	6
1h	0.028	0.731	Uqcrh	Cytochrome b-c1 complex subunit 6, mitochondrial	2	24

3.3 Manuscript IV

Generation and characterization of a *Leishmania tarentolae* strain for site-directed in-vivo biotinylation of recombinant proteins

Klatt S, Hartl D, Fauler B, Gagoski D, Castro-Obregón S, Konthur Z.

Journal of Proteome Research, October 7, 2013. PMID: 24093329

My contribution to this manuscript:

My experimental contributions to this manuscript include the cloning of all genes as well as the expression and purification of all tagged proteins except mTKIN. Additionally, I have designed and performed the autophagy assay, assisted with the TEM results interpretation, helped and partially performed the 2-D PAGE and western blots. Last but not least, I have written the manuscript with Dr. Zoltán Konthur.

4 Discussion

In my thesis, the benefit of the protozoan host *L. tarentolae* for biomedical research was evaluated. In detail, its ability to produce biomedically relevant mammalian recombinant proteins of high homogeneity to the appropriate counterpart was analysed. I conducted the following experiments to enable protein expression and to fulfil my aim: First, new expression vectors were designed to form the experimental basis. Second, to increase the expression-secretion efficiency of secretory proteins, modification of an existing SP was performed and firstly tested with human recombinant antibody fragments (scFv's). Third, to confirm mammalian-like PTM's of recombinant proteins expressed with *L. tarentolae*, the highly glycosylated sAPPalpha was selected and analysed in detail. Fourth, an in-vivo biotinylation system was established to enable one-step purification of intracellular/cytoplasmic recombinant proteins preferably resulting in a pure product.

Furthermore, all manuscripts including all successfully expressed recombinant proteins will be presented here. Finally, an outlook will be given which discusses the possibility to further increase the similarity of expressed recombinant proteins to their originally counterparts. Also it will be examined, why this could expand the community working with this imposing and interesting organism.

Preliminary work

Successful recombinant protein expression requires the knowledge about the natural properties of the target protein as well as of the expression host [150]. Commonly asked questions include: What is the natural function of the protein? Does it have interaction partners? What about its final destination or natural abundance? Is the host able to mimic the natural surrounding and can it provide the protein with adequate PTM? What about protein purification; which affinity tag is best? The more facts one knows, the higher the probability to produce a biologically active recombinant protein. Additionally, since the genome of *L. tarentolae* was sequenced [26], it can be screened for similar protein motifs to allow a first evaluation of metabolic reactions or cross-reactivities. Finally, the choice of the right cloning strategy in combination with the right purification method affects protein expression/-secretion efficacy or rather the protein yield [151, 152].

The culture conditions of *L. tarentolae* (subchapter 1.2.3) as well as different expression vector variants (subchapter 1.2.5) are already well-established. Furthermore, *L. tarentolae* is generally able to posttranslationally modify human recombinant proteins in a mammalian-like manner (subchapter 1.2.4) and is therefore well-suited for the expression of further human proteins.

Based on this preliminary work, I analysed potential protein candidates and adapted available *L. tarentolae* expression vectors to secure and/or to increase their successful expression. Of note: the sequence of the *L. tarentolae* genome was only published in 2012 and therefore the analysis based on genomic data was conducted post expression.

Target proteins

Anti-cancer recombinant monoclonal antibodies account for around one third of the recombinant protein's turnover [7]. Thus, their biomedical relevance is extremely high. In general, antibodies are produced by B-cells of the immune system and bind specifically to intruders marking them for neutralization. For this, they are naturally not occurring in Leishmania parasites, potentially minimizing cross-reactivities with endogenous proteins. Furthermore, they are secretory proteins and are often modified with glycans [153, 154]. To evaluate the general ability of *L. tarentolae* for human antibody expression, I decided to firstly clone and express not full-length Ig's, but their smallest functional entity comprising of the variable domain of the heavy and the light chain as a single-polypeptide known as scFv's. This antibody fragments have a size of around 27kDa and are rarely glycosylated. Here, the four chosen scFv's are derived from a semi-synthetic single-framework phage display antibody library (Manuscript I). In addition, the scFv's and all further expressed recombinant proteins, except the biotin ligase, were linked to a hexahistidine-tag (His_6-tag). The His_6-tag was selected for purification, because this peptide is very short, possesses low cross-reactivity to endogenous proteins, has a highly selective interaction with Ni^{2+}-NTA and the His_6-Ni^{2+}-NTA purification system has been tested with a large number of applications, allowing one-step purification [155-157]. Next, the highly glycosylated human secretory protein sAPPalpha (68.6kDa) was chosen to verify mammalian-like glycosylation characteristics of *L. tarentolae* (Manuscript II+III). sAPPalpha is a good predictor for AD and other neurodegenerative diseases [158, 159] and therefore also relevant in biomedical research. Additionally, two more human proteins were selected as expression candidates: The cellular tumour antigen p53 (p53) as well as the farnesyl-diphosphate farnesyltransferase (FDFT1). p53 (53kDa) is a widespread multifunctional transcription factor

and very well studied [160, 161]. For example, it acts as tumor suppressor in many tumor types, can induce growth arrest, apoptosis or autophagy [162]. FDFT1 (48.1kDa) is a monomeric transferase and important for several biosynthetic pathways like the one of cholesterol [163]. Further selected proteins are the mammalian codon-optimized biotin ligase (BirA) from *E. coli* [164], the enhanced green fluorescent protein (eGFP) derived from the jelly fish *Aequorea victoria* and a cysteine-free mutant of kinesin (mTKIN) from the thermophilic fungus *Thermomyces lanuginosus*. BirA (35.3kDa) is an enzyme that biotinylates specific protein motifs *in-vivo* and is used here as detection marker for further co-expressed recombinant proteins (Manuscript IV). It was expressed without any affinity-tag. eGFP (26.9kDa) is used as marker protein to generally test transfection and expression success. Last but not least, mTKIN (55kDa) is a motor protein of the kinesin family and used as another expression example [165, 166]. By screening the genome of *L. tarentolae* [26] using the platform TriTrypDB (www.tritrypdb.org), endogenous counterparts to FDFT1 (putative farnesyltransferase, 47.5kDa; LtaP31.3430), BirA (biotin/lipoate ligase-like protein, 28.6kDa; LtaP31.1020) and mTKIN (eight putative kinesin like proteins/kinesins, >100kDa; e.g. LtaP22.0920 or LtaP19.0670) were revealed. Thus, interaction between the endogenous protein and its recombinant counterpart is generally possible, but it was not pursued any further. Moreover, their similarity significantly differed and no differences in the growth rate or phenotype of *L. tarentolae* clones could be observed. More information about the proteins can be found on www.uniprot.org or www.phosphosite.org. Altogether, these seven recombinant proteins of different origin, localization, size and function were successfully expressed in *L. tarentolae* (Table 2).

Table 2: Summary of all successfully expressed recombinant proteins in L. tarentolae.

Protein	Origin	Localization	Post-translational Modification	Main Function	Expression vector	Manu-script
Single chain fragment variables **(scFv's)**	Homo sapiens	secretory	/ (glycosylation)	Antigen binders	pLTEX-2 to pLTEX-5	I
Soluble amyloid precursor protein alpha **(sAPPalpha)**	Homo sapiens	secretory	Phospho-rylation, Glycosylation, Disulphide Bond	Alzheimer's disease marker	pLTEX-5	II, III
Enhanced green fluorescent protein **(eGFP)**	*Aequorea victoria*	cytoplasmic	/	Localization marker, expression control	pLEXSY_ sat2, pLTEX-1	IV

Mammalian codon-optimized biotin ligase **(BirA)**	*E. coli*	cytoplasmic	/	In-vivo biotinylation of special proteins	pLEXSY_ hyg2	IV
Cellular tumor antigen p53 **(p53)**	Homo sapiens	cytoplasmic	Phospho-rylation, Acetylation, Methylation, …	Multi-functional; autophagy inducer	pLTEX-1	IV
Farnesyl-diphosphate farnesyltransfe rase **(FDFT1)**	Homo sapiens	ER-membrane/ cytoplasmic	Phospho-rylation, Ubiquitination	Enzymatic active in cholesterol pathway	pLTEX-1	IV
Cysteine-free mutant of kinesin **(mTKIN)**	Thermo-myces lanuginosus	cytoplasmic	?	Motor-protein, intracellular transport	pLTEX-1	IV

Development of new expression vectors to improve recombinant protein expression and secretion

To generally enable expression of above mentioned proteins in *L. tarentolae*, different expression vector variants are available. The two vectors pLEXSY_hyg2 and pLEXSY_sat2 from the constitutive system of JBS were one of the first established vectors (subchapter 1.2.5). Thus, they served here as the basis. The cloning strategy allows continuous expression of either intracellular or secretory target proteins unless the rising concentration of the protein may not disturb metabolic equilibrium (homeostasis) of the parasite, resulting in e.g. insoluble protein aggregates, gene silencing or cell apoptosis [12, 110].

Vector pLEXSY_hyg2 was only used to clone the BirA gene into *L. tarentolae* to generate the biotinylation strain P10-BirA (Manuscript IV) and also to allow the integration and expression of a co-expressed second construct. Parental vector pLEXSY_sat2 was used to generate the vector series pLTEX-1 to pLTEX-5 (Manuscript I). Vector pLTEX-1 is used for the expression of cytoplasmic recombinant proteins (here: eGFP, p53, FDFT1 and mTKIN). Following, vectors pLTEX-2 to pLTEX-5 are based on pLTEX-1 and only differ in their SP sequence (Manuscript I). They are used for recombinant protein secretion. The design of the SP has a great influence on the expression-secretion efficiency of the linked protein. Thus, minor amino-acid changes of the sequence can lead to a higher efficiency. The verified online tool SignalP (http://www.cbs.dtu.dk/services/SignalP-3.0/) can predict the presence and location of SPs and

their cleavage sites in polypeptides [146, 147] and was used here. Several different SPs were designed in-silico, the four most promising candidates were cloned and their functionality was verified by expression-secretion studies with above mentioned scFv's. In the end, the protein yield of scFv's could be increased close to an order of magnitude (Manuscript I), when cloned into pLTEX-5. These results pave the way for further expression-secretion studies in *L. tarentolae*, for example for the expression of full length immunoglobulins or potential highly glycosylated proteins. For this, sAPPalpha was cloned into pLTEX-5.

All *L. tarentolae* expression clones were analysed applying light-microscopy to detect possible phenotype changes caused by the expressed recombinant protein (data not shown). With the only exception of p53, no phenotypic changes in relation to the wild type could be observed. In addition, TEM images of p53 clones displayed autophagosome-like structures, strongly suggesting that p53 is also able to induce autophagy in a heterologous system like *L. tarentolae*. However, the effect of p53 on metabolic activities is known and the cells were still viable enabling subsequent purification of p53 (Manuscript IV).

Even after many weeks in culture, all recombinant proteins could be still purified. Compared to freshly transfected clones, no significant differences in the protein yields were observed, although not further determined. These results suggest that the selected and modified vectors are fully functional, and that genomic integration is performed at the right loci since no gene silencing events were observed. In the case of eGFP, small amounts were found as dimers and multimers, negatively influencing the purification efficacy, and it was concluded that possibly the overall high concentration of eGFP is responsible for this. In the case of secretory proteins (scFv's, sAPPalpha), the His_6-tag has turned out to be the right choice for purification, resulting in a very high purity (Manuscript I + II). In contrast, the same degree of purity was not achieved for intracellularly expressed proteins. However, isolation of intracellular proteins is in most cases more difficult due to a high endogenous protein background. FPLC-based purification strategies for eGFP, p53, FDFT1 and mTKIN (Manuscript IV) could still be improved to increase their purity. To further increase their purity or even their yield, different expression vector variants or affinity-tags can be utilized. Beside the His_6-tag (or the biotin-tag), further possible affinity-tags [152, 167] are for example the FLAG-tag (8 AA's) [168, 169], the c-myc-tag (10-11 AA's) [167] or the T7-tag (11-16 AA's) [170]. In general, peptide-tags can positively as well as negatively influence the yield of the recombinant protein [171, 172], its solubility [173] or its biological activity [174]. For this, the choice of the right affinity-tag is very important. Preparing 2-D gels of *L. tarentolae* lysates and subsequent incubation with tag-specific antibodies would give a first

impression of the background and the general usability of the tag. Also important is the right choice of the expression vector. In my work, all proteins were continuously expressed in *L. tarentolae* using the constitutive expression system of JBS [18]. This can become problematic if the recombinant protein disturbs cellular homeostasis. Except for p53 expression clones, no other recombinant protein had such an effect on *L. tarentolae*. In contrast, toxic proteins cannot be continuously expressed. Here, the inducible systems [15, 139] or rather the cell-free system [140] of *L. tarentolae* are much better suited. These expression vector variants could also result in a higher protein yield for my seven expressed proteins. To sum up, factors like the expression host, the expression vector, the affinity-tag and the recombinant protein strongly influence the folding status, PTMs, solubility, location and yield of the recombinant protein making an optimal choice nearly impossible.

Expression and analysis of the human secretory glycoprotein sAPPalpha

Many eukaryotic proteins are only functional when equipped with adequate PTMs. To date, more than 20 eukaryotic recombinant proteins could be functionally expressed in *L. tarentolae* [17, 18], indicating high homogeneity to their original counterpart. Especially the homogeneity of glycoproteins was often discussed, although rarely proven. Although the homogeneity is higher compared to many other eukaryotic expression systems, Leishmania species generally lack enzymes for the sialic acid biosynthesis pathway (Figure 4, subchapter 1.2.4) [126].

For example, in the publication of Dortay et al. [10], human c-reactive protein (CRP) is produced with *L. tarentolae*. CRP is a secretory glycoprotein naturally produced by the liver and widely accepted as cardiac marker. For functionality, the formation of a pentameric structure through disulphide bridges and a correct glycosylation pattern are necessary. In the study, this could be indirectly shown by a purification method where only intact pentameric molecules are to purify. However, neither disulphide bridges nor glycan structures were analysed. To give another example, the secretory glycoprotein proprotein convertase 4 (PC4; from rat) was enzymatically actively expressed [143]. PC4 plays a key role in mammalian fertilization, sperm maturation and sperm-egg fusion. Again, no glycan analysis was performed. The same applies for the large human secretory heterotrimeric glycoprotein laminin-332 (LM-332) [9] or for the human glycoprotein tissue plasminogen activator (t-PA) [175]. Unfortunately, not even a handful of studies are analysing glycans at all. All above mentioned studies refer to a single publication

from 2002, where mammalian-like glycosylated and biologically active EPO could be successfully expressed [18]. In detail, a mammalian-type biantennary fully galactosylated, core-α-1,6-fucosylated N-glycan was described. However, O-glycosylation of EPO could not be verified. Similar glycosylation profiles were also found in human interferon-γ (IFN-γ) and host major surface protein GP63, suggesting this is a common feature of all recombinant glycoproteins produced by *L. tarentolae* (JBS, LEXSY Essence). Summarized, these few examples are too little to clearly prove the existence of mammalian-like glycans, although above examples show a positive tendency as some of these recombinant proteins are only functional when properly glycosylated [9, 10].

Protein glycosylation is one of the most common PTM to occur in eukaryotic protein synthesis [8] and Leishmania are specifically rich in glycoproteins [121]. Glycosylation occurs within the secretory pathway, where the glycans are attached to proteins by a tremendously big set of transferases. The two most common types of glycosidic bonds are N- and O-linked glycosylations [176]. In the case of N-glycosylation, the oligosaccharyltransferase (OST) transfers the glycan to an asparagine (N) residue inside the known motif N-X-S/T/C. This process happens during protein translocation into the endoplasmic reticulum [177]. For O-glycosylation, the glycan is transferred to a serine (S) or threonine (T) residue inside a sequence area rich in S, T and proline (P). The glycoprotein sAPPalpha was selected due to the presence of nine potential N- and O-glycosylation sites [178-180]. It was analysed in detail to confirm high homogeneity to its human counterpart (Manuscript II). We could identify similar patterns of N-glycan site occupancy as described for sAPPalpha expressed in mammalian systems. However, in contrast to previous studies [18], we were unable to identify the presence of complex-type N-glycans since the only N-glycan structure to be present on *L. tarentolae* expressed sAPPalpha appeared to be the core N-glycan pentasaccharide. This difference could possibly be explained by previous studies reporting Leishmania and Trypanosoma N-glycosylation to be highly variable-dependent on the particular life cycle stage [177]. Nevertheless, the observed N-glycosylation pattern could also depend on protein or growth conditions. Importantly, our work describes the first report on the principle capacity of *L. tarentolae* to perform initial O-glycosylation steps (single HexNAc residues). The exact nature of the HexNAc residues requires further determination. Despite these findings, we were unable to identify any larger O-glycan structures commonly found in mammalian systems. Optimization of the expression conditions could result in higher O-glycosylation occupancy rates, and the single HexNAc residues could potentially be further modified in vitro using specific

glycosyltransferases, if particular structural features are required to be present on N- or O-glycans of the expressed proteins. Together, these results demonstrate that *L. tarentolae* is a well-suited system for the expression of mammalian glycoproteins, but additional studies are still necessary to fully understand the capacity of *L. tarentolae* to glycosylate heterologous proteins.

Could I achieve the aim of my thesis?

With the expression of scFv's, sAPPalpha and p53, I could underpin the ability of *L. tarentolae* to express biomedically relevant human recombinant proteins in a functional fashion. Thus, I could generally achieve my aim. Especially with sAPPalpha, I could demonstrate for the first time that *L. tarentolae* is also able to O-glycosylate a recombinant human glycoprotein and confirm high homogeneity to the appropriate counterpart. However, complex glycan structures were not found. Functionality of scFv's could be verified by ELISA (data not shown) experiments. Furthermore, overexpression of human p53 resulted in the formation of autophago(lyso)somes, which clearly indicates functionality of the protein. In future, further human recombinant proteins should be expressed in *L. tarentolae*, their functionality should be confirmed with appropriate functional assays and their PTMs should be analysed in detail to secure the capacity of *L. tarentolae* to express heterologous proteins functionally, possibly suitable for biomedical applications.

Moreover, in my thesis I could demonstrate that there was and still is room for optimization in vector design. Here I could show that optimization of the expression vector in terms of the secretory signal peptide sequence has a positive effect on the expression-secretion rate. In addition, I could demonstrate that for secretory proteins, the selected His_6-tag was well-suited and that very high purity of the recombinant protein can be obtained.

5 Current state of research and outlook

Like any other expression system, *L. tarentolae* possesses a number of advantages (subchapter 1.2 + 1.2.2) and disadvantages (subchapter 1.2.3). The parasite was originally isolated from the lizard *Tarentola annularis* in 1921 [14], and is the best studied saurian-pathogenic Leishmania species today (most publications in PubMed). Typically, the life cycle of Leishmania species alternates between promastigotes in the insect, and amastigotes in the vertebrate host (subchapter 1.1.2) [40]. Originally, different environmental conditions in the insect and in the vertebrate host gave rise to the development of these two main morphological forms. Promastigotes are perfectly adapted to a life as extracellular motile cells in the gut of insects, whereas amastigotes are perfectly adapted to survive the intracellular conditions of immune cells of their respective vertebrate hosts [181]. For instance, promastigote survival is depending on the adhesion of parasitic surface GPI-proteins with the intestinal wall of the insect vector. In contrast, strong reduction of surface GPI-proteins in amastigotes [75, 76] is an important survival benefit in immune cells resulting in decreased immune response [81]. Hence, without differentiation, amastigotes are not able to survive/reproduce inside immune cells, and promastigotes are not able to survive/reproduce in the gut of insect. The life cycle of *L. tarentolae* is slightly different. It alternates between promastigotes in the insect vector, and most likely amastigote-like forms [19-21] inside immune cells of the lizard as well as free promastigotes in the blood (subchapter 1.2.1 and 1.2.3). The existence of *L. tarentolae* amastigotes is still debated [101, 102]. In cell culture, Leishmania parasites normally occur as promastigotes. Mimicking phagolysosomal conditions [116, 117] can induce amastigote-like cells of *L. tarentolae*. However, induction of amastigote-like cells was only successful with *L. tarentolae* strain LEM-125, and not for the sequenced strain TarII (which was used here) [21]. The question is now: Would amastigotes or even amastigote-like cells – due to their evolutionary acquired adaptations to the vertebrate host – not be even more suitable for human recombinant protein expression? It is known, that during differentiation, the PTM of endogenous proteins changes [182]. This could also be the case for recombinant proteins, resulting in a higher homogeneity to the mammalian counterpart.

To date, recombinant protein expression was almost unexceptionally performed in promastigotes of *L. tarentolae.* Only GFP was episomally expressed in amastigote-like cells, but without recombinant protein analysis [100], showing that despite of parasitic differentiation, the gene was still translated leading to a functional product. Concluding, analysis and comparison of

recombinant proteins expressed in promastigotes and amastigote-like cells of *L. tarentolae* should be performed in the future.

Alternatively, different glycosyltransferases could be overexpressed to increase the homogeneity of recombinant glycoproteins (glycan pattern). Here, sialic acids are most important as these terminal sugar residues directly regulate the immune response and directly influence the serum half-life of glycoproteins e.g. immunoglobulins (Ig) [11]. Leishmania parasites lack enzymes for the sialic acid biosynthesis pathway [126], although terminal sialic acids (α2,3- and α2,6-linked) were found on endogenous membrane-glycoproteins of *L. major* and *L. donovani* [80, 81, 127]. To 'humanize' the glycosylation profile of *L. tarentolae*, the Trans-Sialidase (TS) can be transfected. The TS is an enzyme of Trypanosoma species that can transfer sialic acids from the host to endogenous glycoproteins [183, 184]. Successful expression of the TS and its functionality has already been demonstrated in insect cells [185] and yeasts [186]. Additionally, it was also successfully expressed in *L. major*. Here, the enzyme was used to investigate the influence on the virulence, and not for recombinant protein expression [127]. TS expression in *L. tarentolae* would enable to produce human-like, sialylated glycoproteins with a comparably low work load.

From above arguments, one might derive to the conclusion that *L. tarentolae* is not the best suited Leishmania species for recombinant protein expression due to the lack of amastigote stages and missing human pathogenicity factors (subchapter 1.2.1). In contrast, human pathogenic Leishmania species, like *L. major*, would be better-suited. However, cultivation of amastigotes in cell-free media is difficult and the high risk of infection (S2) makes cultivation very cost intensive. For these reasons it should be investigated in the future whether amastigote-like cells of *L. tarentolae* or amastigotes/amastigote-like cells of human-pathogenic species are better suited (higher homogeneity to the original counterpart) for human recombinant protein expression than promastigotes.

6 Summary

Leishmania tarentolae is a eukaryotic, protozoan parasite of the genus Leishmania. It is transmitted by sand flies and infects only reptiles. It was originally isolated in 1921 and is one of the best studied Leishmania species today. In addition to the research of human leishmaniasis, a strongly growing application area of this parasite in the last years is recombinant protein expression. The increasing use as an expression host is mainly based on its ability to synthetize human recombinant proteins not only in biologically active form, but also providing them with posttranslational modifications of high homogeneity. The quality of expressed proteins could be sufficient in principle to e.g. use them as therapeutic agents in biomedical research. However, this statement is based on little knowledge from very few analysed examples, questioning their possible biomedical application.

Hence, the main goal of my work was to evaluate the potential of the *L. tarentolae* expression system for biomedical research and the suitability of recombinant proteins as human therapeutic agents, respectively. Therefore, successfully established culture conditions as well as different expression vector variants will form a solid initial situation.

The most frequently approved recombinant therapeutic proteins are monoclonal antibodies, proteins with enzymatic activity or generally glycosylated proteins. Their successful expression in *L. tarentolae* would confirm the biomedical potential of the parasitic expression system and therefore the main goal of my work. Here, glycoproteins play a major role as more than 50% of all human proteins are glycosylated. During the secretory process, glycoproteins are glycosylated and are finally incorporated into the cell membrane or are secreted into the extracellular space/culture supernatant. For this, proteins need a secretory signal peptide, whose amino-acid sequence strongly influences expression-secretion efficiency of the protein.

To maximize the yield of secretory proteins, the secretory signal peptide of the acid phosphatase 1 (sAP1) of *L. mexicana* was modified in my work, and the expression-secretion efficiency was tested with recombinant human antibody fragments (scFv's). In the end, the protein yield could be increased close to an order of magnitude (Manuscript 1).

Based on these results, the human glycoprotein 'soluble amyloid precursor protein alpha' (sAPPalpha) was selected as a next example, and cloned together with the optimized signal peptide of manuscript I. Again, the secreted protein could be produced in a biologically active

form with sufficient quantities. Mass spectrometric analyses of sAPPalpha could show for the first time, that *L. tarentolae* is not only able to provide a human recombinant glycoprotein with N-glycosylations, but also with O-glycosylations (Manuscript 2). In addition, other studies could demonstrate, that the glycosylation profile of human proteins expressed in *L. tarentolae* shows, compared to other eukaryotic expression systems (yeast, insect cells, or plant cell culture), the highest homogeneity to the human original. These results support the application of human glycoproteins as therapeutic agents expressed with *L. tarentolae*.

Beside secretory proteins, intracellularly accumulating proteins were also expressed in my work. Due to the interaction with endogenous biomolecules, the choice of the right purification strategy is most important to get a product of high purity. For this, a new *L. tarentolae* strain was generated. By constitutive expression of the biotin ligase of *E. coli*, this strain is now able to biotinylate co-expressed proteins in a sequence-specific manner. Biotinylation is a relatively rare posttranslational protein modification, where the biotin ligase is only linking a biotin sequence-specific to proteins offering a biotin-acceptor peptide-sequence. Here, I could show minimal off-target effects of the enzyme. On the one hand, biotinylation is used to confirm successful protein purification. On the other hand, biotin itself can be used for purification. In my work, all intracellularly expressed proteins were purified via a poly-histidine-tag. Moreover, I could also show that intracellular accumulating human proteins can be functionally expressed in *L. tarentolae*. As example, human p53 could also trigger autophagy in *L. tarentolae* (Manuscript 4). For functionality, p53 needs to be properly phosphorylated, which was not shown here due to the lack of specific antibodies.

In conclusion, my results demonstrate that *L. tarentolae* is well-suited for the production of recombinant proteins of highest quality. In addition, mammalian-like glycosylation profiles of human recombinant proteins indicate that *L. tarentolae* can be utilized as an alternative expression system to cell culture systems like CHO-cells (Chinese Hamster Ovary cells). Concluding, *L. tarentolae* is in the future also eligible for the production of human therapeutically applicable proteins.

7 Zusammenfassung

Leishmania tarentolae ist ein einzelliger, eukaryotischer Parasit, welcher zur Gattung der Leishmanien gehört. Er wird von Sandmücken übertragen und infiziert – im Gegensatz zu seinen humanpathogenen Verwandten – keine Menschen, sondern ausschließlich Reptilien. Er wurde 1921 erstmals isoliert und gehört heute zu den am besten untersuchten Leishmania Arten. Neben der Erforschung der humanen Leishmaniose, wird er vor allem in den letzten Jahren verstärkt für die rekombinante Proteinexpression eingesetzt. Die zunehmende Verwendung als Expressionswirt beruht hauptsächlich auf seiner Fähigkeit, humane rekombinante Proteine nicht nur in biologisch aktiver Form zu synthetisieren, sondern auch mit posttranslationalen Modifikationen hoher Homogenität auszustatten. Die Qualität der hier entstehenden Proteine könnte prinzipiell ausreichen, um sie z.B. als Therapeutikum in der biomedizinischen Forschung einzusetzen. Jedoch fußt diese Aussage auf nur wenigen Erkenntnissen aus sehr wenigen analysierten Beispielen, sodass ihr möglicher biomedizinischer Einsatz nicht eindeutig gesichert ist.

Das Hauptziel meiner Arbeit bestand folglich darin, das Potential des Expressionssystems von *L. tarentolae* für die biomedizinische Forschung bzw. seine Eignung zur Herstellung rekombinanter Proteine als potentielle humane Therapeutika weiter zu evaluieren. Sowohl erfolgreich etablierte Kultivierungsbedingungen, als auch unterschiedliche Expressionsvektoren bilden hierfür eine solide Ausgangssituation.

Die am häufigsten zugelassenen rekombinanten Proteintherapeutika sind monoklonale Antikörper, Proteine mit enzymatischer Aktivität oder generell glykosylierte Proteine. Ihre erfolgreiche Expression in *L. tarentolae* wäre demnach eine Bestätigung des biomedizinischen Potentials des Expressionssystems und somit das Hauptziel meiner Arbeit. Glykoproteine spielen hier eine übergeordnete Rolle, da mehr als 50% aller menschlichen Proteine glykosyliert sind. Sie erhalten während des sekretorischen Prozesses ihre Glykosylierung und werden letztlich entweder in die Zellmembran inkorporiert, oder aber von der Zelle in den extrazellulären Raum/Kulturüberstand sezerniert. Hierfür benötigen die Proteine ein sekretorisches Signalpeptid, dessen Aminosäure-Abfolge in starker Weise die Expressions-Sekretions-Effizienz des Proteins beeinflusst.

Um folglich die Ausbeute sezernierter Proteine zu maximieren, wurde in dieser Arbeit das sekretorische Signalpeptid der sauren Phosphatase 1 (sAP1) aus *L. mexicana* modifiziert, und

die Expressions-Sekretions-Effizienz mit Antikörper-Fragmenten (scFv's) getestet. Hierbei konnte eine Erhöhung der Proteinausbeute um nahezu eine Potenz erzielt werden (Manuskript 1).

Basierend auf diesen Ergebnissen wurde das humane Glykoprotein ‚soluble amyloid precursor protein alpha' (sAPPalpha) als weiteres Beispiel ausgewählt, und mit dem optimierten Signalpeptid kloniert. Auch hier konnte das sezernierte Protein funktionell und in ausreichenden Mengen produziert werden. Anhand von sAPPalpha konnte zum ersten Mal mit diesem Teil der Arbeit gezeigt werden, das *L. tarentolae* nicht nur in der Lage ist ein humanes rekombinantes Glykoprotein mit N-Glykosylierungen, sondern ebenfalls mit O-Glykosylierungen auszustatten (Manuskript 2). Andere Studien konnten außerdem zeigen, dass das Glykosilierungsprofil von in *L. tarentolae* exprimierten humanen Proteinen im Vergleich zu anderen eukaryotischen Expressionssystemen (Hefen, Insektenzellen, oder pflanzlichen Zellkulturen) sogar die höchste Homogenität zum humanen Original aufwies. Diese Ergebnisse sprechen sehr stark dafür, *L. tarentolae* für die Expression humaner Glykoproteine als Therapeutika einzusetzen.

Neben sekretorischen Proteinen, habe ich in meiner Arbeit auch intrazellulär akkumulierende Proteine exprimiert. Aufgrund der Interaktion mit endogenen Biomolekülen, ist hier die Wahl der richtigen Aufreinigungsstrategie besonders wichtig, um am Ende ein Produkt von hoher Reinheit zu bekommen. Dazu wurde ein neuer *L. tarentolae* Stamm generiert, der durch die konstitutive Expression der Biotin-Ligase aus *E. coli* nun in der Lage ist, ko-exprimierte Proteine zu biotinylieren. Biotinylierung an sich ist eine seltene posttranslationale Proteinmodifizierung, bei der die Biotin-Ligase Biotin sequenzspezifisch an Proteine koppelt, die eine Biotin-Akzeptor Peptidsequenz besitzen. Hier konnte ich zeigen, dass es nur wenige Nebeneffekte des Enzyms gibt. Die Biotinylierung dient zum einen als Nachweis der erfolgreichen Proteinaufreinigung, zum anderen kann man über Biotin selbst aufreinigen. Hier wurden die Proteine jedoch über einen Polyhistidin-tag aufgereinigt. Weiterhin konnte ich zeigen, dass auch intrazellulär akkumulierende humane Proteine in Leishmania in funktioneller Form exprimiert werden können. So konnte humanes p53 auch in *L. tarentolae* Autophagie auslösen (Manuskript 4). Dies setzt u.a. die funktionelle posttranslationale Modifikation – hier die Phosphorylierung – des p53 voraus, welches jedoch hier Mangels spezifischer Antikörper nicht gezeigt werden konnte.

Zusammenfassend zeigen meine Ergebnisse, dass *L. tarentolae* für die Produktion von rekombinanten Proteinen höchster Qualität anwendbar ist. Durch das säugetierähnliche Glykosilierungsprofil von humanen rekombinanten Proteinen kann er zudem als alternatives

Expressionssystem zu Zellkultursystemen wie CHO-Zellen (Chinese Hamster Ovary) eingesetzt werden. Somit kann man schlussfolgern, dass *L. tarentolae* in Zukunft auch für die Herstellung für am Menschen therapeutisch anwendbarer Proteine durchaus in Frage kommen kann.

8 References

1. Hacker, D.L., M. De Jesus, and F.M. Wurm, *25 years of recombinant proteins from reactor-grown cells - where do we go from here?* Biotechnol Adv, 2009. 27(6): p. 1023-7.
2. Jayapal, K., K. Wlaschin, and W. Hu, *Recombinant Protein Therapeutics from CHO cells - 20 Years and Counting*. 2007, CHO Consortium, SBE Special Section.
3. Pertea, M. and S.L. Salzberg, *Between a chicken and a grape: estimating the number of human genes.* Genome Biol, 2010. 11(5): p. 206.
4. Pruitt, K.D., et al., *NCBI Reference Sequences: current status, policy and new initiatives.* Nucleic Acids Res, 2009. 37(Database issue): p. D32-6.
5. Boehringer-Ingelheim, *Unternehmensbericht*. 2006.
6. Roche-AG, *Biotechnologie - neue Wege in der Medizin*. 2006.
7. Walsh, G., *Biopharmaceutical benchmarks 2010.* Nat Biotechnol, 2010. 28(9): p. 917-24.
8. Wong, C.H., *Protein glycosylation: new challenges and opportunities.* J Org Chem, 2005. 70(11): p. 4219-25.
9. Phan, H.P., M. Sugino, and T. Niimi, *The production of recombinant human laminin-332 in a Leishmania tarentolae expression system.* Protein Expr Purif, 2009. 68(1): p. 79-84.
10. Dortay, H., et al., *Expression of human c-reactive protein in different systems and its purification from Leishmania tarentolae.* Protein Expr Purif, 2011. 78(1): p. 55-60.
11. Kolarich, D., B. Lepenies, and P.H. Seeberger, *Glycomics, glycoproteomics and the immune system.* Curr Opin Chem Biol, 2012. 16(1-2): p. 214-20.
12. Wurm, F.M., *Production of recombinant protein therapeutics in cultivated mammalian cells.* Nat Biotechnol, 2004. 22(11): p. 1393-8.
13. Sodoyer, R., *Expression systems for the production of recombinant pharmaceuticals.* BioDrugs, 2004. 18(1): p. 51-62.
14. Elwasila, M., *Leishmania tarentolae Wenyon, 1921 from the gecko Tarentola annularis in the Sudan.* Parasitol Res, 1988. 74(6): p. 591-2.
15. Kushnir, S., et al., *Development of an inducible protein expression system based on the protozoan host Leishmania tarentolae.* Protein Expr Purif, 2005. 42(1): p. 37-46.
16. Fritsche, C., et al., *Characterization of the growth behavior of Leishmania tarentolae: a new expression system for recombinant proteins.* J Basic Microbiol, 2007. 47(5): p. 384-93.
17. Basile, G. and M. Peticca, *Recombinant protein expression in Leishmania tarentolae.* Mol Biotechnol, 2009. 43(3): p. 273-8.
18. Reinhard Breitling, S.K., Nico Callewaert, Regina Pietrucha,, G.E. Anett Geyer, Regina Hartung, Angelika Müller, Roland Contreras, and a.K.A. Stephen M. Beverley, *Non-pathogenic trypanosomatid protozoa as a platform for protein research and production.* Protein Expression and Purification, 2002. 25 (2002) 209–218.
19. Breton, M., et al., *Live nonpathogenic parasitic vector as a candidate vaccine against visceral leishmaniasis.* Infect Immun, 2005. 73(10): p. 6372-82.
20. Breton, M., et al., *A recombinant non-pathogenic Leishmania vaccine expressing human immunodeficiency virus 1 (HIV-1) Gag elicits cell-mediated immunity in mice and decreases HIV-1 replication in human tonsillar tissue following exposure to HIV-1 infection.* J Gen Virol, 2007. 88(Pt 1): p. 217-25.
21. Taylor, V.M., et al., *Leishmania tarentolae: utility as an in vitro model for screening of antileishmanial agents.* Exp Parasitol, 2010. 126(4): p. 471-5.

22. Simpson, L., et al., *Mitochondrial proteins and complexes in Leishmania and Trypanosoma involved in U-insertion/deletion RNA editing.* RNA, 2004. 10(2): p. 159-70.
23. White, T.C., et al., *The H circles of Leishmania tarentolae are a unique amplifiable system of oligomeric DNAs associated with drug resistance.* J Biol Chem, 1988. 263(32): p. 16977-83.
24. Ouellette, M., et al., *Direct and inverted DNA repeats associated with P-glycoprotein gene amplification in drug resistant Leishmania.* EMBO J, 1991. 10(4): p. 1009-16.
25. Niimi, T., *Recombinant protein production in the eukaryotic protozoan parasite Leishmania tarentolae: a review.* Methods Mol Biol, 2012. 824: p. 307-15.
26. Raymond, F., et al., *Genome sequencing of the lizard parasite Leishmania tarentolae reveals loss of genes associated to the intracellular stage of human pathogenic species.* Nucleic Acids Res, 2012. 40(3): p. 1131-47.
27. Stuart, K., et al., *Kinetoplastids: related protozoan pathogens, different diseases.* J Clin Invest, 2008. 118(4): p. 1301-10.
28. Teixeira, S.M., *Control of gene expression in Trypanosomatidae.* Braz J Med Biol Res, 1998. 31(12): p. 1503-16.
29. Leishman, W.B., *Note on the Nature of the Parasitic Bodies found in Tropical Splenomegaly.* Br Med J, 1904. 1(2249): p. 303.
30. Jadin, J.M. and J. Creemers, *[The ultrastructure of rosette forms of Leishmania tropica, Wright, 1903].* Ann Soc Belges Med Trop Parasitol Mycol, 1966. 46(3): p. 349-54.
31. (WHO), W.H.O., *Control of the leishmaniases.* Technical Report Series 793; Geneva: WHO, 21-25., 1990.
32. Marco JD, B.P., Calvopiña M, Kumazawa H, Furuya M, Korenaga M, Cajal SP, Mora MC, Rea MM, Borda CE, Basombrío MA, Taranto NJ, Hashiguchi Y., *Species Assignation Of Leishmania From Human And Canine American Tegumentary Leishmaniasis Cases By Multilocus Enzyme Electrophoresis In North Argentina.* Am J Trop Med Hyg. 72(5):606-11, 2005.
33. Gabriele Schönian, I.M.a.E.C., *Is it time to revise the nomenclature of Leishmania?* Trends in Parasitology 2010. Vol.26 No.10.
34. Lainson R, R.L., Shaw JJ. , *Evolution, classification and geographical distribution.* . The Leishmaniases in Biology and Medicine 1987. Vol. 1. Biology and Epidemiology. Academic Press, London. Pp I-120
35. Croan, D. and J. Ellis, *Monophyletic origin of the genus Sauroleishmania.* Archiv für Protistenkunde, 1997. 148(3): p. 269-275.
36. Asato, Y., et al., *Phylogenic analysis of the genus Leishmania by cytochrome b gene sequencing.* Exp Parasitol, 2009. 121(4): p. 352-61.
37. Fraga, J., et al., *Phylogeny of Leishmania species based on the heat-shock protein 70 gene.* Infect Genet Evol, 2010. 10(2): p. 238-45.
38. Requena, J.M., et al., *Sequence analysis of the 3'-untranslated region of HSP70 (type I) genes in the genus Leishmania: its usefulness as a molecular marker for species identification.* Parasit Vectors, 2012. 5: p. 87.
39. Gossage, S.M., M.E. Rogers, and P.A. Bates, *Two separate growth phases during the development of Leishmania in sand flies: implications for understanding the life cycle.* International Journal for Parasitology, 2003. 33(10): p. 1027-1034.
40. Wheeler, R.J., E. Gluenz, and K. Gull, *The cell cycle of Leishmania: morphogenetic events and their implications for parasite biology.* Mol Microbiol, 2011. 79(3): p. 647-62.
41. Killick-Kendrick, R., *Phlebotomine vectors of the leishmaniases: a review.* Med Vet Entomol, 1990. 4(1): p. 1-24.
42. Gluenz, E., M.L. Ginger, and P.G. McKean, *Flagellum assembly and function during the Leishmania life cycle.* Curr Opin Microbiol, 2010. 13(4): p. 473-9.

43. Maroli, M., et al., *Phlebotomine sandflies and the spreading of leishmaniases and other diseases of public health concern.* Med Vet Entomol, 2012.
44. Rose, K., et al., *Cutaneous leishmaniasis in red kangaroos: isolation and characterisation of the causative organisms.* Int J Parasitol, 2004. 34(6): p. 655-64.
45. Dougall, A.M., et al., *Evidence incriminating midges (Diptera: Ceratopogonidae) as potential vectors of Leishmania in Australia.* Int J Parasitol, 2011. 41(5): p. 571-9.
46. Rougeron, V., et al., *"Everything you always wanted to know about sex (but were afraid to ask)" in Leishmania after two decades of laboratory and field analyses.* PLoS Pathog, 2010. 6(8): p. e1001004.
47. Tibayrenc, M. and F.J. Ayala, *The clonal theory of parasitic protozoa: 12 years on.* Trends Parasitol, 2002. 18(9): p. 405-10.
48. Victoir, K. and J.C. Dujardin, *How to succeed in parasitic life without sex? Asking Leishmania.* Trends Parasitol, 2002. 18(2): p. 81-5.
49. Zeledon, R.A., *Hemoflagellates*, in *Medical Microbiology*, S. Baron, Editor. 1996: Galveston (TX).
50. Dujardin, J.C., et al., *Putative Leishmania hybrids in the Eastern Andean valley of Huanuco, Peru.* Acta Trop, 1995. 59(4): p. 293-307.
51. Panton, L.J., et al., *A test for genetic exchange in mixed infections of Leishmania major in the sand fly Phlebotomus papatasi.* J Protozool, 1991. 38(3): p. 224-8.
52. Rougeron, V., et al., *Extreme inbreeding in Leishmania braziliensis.* Proc Natl Acad Sci U S A, 2009. 106(25): p. 10224-9.
53. Akopyants, N.S., et al., *Demonstration of genetic exchange during cyclical development of Leishmania in the sand fly vector.* Science, 2009. 324(5924): p. 265-8.
54. Gibson, W., et al., *The use of yellow fluorescent hybrids to indicate mating in Trypanosoma brucei.* Parasit Vectors, 2008. 1(1): p. 4.
55. Peacock, L., et al., *Intraclonal mating occurs during tsetse transmission of Trypanosoma brucei.* Parasit Vectors, 2009. 2(1): p. 43.
56. Murray, H.W., et al., *Advances in leishmaniasis.* The Lancet, 2005. 366(9496): p. 1561-1577.
57. Reithinger, R., et al., *Cutaneous leishmaniasis.* The Lancet Infectious Diseases, 2007. 7(9): p. 581-596.
58. Bailey, M.S. and D.N. Lockwood, *Cutaneous leishmaniasis.* Clin Dermatol, 2007. 25(2): p. 203-11.
59. Goto H, L.J., *Current diagnosis and treatment of cutaneous and mucocutaneous leishmaniasis.* Expert Rev Anti Infect Ther. 2010 Apr;8(4):419-33., 2010.
60. Aebischer, P.M.K.a.T., *Visceral leishmaniasis: immunology and prospects for a vaccine.* Clin Microbiol Infect., 2011. 2011 Oct;17(10):1462-70.
61. Srividya, G., et al., *Diagnosis of visceral leishmaniasis: developments over the last decade.* Parasitol Res, 2012. 110(3): p. 1065-78.
62. R. Reithinger, M.M., M. Wahid, M. Bismullah, R. J. Quinnell, C. R. Davies, J. Kolaczinski, and J. R. David, *Efficacy of Thermotherapy to Treat Cutaneous Leishmaniasis Caused by Leishmania tropica in Kabul, Afghanistan: A Randomized, Controlled Trial.* Clinical Infectious Diseases, 2005. 40:1148-55.
63. Zhang, W.W., et al., *Comparison of the A2 gene locus in Leishmania donovani and Leishmania major and its control over cutaneous infection.* J Biol Chem, 2003. 278(37): p. 35508-15.
64. kleintier.konkret, E.V. *Die canine Leishmaniose – Erkrankung, Therapie und Prophylaxe durch eine Impfung*. newsletter|forschung.industrie, 2012. 2012; 2: 3 – 6.
65. Agency, E.M., *CaniLeish- Adjuvantierter Impfstoff für Hunde gegen Leishmania infantum*, Z.d.E.f.d. Öffentlichkeit, Editor. 2011.
66. McConville, M.J., et al., *Living in a phagolysosome; metabolism of Leishmania amastigotes.* Trends Parasitol, 2007. 23(8): p. 368-75.

67. Naderer, T. and M.J. McConville, *The Leishmania-macrophage interaction: a metabolic perspective.* Cell Microbiol, 2008. 10(2): p. 301-8.
68. Hsiao, C.H., et al., *The effects of macrophage source on the mechanism of phagocytosis and intracellular survival of Leishmania.* Microbes Infect, 2011. 13(12-13): p. 1033-44.
69. Naderer, T. and M.J. McConville, *Intracellular growth and pathogenesis of Leishmania parasites.* Essays Biochem, 2011. 51: p. 81-95.
70. Ueno, N. and M.E. Wilson, *Receptor-mediated phagocytosis of Leishmania: implications for intracellular survival.* Trends Parasitol, 2012. 28(8): p. 335-44.
71. Holm, Å., et al., *Role of protein kinase C α for uptake of unopsonized prey and phagosomal maturation in macrophages.* Biochemical and Biophysical Research Communications, 2003. 302(4): p. 653-658.
72. Vinet, A.F., et al., *Exclusion of synaptotagmin V at the phagocytic cup by Leishmania donovani lipophosphoglycan results in decreased promastigote internalization.* Microbiology, 2011. 157(Pt 9): p. 2619-28.
73. JEAN-CLAUDE ANTOINE, E.P., COLETTE JOUANNE, AND PIERRE BONGRAND, *Parasitophorous Vacuoles of Leishmania amazonensis-Infected Macrophages Maintain an Acidic pH.* Infection and Immunity, 1990. Mar. 1990, p. 779-787.
74. Gergely L. Lukacs, O.D.R.a.S.G., *Determinants of the Phagosornal pH in Macrophages. In situ assessment of vacuolar H(+)-ATPase activity, counterion conductance, and H+ "leak".* JOURNALOF BIOLOGICAL CHEMISTRY, 1991. Vol. 266, No. 36(Issue of December 25): p. pp. 24540-24548.
75. Blackwell, M.J.M.a.J.M., *Developmental changes in the glycosylated phosphatidylinositols of Leishmania donovani.* THE JOURNAL OF BIOLOGICAL CHEMISTRY, 1991. Vol. 266, No. 23(Issue of August 15): p. PY. 15170-15179.
76. McConville, M.J., et al., *Secretory pathway of trypanosomatid parasites.* Microbiol Mol Biol Rev, 2002. 66(1): p. 122-54; table of contents.
77. Zilberstein, D., *The role of pH and temperature in the development of Leishmania parasites.* Annu. Rev. Microbiol. , 1994. 1994. 48:449-70.
78. Opperdoes, F.R. and G.H. Coombs, *Metabolism of Leishmania: proven and predicted.* Trends Parasitol, 2007. 23(4): p. 149-58.
79. Rosenzweig, D., et al., *Retooling Leishmania metabolism: from sand fly gut to human macrophage.* FASEB J, 2008. 22(2): p. 590-602.
80. Chatterjee M, C.A., Kohla G, Pal S, Merling A, Hinderlich S, Unger U, Strasser P, Gerwig GJ, Kamerling JP, Vlasak R, Crocker PR, Schauer R, Schwartz-Albiez R, Mandal C. , *Identification and characterization of adsorbed serum sialoglycans on Leishmania donovani promastigotes.* Glycobiology, 2003. Vol. 13, No. 5: p. pp 351-361.
81. Chava AK, C.M., Gerwig GJ, Kamerling JP, Mandal C., *Identification of sialic acids on Leishmania donovani amastigotes.* Biol Chem., 2004. 2004 Jan; 385 (1): p. pp 59-66.
82. Li, Y. and X. Chen, *Sialic acid metabolism and sialyltransferases: natural functions and applications.* Appl Microbiol Biotechnol, 2012. 94(4): p. 887-905.
83. Lewis, A.L. and W.G. Lewis, *Host sialoglycans and bacterial sialidases: a mucosal perspective.* Cell Microbiol, 2012. 14(8): p. 1174-82.
84. Singh, N., M. Kumar, and R.K. Singh, *Leishmaniasis: Current status of available drugs and new potential drug targets.* Asian Pacific Journal of Tropical Medicine, 2012. 5(6): p. 485-497.
85. Croft, S.L., S. Sundar, and A.H. Fairlamb, *Drug resistance in leishmaniasis.* Clin Microbiol Rev, 2006. 19(1): p. 111-26.

86. Hart;, D.T. and F.R. Opperdoes, *The occurence of glycosomes (microbodies) in the promastigote stage of four major Leishmania species.* Molecular and Biochemical Parasitology, 1984. 13 (1984) 159-172.
87. Opperdoes, F.R., *Compartmentation of Carbohydrate Metabolism in Trypanosomes.* Ann. Rev. Microbiol. , 1987. 1987. 41 :127-51.
88. Docampo, R. and S. Moreno, *Acidocalcisome: A novel Ca2+ storage compartment in trypanosomatids and apicomplexan parasites.* Parasitology Today, 1999. vol. 15, no. 11, 1999.
89. Docampo, R. and S.N. Moreno, *Acidocalcisomes.* Cell Calcium, 2011. 50(2): p. 113-9.
90. Peter Overath, Y.-D.S.a.M.W., *Endocytosis and secretion in trypanosomatid parasites- tumultous traffic in a pocket.* Trends in CELL BIOLOGY, 1997. Vol. 7, January.
91. Scott M. Landfear, M.I., *The flagellum and flagellar pocket of trypanosomatids.* Molecular & Biochemical Parasitology, 2001. 115 (2001) 1–17.
92. Liu, B., et al., *Fellowship of the rings: the replication of kinetoplast DNA.* Trends Parasitol, 2005. 21(8): p. 363-9.
93. Motta, M.C., *Kinetoplast as a potential chemotherapeutic target of trypanosomatids.* Curr Pharm Des, 2008. 14(9): p. 847-54.
94. Ferguson, M., *The structure, biosynthesis and functions of glycosylphosphatidylinositol anchors, and the contributions of trypanosome research.* Journal of Cell Science, 1999. 112, 2799-2809 (1999).
95. Previato, J.O., et al., *Leishmania adleri, a lizard parasite, expresses structurally similar glycoinositolphospholipids to mammalian Leishmania.* Glycobiology, 1997. 7(5): p. 687-95.
96. Fu, G. and A.A. Kolesnikov, *Leishmania gymnodactyli and Leishmania infantum minicircles contain the same guide RNA genes as do minicircles of Leishmania tarentolae.* Mol Biochem Parasitol, 1994. 67(1): p. 171-4.
97. Ivens, A.C., et al., *The genome of the kinetoplastid parasite, Leishmania major.* Science, 2005. 309(5733): p. 436-42.
98. Peacock, C.S., et al., *Comparative genomic analysis of three Leishmania species that cause diverse human disease.* Nat Genet, 2007. 39(7): p. 839-47.
99. Azizi, H., et al., *Searching for virulence factors in the non-pathogenic parasite to humans Leishmania tarentolae.* Parasitology, 2009. 136(7): p. 723-35.
100. Mizbani, A., et al., *Effect of A2 gene on infectivity of the nonpathogenic parasite Leishmania tarentolae.* Parasitol Res, 2011. 109(3): p. 793-9.
101. Simpson, L. and G. Holz, Jr., *The status of Leishmania tarentolae/Trypanosoma platydactyli.* Parasitol Today, 1988. 4(4): p. 115-8.
102. LUMSDEN, W.H.R. and D.A. EVANS, *Lizard Leishmania*, in *Biology of the Kinetoplastida*, W.H.R. LUMSDEN and D.A. EVANS, Editors. 1979. p. p. 241-268.
103. Mousavi, S.A., et al., *Clathrin-dependent endocytosis.* Biochem J, 2004. 377(Pt 1): p. 1-16.
104. Williams, D.B., *Beyond lectins: the calnexin/calreticulin chaperone system of the endoplasmic reticulum.* J Cell Sci, 2006. 119(Pt 4): p. 615-23.
105. Ilgoutz, S.C. and M.J. McConville, *Function and assembly of the Leishmania surface coat.* Int J Parasitol, 2001. 31(9): p. 899-908.
106. Yao, C., J.E. Donelson, and M.E. Wilson, *The major surface protease (MSP or GP63) of Leishmania sp. Biosynthesis, regulation of expression, and function.* Mol Biochem Parasitol, 2003. 132(1): p. 1-16.
107. Neidhardt, F., *Escherichia coli and Salmonella: Cellular and Molecular Biology*, F. Neidhardt, Editor. 1996, American Society for Microbiology. p. 1-5.
108. Baneyx, F., *Recombinant protein expression in Escherichia coli.* Curr Opin Biotechnol, 1999. 10(5): p. 411-21.

109. Sorensen, H.P. and K.K. Mortensen, *Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli.* Microb Cell Fact, 2005. 4(1): p. 1.
110. Villaverde, A. and M.M. Carrio, *Protein aggregation in recombinant bacteria: biological role of inclusion bodies.* Biotechnol Lett, 2003. 25(17): p. 1385-95.
111. Konthur, Z., M. Hust, and S. Dubel, *Perspectives for systematic in vitro antibody generation.* Gene, 2005. 364: p. 19-29.
112. Jenkins, N., et al., *Strategies for analysing and improving the expression and quality of recombinant proteins made in mammalian cells.* Biotechnol Appl Biochem, 2009. 53(Pt 2): p. 73-83.
113. Mureev, S., et al., *Construction and analysis of Leishmania tarentolae transgenic strains free of selection markers.* Mol Biochem Parasitol, 2007. 155(2): p. 71-83.
114. Fritsche, C., et al., *Development of a defined medium for heterologous expression in Leishmania tarentolae.* J Basic Microbiol, 2008. 48(6): p. 488-95.
115. Carretero, M.A., *Preferred temperatures of Tarentola mauritanica in spring.* Acta Herpetologica, 2008. 3(1): p. 57-64.
116. Saar, Y., et al., *Characterization of developmentally-regulated activities in axenic amastigotes of Leishmania donovani.* Mol Biochem Parasitol, 1998. 95(1): p. 9-20.
117. Debrabant, A., et al., *Generation of Leishmania donovani axenic amastigotes: their growth and biological characteristics.* Int J Parasitol, 2004. 34(2): p. 205-17.
118. Somanna, A., V. Mundodi, and L. Gedamu, *In vitro cultivation and characterization of Leishmania chagasi amastigote-like forms.* Acta Trop, 2002. 83(1): p. 37-42.
119. Alcolea, P.J., et al., *Temperature increase prevails over acidification in gene expression modulation of amastigote differentiation in Leishmania infantum.* BMC Genomics, 2010. 11: p. 31.
120. Rapoport, T.A., *Transport of proteins across the endoplasmic reticulum membrane.* Science, 1992. 258(5084): p. 931-6.
121. Ferguson, M.A., *The surface glycoconjugates of trypanosomatid parasites.* Philos Trans R Soc Lond B Biol Sci, 1997. 352(1359): p. 1295-302.
122. Roth, J., et al., *Protein N-glycosylation, protein folding, and protein quality control.* Mol Cells, 2010. 30(6): p. 497-506.
123. Sola, R.J. and K. Griebenow, *Effects of glycosylation on the stability of protein pharmaceuticals.* J Pharm Sci, 2009. 98(4): p. 1223-45.
124. Pedemonte, C.H., G. Sachs, and J.H. Kaplan, *An intrinsic membrane glycoprotein with cytosolically oriented n-linked sugars.* Proc Natl Acad Sci U S A, 1990. 87(24): p. 9789-93.
125. Reeves, R., D. Chang, and S.C. Chung, *Carbohydrate modifications of the high mobility group proteins.* Proc Natl Acad Sci U S A, 1981. 78(11): p. 6704-8.
126. Ghoshal, A. and C. Mandal, *A perspective on the emergence of sialic acids as potent determinants affecting leishmania biology.* Mol Biol Int, 2011. 2011: p. 532106.
127. Belen Carrillo, M., et al., *Heterologous expression of Trypanosoma cruzi trans-sialidase in Leishmania major enhances virulence.* Infect Immun, 2000. 68(5): p. 2728-34.
128. Sefton, B.M. and S. Shenolikar, *Overview of protein phosphorylation.* Curr Protoc Protein Sci, 2001. Chapter 13: p. Unit13 1.
129. Johnson, L.N., *The regulation of protein phosphorylation.* Biochem Soc Trans, 2009. 37(Pt 4): p. 627-41.
130. Sevier, C.S. and C.A. Kaiser, *Conservation and diversity of the cellular disulfide bond formation pathways.* Antioxid Redox Signal, 2006. 8(5-6): p. 797-811.
131. Kouzarides, T., *Acetylation: a regulatory modification to rival phosphorylation?* EMBO J, 2000. 19(6): p. 1176-9.

132. Cronan, J.E., Jr., *Biotination of proteins in vivo. A post-translational modification to label, purify, and study proteins.* J Biol Chem, 1990. 265(18): p. 10327-33.
133. Sugino, M. and T. Niimi, *Expression of multisubunit proteins in Leishmania tarentolae.* Methods Mol Biol, 2012. 824: p. 317-25.
134. Wen Wei Zhang, H.C.a.G.M., *The expression of biologically active human p53 in Leishmania cells: a novel eukaryotic system to produce recombinant proteins.* Nucleic Acids Research, 1995. Vol. 23, No. 20 4073-4080.
135. *Jena Bioscience GmbH, WO 01/32896 A1. Protein Expression System for non-pathogentic Kinetoplastidae.* 1999.
136. Misslitz, A., et al., *Targeted integration into a rRNA locus results in uniform and high level expression of transgenes in Leishmania amastigotes.* Mol Biochem Parasitol, 2000. 107(2): p. 251-61.
137. Clayton, C.E., *Life without transcriptional control? From fly to man and back again.* EMBO J, 2002. 21(8): p. 1881-8.
138. Donelson, J.E. and W. Zeng, *A comparison of trans-RNA splicing in trypanosomes and nematodes.* Parasitol Today, 1990. 6(10): p. 327-34.
139. Kushnir, S., et al., *Artificial linear episome-based protein expression system for protozoon Leishmania tarentolae.* Mol Biochem Parasitol, 2011. 176(2): p. 69-79.
140. Mureev, S., et al., *Species-independent translational leaders facilitate cell-free expression.* Nat Biotechnol, 2009. 27(8): p. 747-52.
141. Kovtun, O., et al., *Towards the construction of expressed proteomes using a Leishmania tarentolae based cell-free expression system.* PLoS One, 2010. 5(12): p. e14388.
142. Kovtun, O., et al., *Leishmania cell-free protein expression system.* Methods, 2011. 55(1): p. 58-64.
143. Basak, A., et al., *Recombinant proprotein convertase 4 (PC4) from Leishmania tarentolae expression system: purification, biochemical study and inhibitor design.* Protein Expr Purif, 2008. 60(2): p. 117-26.
144. Tereza Cristina Orlando, M.G.M., David A Campbell, Nancy R Sturm, Lucile Maria Floeter-Winter, *RNA polymerase I promoter and splice acceptor site recognition affect gene expression in non-pathogenic Leishmania species.* Mem Inst Oswaldo Cruz 2007. Vol. 102(7): 891-894.
145. Gazdag, E.M., et al., *Purification and crystallization of human Cu/Zn superoxide dismutase recombinantly produced in the protozoan Leishmania tarentolae.* Acta Crystallogr Sect F Struct Biol Cryst Commun, 2010. 66(Pt 8): p. 871-7.
146. Nielsen, H., et al., *Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites.* Protein Eng, 1997. 10(1): p. 1-6.
147. Bendtsen, J.D., et al., *Improved prediction of signal peptides: SignalP 3.0.* J Mol Biol, 2004. 340(4): p. 783-95.
148. Mechold, U., C. Gilbert, and V. Ogryzko, *Codon optimization of the BirA enzyme gene leads to higher expression and an improved efficiency of biotinylation of target proteins in mammalian cells.* J Biotechnol, 2005. 116(3): p. 245-9.
149. Schatz, P.J., *Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in Escherichia coli.* Biotechnology (N Y), 1993. 11(10): p. 1138-43.
150. Structural Genomics, C., et al., *Protein production and purification.* Nat Methods, 2008. 5(2): p. 135-46.
151. Koehn, J. and I. Hunt, *High-Throughput Protein Production (HTPP): a review of enabling technologies to expedite protein production.* Methods Mol Biol, 2009. 498: p. 1-18.

152. Arnau, J., et al., *Reprint of: Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins.* Protein Expr Purif, 2011.
153. Jefferis, R., *Recombinant antibody therapeutics: the impact of glycosylation on mechanisms of action.* Trends Pharmacol Sci, 2009. 30(7): p. 356-62.
154. Arnold, J.N., et al., *The impact of glycosylation on the biological function and structure of human immunoglobulins.* Annu Rev Immunol, 2007. 25: p. 21-50.
155. Zhao, C., et al., *Hexahistidine-tag-specific optical probes for analyses of proteins and their interactions.* Anal Biochem, 2010. 399(2): p. 237-45.
156. Crowe, J., B.S. Masone, and J. Ribbe, *One-step purification of recombinant proteins with the 6xHis tag and Ni-NTA resin.* Mol Biotechnol, 1995. 4(3): p. 247-58.
157. Schmitt, J., H. Hess, and H.G. Stunnenberg, *Affinity purification of histidine-tagged proteins.* Mol Biol Rep, 1993. 18(3): p. 223-30.
158. Perneczky, R., et al., *CSF soluble amyloid precursor proteins in the diagnosis of incipient Alzheimer disease.* Neurology, 2011. 77(1): p. 35-8.
159. Ray, B., et al., *Increased secreted amyloid precursor protein-alpha (sAPPalpha) in severe autism: proposal of a specific, anabolic pathway and putative biomarker.* PLoS One, 2011. 6(6): p. e20405.
160. Lamb, P. and L. Crawford, *Characterization of the human p53 gene.* Mol Cell Biol, 1986. 6(5): p. 1379-85.
161. Muller, P.A. and K.H. Vousden, *p53 mutations in cancer.* Nat Cell Biol, 2013. 15(1): p. 2-8.
162. Maiuri, M.C., et al., *Autophagy regulation by p53.* Curr Opin Cell Biol, 2010. 22(2): p. 181-5.
163. Do, R., et al., *Squalene synthase: a critical enzyme in the cholesterol biosynthesis pathway.* Clin Genet, 2009. 75(1): p. 19-29.
164. Chapman-Smith, A. and J.E. Cronan, Jr., *Molecular biology of biotin attachment to proteins.* J Nutr, 1999. 129(2S Suppl): p. 477S-484S.
165. Sakowicz, R., S. Farlow, and L.S. Goldstein, *Cloning and expression of kinesins from the thermophilic fungus Thermomyces lanuginosus.* Protein Sci, 1999. 8(12): p. 2705-10.
166. Rivera, S.B., et al., *Temperature dependent properties of a kinesin-3 motor protein from Thermomyces lanuginosus.* Fungal Genet Biol, 2007. 44(11): p. 1170-9.
167. Terpe, K., *Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems.* Appl Microbiol Biotechnol, 2003. 60(5): p. 523-33.
168. Knappik, A. and A. Pluckthun, *An improved affinity tag based on the FLAG peptide for the detection and purification of recombinant antibody fragments.* Biotechniques, 1994. 17(4): p. 754-61.
169. Einhauer, A. and A. Jungbauer, *The FLAG peptide, a versatile fusion tag for the purification of recombinant proteins.* J Biochem Biophys Methods, 2001. 49(1-3): p. 455-65.
170. Chatterjee, D.K. and D. Esposito, *Enhanced soluble protein expression using two new fusion tags.* Protein Expr Purif, 2006. 46(1): p. 122-9.
171. Sun, Q.M., et al., *An improved strategy for high-level production of human vasostatin120-180.* Biotechnol Prog, 2005. 21(4): p. 1048-52.
172. Goel, A., et al., *Relative position of the hexahistidine tag effects binding properties of a tumor-associated single-chain Fv construct.* Biochim Biophys Acta, 2000. 1523(1): p. 13-20.
173. Dyson, M.R., et al., *Production of soluble mammalian proteins in Escherichia coli: identification of protein features that correlate with successful expression.* BMC Biotechnol, 2004. 4: p. 32.
174. Fonda, I., et al., *Attachment of histidine tags to recombinant tumor necrosis factor-alpha drastically changes its properties.* ScientificWorldJournal, 2002. 2: p. 1312-25.
175. Hemayatkar, M., et al., *Increased expression of recombinant human tissue plasminogen activator in Leishmania tarentolae.* Biotechnol J, 2010. 5(11): p. 1198-206.

176. Hagglund, P., et al., *A new strategy for identification of N-glycosylated proteins and unambiguous assignment of their glycosylation sites using HILIC enrichment and partial deglycosylation.* J Proteome Res, 2004. 3(3): p. 556-66.
177. Parodi, A.J., *Biosynthesis of protein-linked oligosaccharides in trypanosomatid flagellates.* Parasitol Today, 1993. 9(10): p. 373-7.
178. Perdivara, I., et al., *Elucidation of O-glycosylation structures of the beta-amyloid precursor protein by liquid chromatography-mass spectrometry using electron transfer dissociation and collision induced dissociation.* J Proteome Res, 2009. 8(2): p. 631-42.
179. Halim, A., et al., *Site-specific characterization of threonine, serine, and tyrosine glycosylations of amyloid precursor protein/amyloid beta-peptides in human cerebrospinal fluid.* Proc Natl Acad Sci U S A, 2011. 108(29): p. 11848-53.
180. Liu, T., et al., *Human plasma N-glycoproteome analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry.* J Proteome Res, 2005. 4(6): p. 2070-80.
181. Besteiro, S., et al., *Protein turnover and differentiation in Leishmania.* Int J Parasitol, 2007. 37(10): p. 1063-75.
182. Rosenzweig, D., et al., *Post-translational modification of cellular proteins during Leishmania donovani differentiation.* Proteomics, 2008. 8(9): p. 1843-50.
183. Pereira, M.E., *A developmentally regulated neuraminidase activity in Trypanosoma cruzi.* Science, 1983. 219(4591): p. 1444-6.
184. Schenkman, S., et al., *Structural and functional properties of Trypanosoma trans-sialidase.* Annu Rev Microbiol, 1994. 48: p. 499-523.
185. Marchal, I., et al., *Expression of a membrane-bound form of Trypanosoma cruzi trans-sialidase in baculovirus-infected insect cells: a potential tool for sialylation of glycoproteins produced in the baculovirus-insect cells system.* Glycobiology, 2001. 11(7): p. 593-603.
186. Ryckaert, S., et al., *Development of a S. cerevisiae whole cell biocatalyst for in vitro sialylation of oligosaccharides.* J Biotechnol, 2005. 119(4): p. 379-88.