**Structural and Functional Analyses of Secretory and Excretory Proteins from *Onchocerca volvulus* as Basis for Rational Drug Design**

**DISSERTATION**

zur Erlangung des akademischen Grades eines Doktors der

Naturwissenschaften (Dr. rer. nat.)

**Fachbereich Chemie**

der Universität Hamburg

vorgelegt von

**Amr Moustafa**

aus Sharkia, Ägypten

Hamburg, April 2016

Die vorliegende Arbeit wurde im Zeitraum von April 2012 bis Januar 2016 in der Arbeitsgruppe von Prof. Ch. Betzel im Laboratorium für Strukturbiologie von Infektion und Entzündung am Institut für Biochemie und Molekularbiologie des Fachbereichs Chemie der Universität Hamburg durchgeführt.

- Gutachter: **Prof. Ch. Betzel**
- Gutachter: **J.Prof. H.Tidow**

**Tag der Disputation 15. April 2016**

Für die Seele meiner Mutter, meinem Vater und meiner Familie

# TABLE OF CONTENTS

# Abbreviations

| | |
|---|---|
| 2D | Two dimensional |
| Å | Ångström (unit, $10^{-10}$ m) |
| aa | Amino acid(s) |
| AHT | Anhydrotetracycline |
| AMP | Adenosine monophosphate |
| Amp$^R$ | Ampicillin resistance |
| ANSI | American National Standards Institute |
| approx. | Approximately |
| APS | Ammonium persulfate |
| *aquadest*/dH$_2$O | distilled water |
| BLAST | Basic Local Alignment Search Tool |
| bp | Base pair(s) |
| BSA | Bovine serum albumin |
| c | Concentration |
| °C | Degree Celsius |
| CD | Circular Dichroism |
| cDNA | *Complementary* DNA |
| CIAP | Calf intestinal alkaline phosphatase |
| D | Diffusion coefficient |
| DEPC | Diethylpyrocarbonate |
| DESY | Deutsches Elektronen Synchrotron |
| DLS | Dynamic light scattering |
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxyribonucleic acid |
| dNTPs | 2'-deoxynucleoside-5'-triphosphate |
| ds | double strand (DNA/RNA) |
| DTT | dithiothreitol |
| *E.* | *Escherichia* |
| EDTA | Ethylenediaminetetraacetic acid |
| EMBL | European Molecular Biology Laboratory |
| *et al.* | *et alii* |
| EtBr | Ethidium bromide |
| EtOH | Ethanol |
| f.c. | Final concentration |
| gm | Gram (unit) |
| g | Relative centrifugal force (rcf) as multiples of the gravitational acceleration on earth |
| GST | Glutathione S-transferase |
| h | Hour |
| MES | 2-(*N*-morpholino)-ethanesulfonicacid |

| | |
|---|---|
| HEPES | 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid |
| HOAc | Acetic acid |
| Hz | Hertz (unit) |
| I | Intensity |
| IPTG | Isopropyl-β-D-thiogalactopyranoside |
| k | Kilo (multipliedby $10^3$) |
| K | Boltzmann constant |
| K | Kelvin (unit) |
| kb | Kilobase(s) |
| kDa | Kilodalton |
| l | Litre (unit) |
| LB | Luria Bertani |
| M | Molar |
| m | Milli-(multipliedby $10^{-3}$) |
| max | Maximum |
| Ov | *Onchocerca volvulus* |
| Ral-2 | Immunodominant hypodermal antigen |
| L1 | Molting larva stage 1 |
| L2 | Molting larva stage 2 |
| L3 | Molting larva stage 3 |
| mf | Microfilaria |
| PCR | Polymerase Chain Reaction |
| ELISA | Enzyme Linked Immunosorbant Assay |
| Ig | Immunoglobulin |
| TGF | Transforming growth factor |
| CD | Cluster of differentiation |
| MHC | Major histocompatability complex |
| DEC | Diethylcarbamazine |
| SEP$_S$ | Secretory and excretory products |
| Poly-Q | Polyglutamine |
| SAD | Single-wavelength anomalous diffraction |
| DUF | Domain of unkown function |
| Th | T helper |
| ROS | Reactive oxygen species |
| s | Scattering vector |
| sec | Second (s) |
| S | Svedberg (unit) |
| SAXS | Small angle X-ray scattering |
| SDS | Sodium dodecyl sulfate |
| SEC | Size-exclusion chromatography |
| *spec.* | Species |

| | |
|---|---|
| T | Temperature (K) |
| vol | Volume |
| w/v | Weight per volume |
| β-ME | β-mercaptoethanol |
| μ | Micro-(multiplied by $10^{-6}$) |
| TAE | Tris-acetate-EDTA |
| *Taq* | *Thermus aquaticus* (polymerase) |
| TCEP | Tris(2-carboxyethyl)phosphine |
| TE | Tris-EDTA |
| TEMED | *N,N,N',N'*-tetramethylethylenediamine |
| TEV (protease) | Tobacco Etch Virus (protease) |
| Tm | Melting temperature |
| TMD | Transmembrane domain |
| TOF | Time of flight |
| Tris | Tris (hydroxymethyl) aminomethane |
| U | Unit (enzyme activity) |
| UV | Ultraviolet |
| V | Volt |
| v/v | Volume per volume |
| vis | Visible |
| SP | Signal peptide |

## Abbreviationsofnucleotides

| | |
|---|---|
| A | adenine |
| C | cytosine |
| G | guanine |
| T | thymine |
| U | uridine |

**Abbreviations of amino acids**

| | | |
|---|---|---|
| A | Ala | alanine |
| C | Cys | cysteine |
| D | Asp | aspartate |
| E | Glu | glutamate |
| F | Phe | phenylalanine |
| G | Gly | glycine |
| H | His | histdine |
| I | Ile | isoleucine |
| K | Lys | lysine |
| L | Leu | leucine |
| M | Met | methionine |
| N | Asn | asparagine |
| P | Pro | proline |
| Q | Gln | glutamine |
| R | Arg | arginine |
| S | Ser | serine |
| T | Thr | threonine |
| V | Val | valine |
| W | Trp | tryptophan |
| Y | Tyr | tyrosine |

# 1 INTRODUCTION

## 1.1 Introduction to filarial parasites:

The nematodes (roundworms) are a diverse animal phylum. They cover about one million species; untill now only 25,000 species of them have been described (1). Approximately 35 species parasitize in humans, among them are eight known human pathogenic species. Nematodes are one of the most diverse and richest individual groups in the animal kingdom. Many nematodes have been transferred in the course of evolution to the parasitic life in humans, animals, and plants. They are divided into six orders and 16 superfamilies. Superfamily *Filarioidea* is parasitic in vertebrates. Today worldwide, more than 150 million people are infected with these filarial species, where *O. volvulus* represents the main type (2). Filarial worms have two common features. The first is the adult stage, which lives outside the digestive tract. And the second stage is characterized by its need to have an obligate a hematophagous arthropod vector as an intermediary host.

The filariae are typically classified into three groups.

**(A) The lymphatic filariae**, comprising of *w. bancrofti and b. malayi*, within those the adult worms can stay in the lymphatic vessels (2). The infection of these filariae has acute and chronic form. Most of the symptoms resulted from the obstruction of the lymphatic vessels that finally leads to elephantiasis.

**(B) Onchocerciasis**, also called river blindness. The liver blindness caused by the infection by the filarial parasite *O. volvulus*. The mature form of *O. volvulus* inhabits eyes and form nodules under host skin, sometimes these nodules can be seen associated with larger bones.

**(C) Other filarial species**, include *Mansonella perstans, Loa loa, Mansonella streptocerca and Mansonella ozzardi*. These filarial species have low significance as a source of infection. *L. loa* can produce temporary inflammatory swelling, accompanied by eosinophilia and other allergic responses. The *L. loa*'s migrates widely throughout the body and is remarkably disturbing it (3).

Filarial worms resemble a challenge to those organizations which are concerned with public health issues and setting the control programs in affected areas of the world.

## 1.2. Shape and morphology:

In 1975 filarial microorganisms were examined by application of the electron microscopy. A typical adult filariae shape was found as elongated, slender structure with distinct sexual dimorphism; while the females were longer than males. Mature worms were infrequently observed because of their deep location inside the host, therefor the microfilariae is the

# 1 INTRODUCTION

diagnostic stage. Microfilariae are 250 to 300 μm long and surrounded by a sheath in some species (2). They are commonly not infective for non-human vertebrate hosts with no further maturation. In humans they survive for 1-2 years and in some cases up to 15 years. The adult male worms are shorter (2-4 cm) in length than adult female worms (40-50 cm in length). The adult male worms are 0.02 mm in diameter while the adult female worms are 0.04-0.05 mm.

## 1.3. River blindness; Transmission and geographical distribution:

Human onchocerciasis has a worldwide distribution; however approximately 96 % of all patients are in western part of Africa. About 36 countries are recorded by WHO as endemic, out of them 30 are in sub-Sahara Africa (plus Yemen) and the others are in America. Some endemic foci also exist in Guatemala, Mexico, Ecuador and Venezuela. It total about 18 million people around the world (99 % living in Africa) are affected, of them 270,000 are blind and 500,000 people are visually impaired (4). River blindness is considered as the second main infectious cause of blindness in the world. The disease is caused by the infection with *Onchocerca volvulus* and is biologically transmitted by the black fly (*genus Simulium*). The disease is further characterized by eye lesions with a severe itching and pruritic skin condition with socio-economic problems of the rural inhabitants in these endemic areas (5). Currently more than 120 million peoples are under risk and about 18 million are already infected (6). As a result of complications about one million people suffer from visual impairment as a result in complication of onchocerciasis. About 46,000 individuals every year lose their ability to see inconsequence of this disease (7). This disease mainly affects small isolated communities. In highly endemic areas the infection rates can reach 80-100 %, with clinical manifestations peaking at 40-50 years of age. A favorable ecological environment for the intermediate host determines the geographical distribution of the disease. The vector spreads around the fast-flowing rivers and streams, hence the disease is named as "river blindness". It acts as a host for the larval stage vector spreading of the disease (7). There are many sub-species of this vector that are distributed worldwide. The transmission of onchocerciasis carried out by the female fly. The male have no role in transmission. The female flies are able to spread the disease over hundreds of miles with the assistance of the wind from one river to another. It was believed that onchocerciasis has been only associated with the blindness that caused by this disease,

# 1 INTRODUCTION

but now it is obvious that the onchocerciasis affects mental health and causing social problems, leading to an adverse impact on the life quality (7, 8 and 9).

## 1.4. Life Cycle of *Onchocerca volvulus:*

The life cycle of *Onchocerca volvulus* occurs in two different hosts: In black flies and in human. The humans represent atypical and evolutionarily primary hosts while breeds of black fly act as an intermediate host. The life cycle of the parasite within the vertebrate host is still unclear, due to the lack of a truly suitable comparable animal model and the scarcity of the post-mortem information. *Onchocerca* is a long-living parasite going into a five-stage life cycle. The long-term persistence of the worm refers to highly adapted mechanisms of immune evasion.
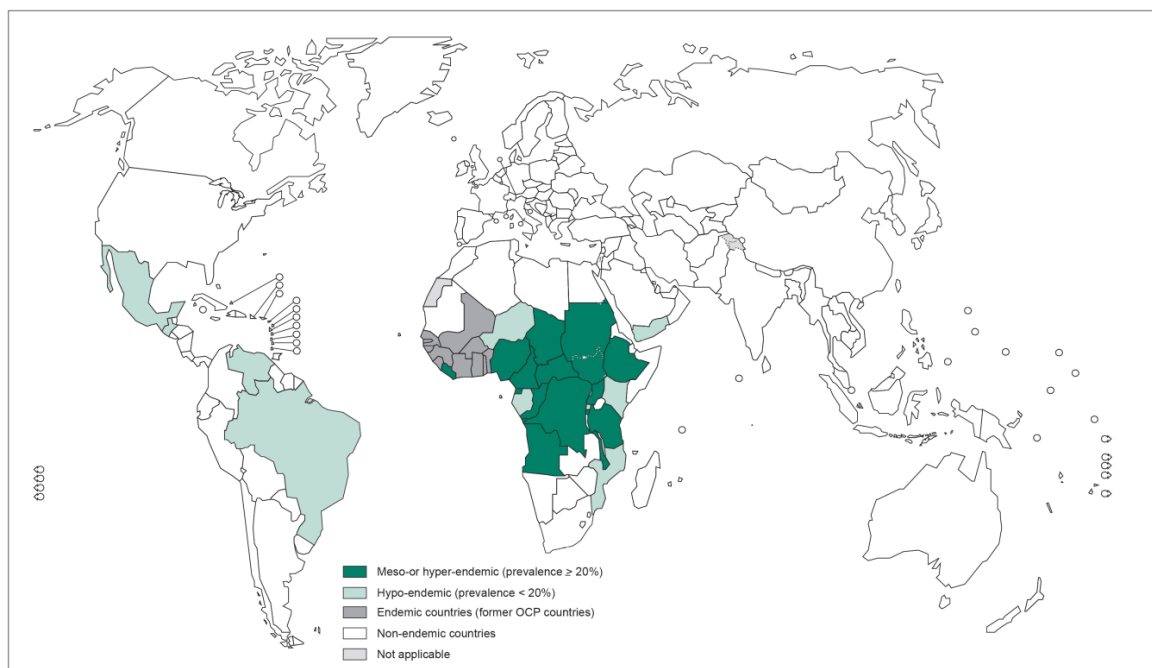


Fig.1: Worldwide distribution of onchocerciais (WHO, 2013).

The *Onchocerca* life cycle is similar to those known for other parasitic filarial parasites with different intermediate hosts. Humans are the only definitive hosts in the life cycle of allthe filarial parasites. In endemic areas nearly 100 % of the population are infected. The life cycle starts when the female blackfly (*Simulium damnosum*) receives the microfilariae during a blood meal from an infected vertebrate host. In the human host, the L3 migrates through connective tissue and skin and occasionally through the eyes. Adult worms survive for years encapsulated in subcutaneous nodules. Often there are several filarial found in one nodule. The females produce daily about a thousand microfilariae, which migrate

# 1 INTRODUCTION

through the lymph of the connective tissue and are eventually washed into the blood stream. In heavily infected patients the number of microfilariae may reaches 100-150 million (10) and they can be detected in blood and urine (11). The microfilariae infect the eye, whereby the principal harmful effects (blindness) are concluded. The positive microscopical examination reveals the presence of the microfilariae between the red blood cells. The size of microfilariae ranged from 220 to 360 µm by 5 to 9 µm and unsheathed, with a life span up to 2 years.

The cycle of development inside the black fly consumes about seven days at high humid temperature, but requires longer time at lower temperatures. The cycle completely stops when the temperature falls below 16 ℃ (12). The *O. volvulus* can also be harmful to the black fly itself; this occurs after massive uptake of microfilaria. The fly may ingest several hundred mf, which can damage the fly's tissues during migration and molting (13, 14).
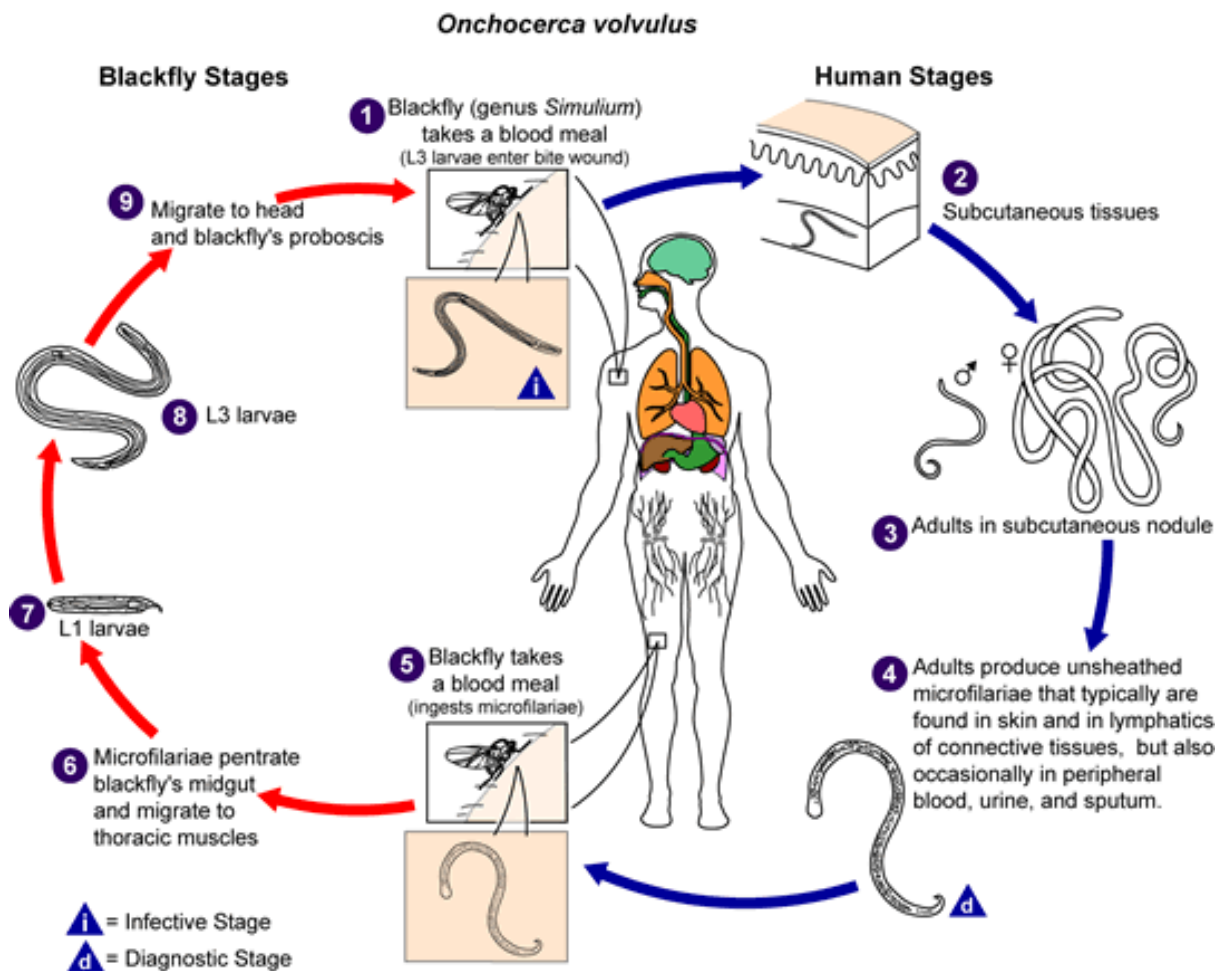


Fig.2: Life cycle of the *Onchocerca volvulus.* Interestingly, the life of adult *O. volvulus* is dependent on the *Wolbachia bacteria. Wolbachia* is living in the cytoplasm of some filarial nematodes and can display, as in the case of *O. volvulus*, a crucial mutualistic relationship with female nematodes (15). Obtained from the CDC Public Health Image Library (16).

# 1 INTRODUCTION

## 1.5. Clinical Symptoms:

The clinical picture observed in human onchocerciasis results from the development of *O. Volvulus* stages inside the human host; this occurs in the skin and eyes. This accompanied by a severe immune response and adverse allergic condition. Over a period of years, the adult female worms become enveloped by host tissue; forming a characteristic subcutaneous nodule (onchocercoma). Reactions to dead microfilariae around these nodules resulted in several adverse conditions. Dermatologic changes are the initial presenting symptoms. The disease characterized by acute popular onchodermatitis of the dermal and epidermal regions (cigarette-paper appearance) appears, which is complicated by itchy skin rashes due to pruritic nature of the disease. Later on, the skin becomes thicker as a result of the breakdown of the elastic tissues and the formation of thick folds. There is also often a depigmentation in advanced cases (leopard skin). The consequent pathology is an outcome of a long-standing interplay between both host and parasite and ending with skin atrophy. The disease is known by severe skin changes and its ocular form, which can damage the eye, leading to irreversible loss of vision. However, an obligatory endosymbiotic bacterium *Wolbachia* is considered to be a potential factor for the development of river blindness (16, 17 and 18). This disease characterized by subcutaneous nodules that mainly seen on the hip, iliac crest, ribs, and skull where the adult male and female worms are encapsulated. Symptoms of onchocerciasis appear when the L3 larvae grows into adult worms. This takes between 9 months and 2 years. This period is known as the prepatent phase. There are no clinical symptoms but the immature worms are able to stimulate some immunological responses (19). The time from infection by L3 to the appearance of clinical signs (the incubation period) is usually longer than the prepatent period and may last for many years (6).



Fig.3: Manifestations forms of onchocerciasis: chronic lichenification with papular dermatitis (left); sclerosing keratitis (right).

# 1 INTRODUCTION

## 1.6. Role of *Wolbachi*a in the pathogenesis of river blindness:

The discovery of the relationship between *Wolbachia* intracellular bacteria with the filarial nematodes leads to a great shift in understanding the biology of the parasites. *Wolbachia* is essential for femal fertilization and development of larvae and embryos. Removal of the endosymbionts by antibiotic treatment causes blockage of worm embryogenesis and development (20). After 6 weeks of doxycycline treatment, the worms contain virtually no *Wolbachia* and have abnormal embryogenesis, and then becomes sterilized. For the worms, this shows that they need *Wolbachia* to survive and fertilize (21). Recently, *Wolbachia* has also been classified as the principle key player of innate and adaptive Th1 immunity. Doxycycline-treated patients showed significantly lower stromal thickness and neutrophil infiltration than untreated patients (22). These results suggest the *Wolbachia* component of *O. volvulus* in the development of eye pathology in onchocerciasis (23).

Eye lesions resulted directly from the invasion of the eyes by the adult worms and host`s inflammatory response that resulted in keratitis, cataracts, secondary glaucoma, chorioretinitis and optic atrophy. There is no favorable location for the worm inside the eye, however mainly the anterior and posterior chambers are affected (24). Also, new clues indicate that white cells activation by *Wolbachia* bacterium contributes to the pathogenesis of the ocular form of the disease (25), as shown in figure 4. Further, the immunological responses to bacterial lipopolysaccharide (LPS) and *Wolbachia*-derived molecules contribute to the development of eye disease (23).
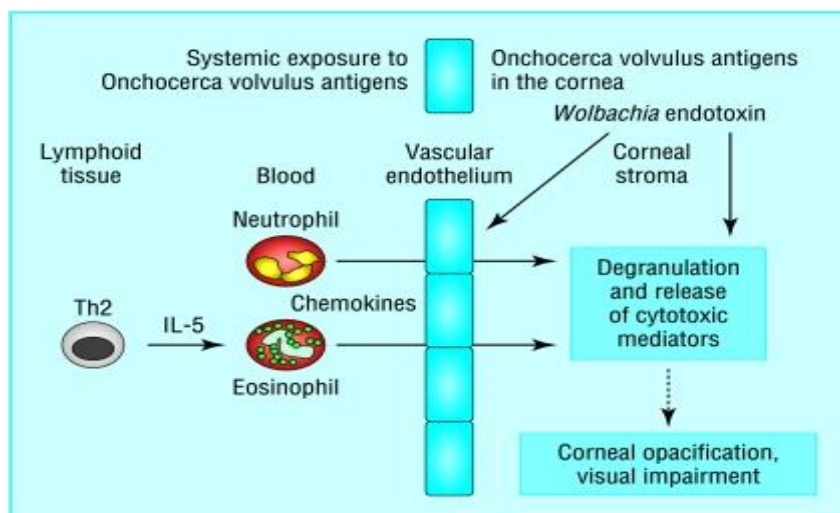


Fig.4: Role of *Wolbachi*a in the pathogenesis of river blindness. Lipopolysaccharide-like molecules from *Wolbachia* endobacteria seem to induce expression of chemotactic cytokines and vascular adhesion molecules that mediate neutrophil recruitment to the corneal stroma, adopted from (23).

# 1 INTRODUCTION

## 1.7 Immunology and Immunopathology:

### 1.7.1. Immunity to Onchocerciasis:

A delicate immunological balance between the parasite and human host can only allow the long-standing persistence of the parasite inside the human body (26). Most of the pathologic symptoms of *O. volvulus* infections resulted from the host immune responses including magnitude and quality of body reactions (27). A scientific discussion is ongoing about the existence and form of the immunity against *O. volvulus* in humans. As mentioned before, the investigations have been hindered so far by the lack of a suitable animal models. However, now adays some progress is already achieved (28, 29).

The presence of a few number of uninfected humans in hyper-endemic areas indicate the presence of immunity against onchocerciasis. Further, immunological investigations try to explain the humoral and cellular responses in these individuals (30, 31 and 32). The results obtained already showed increased IFN-gamma in response to *O. volvulus* antigens (33, 34). However, some other reports relate the immunity to the production of IL-5 (35). Clearly, the existing immunity supports a state of protection. A detailed understanding of the protective immune mechanisms in endemic populations will support more accurate predictions about the eventual success of control programs.

### 1.7.2. Parasite-specific immunomodulation:

The persistence of *O. volvulus* in humans is due to evolutionary processes, referred to as an adaptation and evasion process (36). The aim of these processes is to modulate or inhibit the immue response, in order to obtain an immunological tolerance of host's immune system. The defense mechanisms of the host and protection strategies of the parasites seem to be in a relatively stable balance. This allows up to 15 years persistence of the worm in the human tissue (37). The interaction of the parasite to the host's immune system triggers both a humoral and a cellular immune response in the host.

Understanding how the host response to filarial parasites is highly significant for determining the clinical manifestation of these infections and helping in diagnosis and protection against them. Helminths are presumed to have evolved for survival in their human host and down-regulation of immune cells responses through stimulation of regulatory T cells or the anti-inflammatory cytokines and TGF (38, 39). The successful persistence of the parasite inside the host for a long time is clearly a cellular and humoral immune mechanism. Evidence for similar diminished cellular responsiveness to parasite

# 1 INTRODUCTION

antigens in patients with generalized onchocerciasis is also available (40, 41 and 42), but till now it is not obvious whether this decreased responsiveness is limited to parasite antigens or is broader and less specific (43, 44 and 45). In general, this unresponsiveness happens not because the patients fail to become sensitized to filarial antigens, but because the different modulating mechanisms develop that can specifically suppress responses to these antigens. The mechanisms described include serum suppressor factors that are still incompletely defined (46). The CD4+ Th2 cells, plus the cytokines L-10, IL-9, IL-5, IL-4, and IL-13; the antibody isotypes IgG4, IgG1 and IgE and broad populations of basophils, eosinophils, mast cells and alternatively activated macrophages are the main key player in immunity against parasites (47, 48). Among the mechanisms utilized to avoid immune-mediated regulation are those of suppression, elimination, or block of the immune pathways (49) leading to a decreasing pathology and lowering the tolerance and prolong the persistence of the parasites, which is associated with a weak immune response. Such condition of asymptomatic infection ensures the long-term survival of the parasite inside the host and therefore, secures parasite-feeding and ensure the completeness of the life cycle and successful reproduction (39, 47).

Other modulators from helminths, such as prostaglandins can modulate the T-cell phenotypes (50). The serpins and cystatins are protease inhibitors from helminths that have an immunomodulatory effect. Mammalian cysteine proteases are vital for processing of filarial secretory and excretory products on MHC class II to initiate a proper T cell response. Cystatins from *O.Volvulus* have been proved to inhibit this process (51). Bm-CPI-2, a cystatin from *B. malayi* can stop antigen processing, which result in a low epitopes number exhibited to T cells *in vitro* (52). Studies showed that onchocystatin (rOv17) diminished antigen-driven proliferation of mononuclear cells in a monocyte-dependent manner (53). Bm-SPN-2 is a serpin produced by *B. malayi* and *O. Volvulus* microfilariae, which could stop proteases of human neutrophils (54). The recent genomic study shows that analysis of *L. loa* genes has identified some human cytokines and chemokine antagonists. For example, the genes encoding for the macrophage migration inhibition factor (MIF) families, signaling molecules, members of the interleukin-16 (IL-16) family, transforming growth factor and their receptors, an IL-5 receptor antagonist, a homolog of suppressor of cytokine signalling 7, an interferon regulatory factor and two members of the chemokine-like family were identified (51). In summary, the immunomodulation occurs through the release of soluble substances that degrade, ligate,

mimic or interact with host immune cells and molecules. This inhibits the inflammatory responses and promotes the persistence of worm parasites in host organisms and its interaction with the immune system.

## 1.8. Respiratory burst and antioxidant protection mechanism:

Respiratory burst has an important role in the immune system and helps to degrade internalized particles and bacteria by releasing reactive oxygen species (superoxide radical and hydrogen peroxide) from macrophages and neutrophils. As superoxide is toxic, it is produced in substantial amounts by the enzyme NADPH oxidase to be used in oxygen-dependent killing mechanisms of parasites and other pathogens. The designated lipid peroxidation process can lead to the complete destruction of cell membranes. Oxidative changes may disrupt essential biological processes. This is especially true for uncontrolled oxidation of thiol groups in many enzymes, affecting its activity status and can lead to functional disturbances in cells. *O. Volvulus* is equipped by an extracellular superoxide dismutase; this enzyme protects the worm from being attacked by this reactive oxygen species. Mice deficient with cytosolic SOD are viable but suffer from short life span, cancer, atrophy, cataracts, haemolytic anemia, thymic involution, rapid age-dependent and decline in female fertility (55). Catalase (CAT), peroxiredoxins (PRX), glutathione peroxidases (GPX) and thioredoxin peroxidases (TRX) are available for enzymatic neutralization of oxidants. SODs dismutate the superoxide anion radical, while the group of peroxidases and catalases implement hydrogen peroxide to water and oxygen. Inhibiting of detoxification enzymes would prevent the protective effect and thus lead to the weakening or killing the parasite. A prerequisite for this step is the precise characterization of the detoxifying enzymes.

## 1.9. Disease control:

Many strategies to eliminate onchocerciasis have been initiated, based on vector control and massive treatment with the microfilaricidal drugs. The hope of scientists in future is to use perform a drug treatment in combination with a prophylactic vaccine. Therefore, many efforts were done to characterize antigens playing critical roles in parasitic development or transmission. This parallel approach can help to eradicate this disease and decrease the incidence of disease in the future.

# 1 INTRODUCTION

## 1.9.1. Treatment of disease:

Although most of the pathology is caused due to reaction of dead microfilaria, patients who are infected with *O. volvulus* must be treated to prevent long-term complications. The treatment has been a problem over the last decade, and the result is unsatisfactory. Ideal drugs should affect both adult and larval stages, with minimal side effects. Suramin and diethylcarbamazine were used before to control the disease (56). Suramin was used as an intravenous injection as a filaricidal, but the drug had toxic effects. DEC (diethylcarbamazine) is used orally as an effective compound, but has a life-threatening severe allergic-like clinical response (57). In some patients, diethylcarbamazine complicates the condition and can accelerate the complete loss of vision (58). After that ivermectin was approved, this is applied every 12 months or till the lesions disappear (57). Ivermectin is a semi-synthetic macrocyclic lactone, defined as fermentation byproduct of *Streptomyces* avermitilis (59). It is estimated as an important step of parasitic disease treatment, as it is an effective microfilaricidal with low side effects (56, 60 and 61). The drug is already used successfully in mass treatment programs (62). It dramatically decreasesthe number of mf and improves skin lesions unless for depigmentation (63, 64). Recently, the administering of the antibiotic tetracycline together with ivermectin resulted in an improved macrofilaricidal effect. This improvement might be due to the effect of tetracycline on *Wolbachia* endobacteria (65). This combination program was usefully applied to increase the efficacy to control human onchocerciasis. Antihistamines or corticosteroids may benefit to relieve the allergic reactions from dead microfilariae and *Wolbachia* products that following have seen the treatment of onchocerciasis (66). Avermectins block the specific glutamate-gated chloride channel muscles and invertebrate nerves (67). A few cases of recurrence were noticed in a small proportion of treated patients with ivermectin after only one month of treatment. This may raise the possible role of the immune responses in the microfilaricidal mechanism of ivermectin. Ivermectin is today classified as a safe drug; many studies demonstrated the absence of immunopathological responses induced by a long-term repeated ivermectin treatment (32). Previous articles report that ivermectin may be needed to enhance immunity against onchocerciasis (68). Ivermectin treatment should considered when the patients are co-infected with the filarial parasite *Loa loa*, due to the involvement of the central system (69, 70). In areas where the lymphatic filariasis and onchocerciasis are co-endemic, the introduction of the drug albendazole is necessary, as a part of the control program (71). Nowadays, a lot of efforts are concentrating to discover new drugs that can kill adult stage

# 1 INTRODUCTION

without adverse side effects, or, at least, have long-term sterilizing action. A new emerging potential drug called moxidectin has a macrofilaricidal effect. It has been used as safe and well-tolerated drug in humans (72). However, additional treatments of *Wolbachia* are necessary (73).

## 1.9.2. Vaccine development:

The Edna McConnell Clark Foundation (EMCF) contributed about 21.6 million dollars to initiate the development of effective control tools. However, for vaccine development, it is necessary to understand the mechanisms of the protective immunity to onchocerciasis in humans. Another challenge is the difficulty in obtaining a live fresh larval stage. An alternative approach is to isolate potentially protective antigens include targeting molecules that are considered to be essential for the infection process or play a critical role in worm protection against the immune system. Complementary DNA (cDNA) libraries have been identified for the adult and larval stage of the parasite to be used as a source of recombinant antigens (74). Various strategies to identify potential vaccine candidates against river blindness resulted in cloning of recombinant antigens, which confer protection in vaccinated mice or other animal models. The identification and characterization of larval antigens that play a role in stimulating host immunity against the infection were determined to be an essential first step. Selected recombinant *O. volvulus* antigens were shown previously to exhibit varying degrees of protection. Studies of other filarial worms have shown that the targets of the protective host-immune response were the L3 and the molting L3 (75, 76) and that excretory/secretory (ES) products from the developing parasites induce host protection (77). These studies also suggested that the protective immune responses could inhibit the development of L3 to L4 (76, 78 and 79). Assuming that there are common protective mechanisms against filariae and that the target antigens are similar, it was also recommended that antigens, which appear to be protecting against parasites, causing either lymphatic filariasis or dirofilariasis, should also be further investigated as potentially relevant to onchocerciasis. In addition, it was postulated that an effective *O. volvulus* vaccine should contain proteins found in the ES products or on the surface of larval stages.

## 1.9.3. Identification of protective *O. volvulus* antigens:

Two major strategies were applied to identify *Onchocerca* larval and adult target-antigens that could be possible vaccine candidate. The first strategy was based on the potential

# 1 INTRODUCTION

involvement of antibodies in protective immunity; thus, attempts were made to identify the target proteins by immunoscreening cDNA libraries using immune sera from human or animal hosts. Although this approach was used with success, it did not take into account carbohydrates and other non-protein determinants important in other helminth systems (80–81), which could be crucial to the generation of protective immunity. 26 recombinant antigens were identified by immunoscreening and were further tested in the *O. volvulus* mouse model, 12 induced partial but significant protection (39–69 %) in the presence of Block Copolymer (BC), alum or Freund's complete adjuvant (FCA). Five of the protective antigens (Ov-CPI-2, Ov-ALT-1, Ov-RAL-2, Ov-FBA-1 and Ov-B8) were protective in multiple experimental conditions. The second strategy was to isolate molecules that were thought to be vital in the infection process. These would include proteins with vital metabolic functions or defense properties, which permit the parasite to survive in immunocompetent hosts. Targeting such molecules would block the establishment of the parasite in the host.

## 1.9.4. Other approaches to rational drug development:

Rational approaches to identify drug target candidates were aimed at cloning: for example [1] enzymes that might have a role in defense against host-killing mechanisms, [2] proteins involved in immune evasion and [3] hidden antigens [4] proteins involved in vital processes like metabolic pathways or ions transportation like chloride channels. Helminth antioxidant and detoxification enzymes [thioredoxin peroxidases (TPX), glutathione peroxidases (GPX), catalase (CAT), superoxide dismutases (SOD) and glutathione S-transferases (GST)] have been postulated to play a role in the defense systems of parasites against host-killing mechanisms (82). Although these enzymes were initially studied to establish their potential as targets for chemotherapeutic intervention, their immunogenicity in humans and their anatomical localization in the parasite suggested their potential as targets for immunological interference (83). An example of this is the 28 kDa GST of *Schistosoma mansoni*, which is a primary vaccine candidate in trials against schistosomiasis (84). Drug discovery to treat onchocerciasis is a big challenge. Variant compounds from different sources were screened against *Onchocerca gutturosa* adult male worms *in vitro*; *Onchocerca lienalis* microfilariae and *Onchocerca lienalis* microfilariae in mice. Since 2006 about 5000 compounds have been tested *in vitro*, producing ~100 promising compounds for follow-up investigations and around six lead compounds for further optimization were identified (85).

# 1 INTRODUCTION

## 1.10. History of protein crystallography:

The method of protein crystallography is established after the discovery of X-rays by Conrad Röntgen (he obtained the Nobel prize in Physics for this discovery in 1901), followed by the subsequent developments by Max von Laue, who was first to observe diffraction of X-rays and revealed the wave nature of X-rays. These discoveries were followed by the experiments undertaken by William Henry Bragg and his son, William Lawrence Bragg, who showed that X-ray diffraction could be used in the determination of the atomic structure of matter. Until the early 1930s the British physicist Paul Dirac and his assistant Dorothy Hodgkin (she received the Nobel Prize for Chemistry in 1964), succeeded to obtain sharp diffraction images from protein crystals and could further solve the first three-dimensional protein structure. However, a large computational effort was required, which came only easier to cope with the development of computers. The first structure of a protein (myoglobin) was elucidated by John Kendrew using X-ray crystallography. He obtained Nobel Prize for Chemistry in 1962, along with Max Perutz, who developed the method decisively. The X-ray crystallography provides a tool which revolutionized our understanding of the structure of matter ranging from minerals, pharmaceutical materials, DNA, proteins to viruses.

## 1.11. Theory of protein crystallization:

The laws of physical chemistry and thermodynamics control the process of crystallization. In order to crystallize a protein, we need to bring the solution into a supersaturated state. In the protein crystallization two steps are necessary. The first step is the nucleation, which is followed by crystal growth. Under appropriately chosen conditions nucleation is taking place in the supersaturating range of the phase diagram. Initial nuclei are formed only in a supersaturated solution (unstable area); unfavorable for the growth area. The emergence of many small crystals can result. To grow large and X-ray suitable crystals (100-500 μm), the solution should be kept in a metastable zone. Ideal is a quick generation of the unstable region with a subsequent drop in the protein concentration, followed by a long maintenance of the solution in the metastable area (arrows in the diagram in figure 6). The transition from, the stable to the supersaturated solution, can be achieved mainly by changing the protein concentration or precipitant concentration. It can furthermore be obtained by the influence of salts, organic solvents, pH values, and polymers such as polyethylene glycol (PEG).

# 1 INTRODUCTION
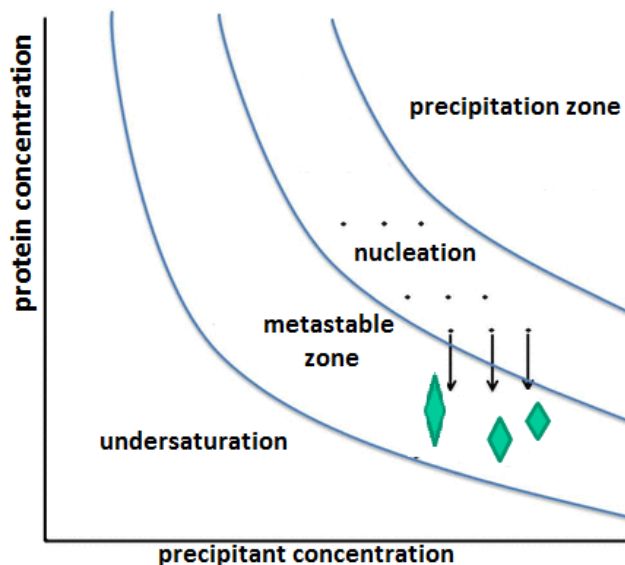


Fig.6: Crystal formation as a function of protein concentration (Y-axis) and precipitant concentration (X-axis). The arrows represent the optimal concentration-time curve during the crystallization process.

The basic requirements for the production of protein crystals are sufficient quantities of highly pure homogenous non aggregating protein (proteins can be purified by a combination of precipitation methods, chromatography or preparative electrophoresis separated from other proteins). The crystallization conditions are found by mixing a highly concentrated protein solution (2-40 mg protein/ml) with different precipitant solutions. They consist mostly of very high concentrations of salts, alcohols or polyethylene glycol (PEG). The most common way to grow protein crystals is by applying vapor diffusion techniques. In this case, a protein solution is mixed with the precipitant in a small sample drop and then placed in a larger well containing a second less efficient solvent, is then sealed and left to stand over days, for weeks, or even months at a constant temperature. This method is used because it allows gentle and gradual changes in the concentration of protein and precipitant, which can promote the growth of large and well-ordered crystals.

## 1.12. The X-ray diffraction and Bragg's law:

The X-ray diffraction is caused by the interaction of the electromagnetic waves with the matter inside the crystals, and in particular with the electrons. These waves scattered by the electrons become a small X-ray source. The scattered waves from all the electrons within each atom are merged to each other, giving the diffracted waves from all atoms. When the

scattered waves are combined, they may result in stronger or weak or even cancel each other. The X-ray detector detects those who get stronger. Although these waves cancel one another out in most directions through destructive interference, they add constructively in a few specific directions, these waves are determined by **Bragg's law**: $n\lambda = 2d\sin\theta$

Here **d** is the spacing between diffracting planes, **θ** is the incident angle, **n** is any integer, and **λ** is the wavelength of the beam.

Next step after the data collection is the processing of the data, which aims to obtain the relative intensities of the diffracted X-ray beam and to solve the protein structure. Different computer programs are used for this purpose, like for example the CCP4 package.

## 1.13. Structural biology and drug development:

X-ray crystallography is a technique that uses X-rays to determine the high resolution structures of proteins. However, proteins must be present in a crystallized form to solve the protein structure. When protein crystals are exposed to X-rays, the X-rays are scattered in specific directions displaying a diffraction pattern that cover information about the structural arragments in the crystals, as well as about the detailed 3D structure. X-ray crystallography is now applied routinely to analyse the structure of biomolecules and to determine how pharmaceutical drugs bind to their target proteins and what changes might be considered to inhibit or improve their functions. More progress was achieved in the last decade to make the process of solving protein structures more convenient for solving the structure of vital proteins that play an essential step in a disease progress, or perform a critical process in the organism and perception. It is also questioned, how each special function is correlated to the specific protein's structure. Based on this evidence, scientists can design and screen small molecules that can bind to special and defined parts of the protein. These small molecules can affect the protein's function; potentially stop its contribution to diseases. Some of these small molecules and inhibitors and undergo clinical trials to evaluate their potential as new possible drugs to be approved.

Although magnificent advancements in structural biology over the last decade, about 2 % of the structures in the protein data bank emerge from eukaryotic parasites and less than 0.5 % come from multicellular parasites. Nevertheless, nearly one third of the global human diseases originate from parasites. It is time to spend more efforts and resources in structural biology to succeed in the fight against parasitic diseases. Using the benefits of

# 1 INTRODUCTION

the recent technological and methodological progresses and applying more concerted effort to solve the macromolecular structures from parasite pathogens would provide more
 insights into the vital function of the proteins, in the end; this would recommend new strategies for controlling infections.

# 2 AIM OF THE WORK

## 2. AIM OF THE WORK:

In 1985, the Edna McConnell Clark Foundation launched a 15-year research program on immunological and molecular biology approaches to investigate the potential existence of protective immunity towards *O. volvulus* in humans and its mechanisms to identify and characterize promising vaccine antigens and to develop animal models. As a result of this program some partially protective antigens could already be identified. However, till now, its functions and 3-dimensional structures characterized on the molecular level are unknown or only poorly presented. In terms of this Ph.D. research project, it was intended to contribute towards the understanding of the structure-function-relationship of selected proteins. The following proteins are of particular interest and were targeted:

• OvEC-SOD is a secreted Superoxide-Dismutase (21 kDa homodimer) important to resist the toxic free radicals within the environment around the parasite. SODs are metal-containing enzymes that catalyze the converting of superoxide into oxygen and hydrogen peroxide. Thus, these proteins are important for the antioxidant defense mechanism in nearly all cells exposed to oxygen. In previous studies, SOD was identified as a target for drug action and some studies showed that certain chemical modifications at the 2-carbon (2-OH, 2-OCH$_3$) of the certain estrogen derivatives can lead to SOD inhibition (86). This implicates that selective inhibition of SOD is possible and may be a promising strategy for the selective killing of parasites, suggesting the role of SOD as an emerging therapeutic target. Therefore, the determination of the OvEC-SOD 3D structure is one of the most important outcomes of this project and a very important step towards designing a prodrug that interferes with OvEC-SOD without affecting human SODs.

• Another interesting protein for drug discovery investigations is called Ov Ral-2, an immunodominant hypodermal antigen 17 kDa protein, present in both adult and larval stages of the parasite. Ov Ral-2 is a protein belonging to the SXP/RAL-2 family; it is present exclusively in nematodes and so far no homology is found in human or outside the nematodes. Therfore these proteins may be an interesting target for controlling these parasites. Structural and functional informations were so far not available until the 3D structure of *Anisakis simplex* Ani s 5 was determined in 2014. This structure can be employed for modeling Ov Ral-2. The OvRal-2 shows approx. 44 % sequence homology with *Anisakis simplex* Ani s 5. Considering the complete absence of functional information about these proteins, any data about the structure-function-relationship will be supportive,

# 2 AIM OF THE WORK

as they can guide insights about its importance in the parasite life cycle. Solving the 3D structure of Ov Ral-2 opens new ways to investigate the possible function of Ov Ral-2 and other protein families. Besides this evident, the data may lead to the finding of new therapeutic and diagnostics tools and will collaborate towards a better understanding the immunotherapeutic strategy.

The aim of this work was to analyze the structure-function-relationship of the proteins mentioned before, applying complementary analytical techniques, as particular X-ray crystallography and small-angle solution scattering (SAXS), and contributing toward the discovering of a novel therapeutic agents and new vaccines.

# 3 MATERIALS AND METHODS

## Materials and Methods:

### 3.1. Material and devices:

#### 3.1.1. Devices:

| | |
|---|---|
| **Centrifuge** | Centrifuge 5804R/5810R/5415R/5424 (Eppendorf, Germany), Centrifuge Minispin® Plus (Eppendorf, Germany), Optima TL ultracentrifuge (Beckman Coulter, USA) |
| **Incubator** | 37-30 °C incubator Kelvitron® T (Thermo scientific, USA), 4 °C incubator (Rubarth, Germany), 20 °C incubator (Rubarth, Germany) |
| **PCR machines** | UNO II (Biometra, Germany) |
| **Spectrophotometer** | GeneQuant 1300 (GE Healthcare, UK), Nanodrop 2000c (Thermo Scientific, Peqlab, Germany), GENios microplate reader (Tecan, Schweiz) |
| **Thermomixer** | Thermomixer comfort (Eppendorf, Germany) |
| **Sonifier** | Branson Sonifier 250/450 (Emerson Electric Co, USA) |
| **Micropipette** | Micropipette Research (Eppendorf, Germany) |
| **Balance** | TE3102S (Sartorius AG, Germany), LP224S-0CE (Sartorius AG, Germany) |
| **pH meter** | SevenEASY (Mettler Toledo, USA) |
| **Microwave** | Microwave MR-6450 (Hitachi, Japan) |
| **Thermocycler** | Mastercycler® gradient, Mastercycler® personal (Eppendorf, Germany) |
| **Electrophoresis power supply** | EV 231 (Peqlab, Germany), Power PAC 200 (Bio-Rad, Germany) |
| **SDS-PAGE** | EV734 Power Supply (Consort, Belgium) |
| **Microbalance** | Sartorius CP224S-OCE (Sartorius, Germany) |
| **Hot-plate magnetic stirrer** | VMS-A (VWR, USA), MR 3001 (Heidolph, Germnay) |
| **shaker** | IRC-1-U (Adolf Kühner AG, Switzerland), Innova® 43/43R (New Brunswick Scientific, USA), Innova® 4330 (New Brunswick Scientific, USA), GFL 3017 (GFL, Germany) |
| **CD spectrometer** | J-815 (Jasco, UK) |
| **Freezer (-20 °C)** | Liebherr premium (Liebherr, Germany) |
| **DLS instrument** | SpectroSIZE 300 (Xtal-Concepts) |
| **Acryl amide gel chamber** | SE275 (Hoefer) |
| **Pipetting robot** | Honeybee 961 (Genomic Solutions, USA), Oryx 8 (Douglas, UK) |
| **UV-light source** | CrystalLIGHT 100 (Nabitec, Germany) |
| **SONICC** | (Benchtop, Formulatrix.Inc, USA) |
| **Crystal plate incubator** | RUMED 3001 (Rubarth, Germany) incubators |
| **Crystal imaging** | CrystalScore (Diversified Scientific Inc., USA), microscope SZX12 with camera DP10 (both Olympus, Japan) |

**3.1.2**. Plastic consumables reaction tubes and pipette tips were obtained from Sarstedt (Germany) and Eppendorf (Germany).

**3.1.3 Plasmids:** Two different vectors were used in these study:

# 3 MATERIALS AND METHODS

**(1) pJC40:** is a high-copy-number expression vector, encodes an amino-terminal histidine sequence of 10 residues and a factor Xa cleavage site, allowing the purification of a recombinant protein by metal chelate chromatography.

**(2) pASK-IBA16:** allows the expression of *Strep*-tag®-fusion-proteins in *E.coli*. The vector carries the inducible tetracycline promoter/operator for the regulated expression of proteins, the ompA signal for periplasmic secretion of the recombinant protein, the *Strep*-tag® for N-terminal fusion to the recombinant protein and the ampicilin resistance cassette. It can be used with any *E. coli* strain because the *tet*-promoter works independently from the genetic background of *E.coli*.

## 3.1.4. Primers for cloning

| Target | vector | type | Sequence |
|--------|--------|------|----------|
| OvEC-SOD | pASK-IBA16 | Forward | **G GGC AAG AAT TCC ATG GCT AGA AGA GCA GTA GCA GT** |
| | | Reserve | GGG CAA GGA TCC TCA AGC AGC AAT GCC AAT AAC ACC |
| Ov Ral-2 | | Forward | **AAGAATTCCCGTGATGAACGTGAAATA** |
| | | Reserve | AAGGATCCTTATTCTTGTGGGCTCAATGCATTAAT |
| OvEC-SOD | pJC40 | Forward | **GA ACC ATG GAA CAT CAT CAT CAT CAT CAT GCT AGA AGA GCA GTA GCA** |
| | | Reserve | GGG CAA GGA TCC TCA AGC AGC AAT GCC AAT AAC ACC |
| Ov Ral-2 | | Forward | **AACCATGGAACATCATCATCATCATCATGGCCGTGATGAACGTGAAATA** |
| | | Reserve | AAGGATCCTTATTCTTGTGGGCTCAATGCATTAAT |

## 3.1.5. Bacterial strains:

| Strain | Advantage | resistance/ selection marker | supplier |
|--------|-----------|------------------------------|----------|
| BL21 Star™ (DE3) | Transformation efficiency: 1–5 x $10^7$ cfu/µg pUC19 DNA | / | Invitrogen |
| BL21 DE3 | Deficient in proteases Lon and OmpT | / | Invitrogen |
| BL21 AI | High yields of protein | / | Invitrogen |
| DH5α™ | Greatly increased plasmid yield and quality due to endA1 mutation | Nalidixic acid | Invitrogen |
| XL10-Gold | For transformation of large DNA molecules with high efficiency. | Nalidixic acid, tetracycline, chloramphenicol | Stratagene |
| B834 (DE3) | Methionine auxotrophic | / | Novagen |

## 3.1.6. Buffers and solutions:
For all buffer preparations double distilled water was used, pH was adjusted with HCl or NaOH.

# 3 MATERIALS AND METHODS

### General buffers:

| Solution | Composition |
|---|---|
| Agarose gel electrode buffer | 80 mM Tris, 40 mM acetic acid, 2 mM EDTA |
| DNA loading buffer | 30 % glycerol, 0.25 % Bromophenol Blue, 0.25 % Xylene Cyanol |
| 2x protein loading buffer | 500 mM Tris-HCl, pH 6.8, 5 mM DTT, 10 % SDS, 0.5% Bromophenol blue, 50 % Glycerol |
| 1x TBE buffer | 90 mM Tris, 90 mM borate, 2 mM EDTA, pH 8.0 autoclaved |
| Crystallization buffer/OvEC-SOD | 100 mM NaCl, 50 mM MES, 1 mM $ZnCl_2$, 0.1 mM $Cu_2So_4$, pH 6.5 |
| Crystallization Buffer/Ov Ral-2[17-148] | 50 mM NaCl, 50 mM Tris, pH 8 |
| Buffer W | 100 mM Tris, 100 mM NaCl, pH 8.0, autoclaved |
| LB-medium/agar | 10 % Bacto-Trypton, 0.5 % Bacto-Yeast-Extract, 1.0 %, NaCl in $dH_2O$, Autoclaved. For LB-agar: 1.5 % agar in LB-medium |
| Selective antibiotic | 100 mg/ml ampicillin in 10 ml water; stock solution 34 mg/ml chloramphenicol in 100 % ethanol (abs.); stock solution |
| IPTG | 1 M in $dH_2O$, Autoclaved; stock solution |
| Anhydrotetracycline | 2 mg/ml in DMF; stock solution |
| Protease inhibitor | 1 M PMSF in isopropanol |
| His-lysis buffer | Buffer W with 20 mM imidazole (imidazole addition after autoclaving |
| His-washing buffer | Buffer W with 50 mM imidazole (imidazole addition after autoclaving) |
| His-elution buffer | Buffer W with 250 mM imidazole (imidazole addition after autoclaving) |
| Strep-elution buffer | 2.5 mM desthiobiotin (buffer E, IBA, Germany) in buffer W |
| TEV storage buffer | 1 mM HABA in buffer W |
| Regeneration buffer | 1 mM HABA in buffer W |

### Competent cells:

| Solution | Composition |
|---|---|
| $CaCl_2$ | 100 mM $CaCl_2$, autoclaved |
| $CaCl_2$ buffer | 100 mM $CaCl_2$, 10 % glycerol, (glycerol addition after autoclaving) |

### Agarose gel electrophoresis:

| Solution | Composition |
|---|---|
| 1 % Agarose | 1 g in 100 ml in $dH_2O$ |
| Loading dye | 0.05 % bromophenol blue, 0.25 % xylene cyanol, 1 mM EDTA, 50 % glycerol |
| Ethidium bromide | 10 mg/ml in $dH_2O$ |

### Molecular weight markers:

| Solution | Composition |
|---|---|
| Protein marker | Bench Mark prestained protein ladder (Invitrogen, Karlsruhe, Germany) |
| DNA marker | 1kb DNA ladder, (Invitrogen, Karlsruhe, Germany) |

# 3 MATERIALS AND METHODS

### SDS-PAGE preparation:

| Solution | Composition |
|---|---|
| 2x sample buffer | 2.0 ml (10 %) SDS (w/v), 1.25 ml (0.5M) Tris, pH 6.8, 2.5 ml (100 %) glycerol, 0. 2 ml (0.5 %) bromophenol blue (w/v), 0.05 % β-mercaptoethanol (v/v) or DTT in 3.55 ml dH$_2$O |
| APS | 10 % in dH$_2$O |
| Electrode buffer 10× | 30.0 g Tris, 1440 g Glycine, 10.0 g SDS, in 1L dH$_2$O |
| Separating gel buffer | 1.5 M Tris/HCl, pH 8.8 in 100 ml dH$_2$O |
| Stacking gel buffer | 0.5 M Tris/HCl, pH 6.8 in 100 ml dH$_2$O |
| SDS buffer | 10 % (w/v) in dH$_2$O |
| Staining solution | 0.25 % (w/v) coomassie brilliant blue G-250, 25 % (v/v) 2-propanol, 10 % (v/v) acetic acid |
| Distaining solution | 20 % (v/v) acetic acid |
| Circular dichroism buffer | 50 mM KH$_2$PO$_4$, 10 mM NaF, pH 7.5 |
| Protein quantification buffer | Bradford reagent 0.01 % Coomassie Brilliant Blue G-250 (w/v), 4.7 % ethanol (w/v), 8.5 % (v/v) H$_3$PO$_4$ |
| Ethanol buffers | 25, 50, 75 % ethanol (den.) in dH$_2$O |

### Ni-NTA agarose regeneration:

| Solution | Composition/Supplier |
|---|---|
| SDS buffer | 2 % (w/v) in dH$_2$O |
| EDTA buffer | 100 mM, pH 8.0 in dH$_2$O |
| NiSO$_4$ buffer | 100 mM in dH$_2$O |
| Regeneration buffer | 6 M guanidine chloride, 0.1 M acetic acid |

### Western Blot:

| Solution | Composition / Supplier |
|---|---|
| **Transfer buffer** | 25 mM Tris, 192 mM Glycin, 20 % (v/v) |
| | Isopropanol, pH 8.3 |
| **Marker** | PageRuler Plus Prestained Protein Ladder |
| | #26619, 10-250 kDa (Thermo Fisher Scientific, |
| | Germany) |
| | PageRuler Unstained Protein Ladder |
| | (Fermentas, Germany) |
| **Nitrocellulose membrane** | Roti®-NC, 0.2 μm (Carl Roth, Germany) |
| **Ponceau staining solution** | 0.2 % (w/v) Ponceau-S |
| | 2 % (v/v) Acetic acid |
| **PBS (Phosphate buffered saline)** | |
| | 140 mM NaCl |
| | 25 mM KCl |
| | 0.5 mM MgCl$_2$ |
| | 1 mM CaCl$_2$ |
| | 10 mM Na$_2$HPO$_4$ (pH 7.5) |
| **TBS (Tris buffered saline)** | 1 M Tris/HCl, 750 mM NaCl, pH 7.5 |
| **PBS-T/TBS-T** | 1x PBS/1x TBS |

# 3 MATERIALS AND METHODS

| | |
|---|---|
| | 0.05 % (v/v) Tween 20 |
| **Blocking solution** | 4 % (w/v) BSA (Bovine serum albumin) |
| | in PBS or TBS buffer |
| **Antibodies and WB detection:** | |
| **1st antibodies** | Murine Anti-Strep-tag II antibody, IgG1; |
| | #2-1507-001, 0.2 mg mL$^{-1}$ in PBS (IBA, Germany) |
| | final dilution: 1:2000 |
| | Murine Anti-His-tag antibody, IgG1; #65250, |
| | 0.5 mg ml$^{-1}$ in PBS, 0.09 % (w/v) NaN$_3$ |
| | (BioLegend, US) |
| | final dilution: 1:10000 |
| **2nd antibodies** | Goat anti-mouse IgG-HRP (horseradish |
| | peroxidase) conjugated, #SC-2005 |
| | (Santa Cruz Biotechnology, US) |
| | final dilution: 1:3000 |
| | Goat anti-mouse IgG-AP conjugated, #A3562 |
| | (Sigma, Germany) |
| | final dilution: 1:30000 |
| **AP reaction buffer** | 100 mM Tris/HCl, 4 mM MgCl$_2$, pH 9.5 |

**Protein labeling:**
**Feed-back inhibition amino acids mix:**

1.0 g of lysine, threonine, phenylalanine
0.5 g of leucine, isoleucine, valine
0.5 g of L (+) selenomethionine (Sigma)
The mix was divided into ten 0.5 g portions and the amino acids were added to the culture flasks

**Minimal medium (per liter):**

1 L M9 medium
2 ml 1M MgSO$_4$ (2 mM)
20 ml 20 % glucose (0.4 %)
Antibiotic (s)
1ml vitamins 1000x
10ml trace elements100x

| 1L 100x trace elements | 500ml 1000x vitamins |
|---|---|
| 5 g EDTA, 0.8 g FeCl$_3$, 0.05 g ZnCl$_2$, 0.01 g CuCl$_2$ , 0.01 g CoCl$_2$, 0.01 g H$_3$BO$_3$, 1.6 g MnCl$_2$, Some Ni$_2$SO$_4$, Some molybdic acid pH adjusted to 7.0 with NaOH, sterile filtered | 0.5 g riboflavin<br>0.5 g niacinamide<br>0.5 g pyridoxine monohydrate<br>0.5 g thiamine<br>sterile filtered, chilled |

# 3 MATERIALS AND METHODS

### 3.2. Molecular Biology methods:

### 3.2.1. PCR (polymerase chain reaction):

PCR is a widely applied technique to amplify a single copy or a few copies of a segment of DNA. PCR was performed for DNA fragment amplification using *pfu/taq*-polymerase (Invitrogen, USA). The reaction was carried out by applying a thermocycler (Eppendorf, Germany) and the samples were applied according to the standard's protocol. For the reaction, primers (Metabion, Germany) were diluted to a final concentration of 10 pM and 1 μl of each forward and reverse, 1 μl (approximately 100 ng) template was used. The first step in the amplification reaction was denaturation for 10 min at 96 °C followed by 30 cycles of denaturation for 45 sec at 96 °C, annealing for 1 min at 55 °C (based on the oligonucleotide annealing temperature, determined by the supplier) and followed by elongation for 45 sec at 72 °C (this step could be modified based on the number of base pairs to be amplified, as *Taq*-polymerase is able to amplify 1,000 base pairs per minute). When the reaction was carried out overnight, samples were stored at 4 °C within the PCR device. The PCR products were applied to 1 % agarose gel. The components listed below were typically mixed in a 0.5 ml reaction tube for a PCR and placed in a thermocycler.

| components | amount |
|---|---|
| DNA template up to 100 ng | 1 μl |
| Polymerase buffer (10×) (NEB) | 5 μl |
| dNTPs, (2 mM) | 5 μl |
| Forward primer | 1 μl |
| Revers primer | 1 μl |
| *taq* polymerase (2.5 U μl$^{-1}$ / 1 U μl$^{-1}$) (NEB) | 1 μl |
| MgCl$_2$ (25 mM) | 5 μl |
| DMSO | 1 μl |
| dH$_2$O | ad **50 μl** |

### 3.2.2. DNA purification:

To purify PCR products from salts and DNA-fragments, the purification of the nucleic acid after PCR was performed using the PCR clean–up kit (NucleoSpin Extract II Kit, Macherey-Nagel). All steps were performed according to the manufacturer's specifications. Instead of elution buffer, 50 μl dH$_2$O were used to elute the DNA.

### 3.2.3. Restriction digestion of the DNA fragments:

Restriction enzyme digestion was used to either prepare DNA fragments for ligation preparation into a plasmid or to examine the success of the ligation. In the cloning procedure, all fragments were ligated into pJC40 and pASK-IBA16vectors (addgene,

# 3 MATERIALS AND METHODS

Germany). PCR products and vectors were digested according to manufacturer's protocol for 2 h at 37 °C. After digestion, the vectors were dephosphorylated by an addition of 1 µl calf intestinal alkaline phosphates (CIAP) followed by incubation at 37 •C for 1 h, while the digested DNA fragments were stored on ice.

The PCR products and vectors were purified (see DNA purification) separately and eluted in 50 µl dH$_2$O for the PCR products and 30 µl dH$_2$O for the vectors. To verify success of cloning, 0.5 µg of DNA preparation after transformation in *E. coli*, were digested with relevant restriction endonucleases, typically BamH1 and EcoR1 (NEB, USA) and visualized by an agarose gel.

## 3.2.4. Agarose gel electrophoresis:

Agarose electrophoresis is a technique used to separate DNA fragments based on their size. Negatively charged DNA is attracted by the anode and moves through an agarose gel depending on agarose concentration, size and conformation of the fragment and applied power. The DNA samples were applied to a 1 % agarose gel and analyzed electrophoretically. The gel was prepared by dissolving 1 % (w/v) agarose in 1× TAE buffer supplemented with 5 µl ethidium bromide solution (Sigma, USA) to visualize DNA fragments with UV-light. The sample was mixed with 6× DNA loading dye and applied to the gel. A suitable size marker was used to estimate the length of the DNA fragments. The electrophoresis run was performed at a constant voltage of 100 V. The DNA fragments from PCR and restriction digestion were exposed to UV light for detection, cut with a scalpel from the gel and purified using the peqGOLD gel extraction kit (PEQLAB Biotechnology GmbH).

## 3.2.5. Ligation:

Plasmid vectors and DNA fragments were ligated using a molecular ratio of 1:5 with the addition of 1 µl of T4 ligase and 2 µl of 10x ligation buffer in a total volume of 20 µl. The reaction mixtures were incubated overnight at 18 °C. Afterward, the ligation mixtures were directly incubated at 65 °C for 15 min to stop the reaction of the T4 ligase. The reaction mixtures were transformed into XL10-Gold or DH5α *E.coli* cells. The cells were plated out onto agar plates containing 100 mg/ml ampicillin and incubated overnight at 37 °C.

$$\frac{m_{Vector}[ng] \times l_{Insert}[kb]}{l_{Vector}[kb]} \times m_{\frac{Insert}{Vector}} = m_{Insert}[ng]$$

**I**$_{insert}$ = length of the insert DNA, **I**$_{vector}$ = length of the vector DNA,
**M**$_{Insert}$ = mass of insert DNA, **m**$_{verctor}$ = mass of vector DNA,
**mr**$_{Insert / vector}$ = molar ratio insert: vector

# 3 MATERIALS AND METHODS

### 3.2.6. Preparation of chemically competent cells with $CaCl_2$:

A single *E. coli* colony or a glycerol stock (200 μl) was used to inoculate 100 ml Luria Bertani (LB)-medium and incubated at 37 °C overnight. The overnight culture was diluted 1:50 in LB-medium (10 ml of overnight culture were added to 500 ml LB-medium) and grown at 37 °C to an optical density at 600 nm (OD600) of 0.6-0.8. Reaching this OD600, the solution was cooled on ice for 10 min and centrifuged at 4 °C and 4,000 rpm for 10 min. The supernatant was discarded and cells resuspended in a sterile solution containing 0.1 M $CaCl_2$ and incubated for 15 min on ice. This suspension was again centrifuged at 4 °C and 4,000 rpm for 10 min and the supernatant was discarded. The cells were resuspended in 5 ml cold 0.1 M $CaCl_2$/10 % glycerol containing buffer. Aliquots of 200 μl were shock-frozen in liquid nitrogen and stored at -80 °C.

### 3.2.7. Isolation and purification of plasmids:

The XL10-Gold and *E. coli* DH5α strains were used for the preparation of DNA plasmids. DNA plasmid amplification was performed on a preparative scale incubating a 5ml overnight *E. coli* culture at 37 °C. The amplified DNA was isolated from cells using the peqGOLD mini-prep kit1 (PEQLAB Biotechnology GmbH). After preparation of the samples, the nucleic acid concentrations were determined using a Nanodrop spectrophotometer at an adsorption of 260 nm. The purity of the isolated DNA was determined by calculation of the ration between absorption at 260 nm and 280 nm.

### 3.2.8. DNA-Sequencing:

To investigate the success of cloning or mutation insertion, plasmid DNA was sequenced by SeqLab by extended hotshot sequencing. Samples were prepared mixing 6 μl of DNA with 1 μl of sequencing forward or reverse primer.

### 3.2.9. Transformation of plasmids into *E.coli* cells:

An aliquot of chemically competent cells was thawed on ice and incubated with 100 ng of plasmid DNA for 10 min on ice. A water bath (GFL, Germany) was preheated to 42 °C and cell-DNA mixture was incubated for 1 min at 42 °C. Subsequently, the sample was cooled on ice for 1 min and 1 ml of sterile LB-medium was added. The suspension was then incubated for 30 min at 37 °C. After that, it was centrifuged at room temperature (RT) for 30 s and the supernatant was decanted. Cells were resuspended with residual LB-medium and plated in a sterile environment. Plates contained 1.5 % agar-agar and the

# 3 MATERIALS AND METHODS

respective antibiotic for selection. After incubation overnight at 37 °C, colonies were picked and further analysed. For subsequent DNA isolation, competent *E. coli* XL10-Gold or DH5α cells, for expression BL 21 (DE3) cells were used for transformation.

## 3.3. Biochemical methods:

### 3.3.1. Recombinant expression of proteins:

For recombinant Step-tagged protein expression, the plasmid encoding for the target protein under control of a tetracycline promoter was transformed into *E. coli* (DE3) expression cells. A single colony or 10 μl of a glycerol stock were used to inoculate an overnight culture containing a 50 μg/ml ampicillin and incubated at 37 °C and shacked at 180 rpm. This overnight culture was diluted 1:50 in medium with the same antibiotic concentration and incubated at 37 °C until the OD600 reached 0.5 and induced with anhydrotetracycline (AHT) (IBA, Germany) to a final concentration of 2 ng/ml. After 4 h of expression, the culture was centrifuged at 8,000 g for 10 min. Cells were resuspended in 30 ml buffer W in case of Strep-tag protein expression and stored at -20 °C. While for His-tagged protein expression, protein expression was induced by adding IPTG to a final concentration of 0.5 mM, followed by cultivation at 30 °C for 5hours. 1 ml of the cell suspension was collected for analysis of the protein expression level. The collected samples were analyzed by SDS-PAGE after centrifuged at 4 °C for 10 min at 12.000 rpm and discarding the supernatant. The cell pellet was resuspended in 50 μl of 2× sample buffer and boiled at 96 °C for 10min. The denatured samples were loaded on a 10 % SDS-PEGE to be examined. The cells of the large-scale expression culture were harvested after 5 hrs of incubation by centrifugation at 4 °C for 30 min at 12.000 rpm. The supernatant was discarded and the pellet was stored at -20 °C.

### 3.3.2. Expression of selenomethionine substituted proteins in methionine auxotrophic *E. coli*:

The expression of the selenomethionine-labeled protein (Ov Ral-2 contains 3 Met residues) was achieved applying a transformation of the construct into the methionine auxotrophic *E. coli* strain B834 (DE3) (Novagen). The SeMet protein was expressed in M9 minimal media containing 30 mg l$^{-1}$ L-SeMet and purified as described for the native protein, except that 2 mM TCEP were added to the buffer after the eluting step from the Ni-NTA affinity column to prevent SeMet oxidation.

# 3 MATERIALS AND METHODS

### 3.3.3. Tobacco etch virus (TEV) protease expression:

Glycerol stocks of TEV-protease mutant S219V overproducing BL21 (DE3)-RIL cells were obtained from **Mrs. Rana Hussein**. The TEV protease was encoded on pRK793 plasmid under *lac*-operon control. An overnight culture was grown at 37 °C in LB medium with an addition of 34 µg/ml chloramphenicol (Cam) (Sigma, USA) and 100 µg/ml ampicillin (Amp) (Roth, Germany).

The overnight culture was diluted 1:50 in the same medium as used for the overnight culture and grown until OD600 reached 0.6-0.7. For induction, the expression was induced by adding β-thiogalactopyranoside (IPTG) (Roth, Germany) to a final concentration of 1 mM, followed by cultivation at 30 °C for 4 hours. The cells were harvested after 4 hrs of incubation by centrifugation at 4 °C for 30min at 12.000 rpm. The supernatant was discarded and the pellet was stored at -20 °C.

### 3.3.4. Preparation of TEV protease glycerol stocks:

The respective plasmids were transformed into *E. coli* BL21 (DE3) expression cells and plated on LB-Amp agar and incubated overnight. A 100 ml sterile culture was inoculated with a single colony and incubated at 37 °C. The medium is centrifuged and 800 µl of sediment was mixed thoroughly with 200 µl of glycerol, frozen in liquid nitrogen and stored at -80 °C.

### 3.3.5. Cells disruption and solubility test:

A bacterial pellet was thawed and resuspended in 30 ml of lysis buffer at RT. After that, a spatula tip of lysozyme powder was added and the suspension incubated on ice for 10 min. The cells were mixed with the protease inhibitor phenylmethylsulfonylfluoride (PMSF) (Sigma, USA) to a final concentration of 100 µM and transferred to a glass beaker. Sonification was carried out with a Branson sonifier 250 (Emerson Electric Co, USA) for 10 min at 40 kHz output frequency followed by 10 sec incubation on ice in between sonication. To check whether the target protein is expressed in a soluble or an insoluble form as inclusion bodies, the lysate was centrifuged at 50,000 g at 4 °C for 50 min. and both of the supernatant and the pellet were analyzed on a 12 % SDS-PAGE gel and stored at 4 °C.

# 3 MATERIALS AND METHODS

### 3.3.6. Purification of the His-tagged proteins:

After cell disruption and centrifugation, the supernatant was incubated at 4 °C with previously equilibrated Ni-NTA agarose (Qiagen, Germany) with washing buffer for 30 min under rotation. For washing, the matrix was pelleted at 800 rpm, for 1 min, at 4 °C and incubated for 20 min with His-washing buffer at 4 °C. In a second washing step, the matrix was transferred to an empty column and protein was eluted with 5 ml His-elution buffer, reusing 3x the same buffer. In the end, 1-2 ml of fresh elution buffer depending on matrix volume was added to the matrix and protein dialyzed against 2 L of the respective imidazole-free buffer overnight at 4 °C.

### 3.3.7. Purification of TEV protease:

TEV protease was fused to a 6x His-tag and was sonified and purified. All subsequent purification steps were performed at 4 °C. The matrix was washed with 4 CVs of column buffer followed by an elution step with 4 CVs of the same buffer containing 250 mM Imidazole. The fractions were collected in pre-chilled (4 °C) Falcon tubes and analyzed by SDS-PAGE. The protein concentrations in the purified fractions were determined using a Nanodrop spectrophotometer at a wavelength of 280 nm. Due to TEV protease instability against oxidation, elution buffer was removed by dialyzing and exchanged by TEV protease storage buffer containing 1 mM dithiotreitol (DTT) (Biomol, Germany). TEV protease was stored at -20 °C.

### 3.3.8. TEV protease digestion and separation of TEV cleaved proteins:

TEV protease was used to remove His- tags from proteins when cleavage site was cloned with a His-tag. The cleavage was achieved by incubating proteins with TEV protease at a ratio of 1:50 and 1:10, from purified aliquots at -80 °C, to test cleavage. Digestion was carried out at room temperature for 4hrs and overnight at 4 °C. Samples were collected and analyzed by SDS-PAGE to determine the cleavage progress. Due to the instability of the proteins during cleavage, the cleaved proteins were dialyzed with a stabilizing buffer before TEV removal. After dialysis of the cleaved proteins, the mixture of TEV protease and proteins for His-tag removal, were incubated with previously equilibrated Ni-NTA resin on a roller rocker at 4 °C for 45min. The collected flow through contained the cleaved proteins, the TEV protease and cleaved His-tags where bound to Ni-NTA matrix which were eluted by buffer containing imidazole.

# 3 MATERIALS AND METHODS

### 3.3.9. Regeneration of Ni-NTA matrix:

The reuse of Ni-NTA resin depends on the nature of the recombinant proteins. Ni-NTA matrix was washed first with 2 CVs of regeneration buffer, and then 5 CVs of water and after that with 3 CVs of 2 % sodiumdodecylsulfate (SDS) solution. An ethanol (Roth, Germany) gradient from 25 % (1 CV per washing step) to 100 % (5 CVs) and back to 25 % was applied. To remove ethanol, the resin was washed with 1 CV of water. To remove the binded $Ni^{2+}$ ion, the matrix was incubated with 5 CVs of 100 mM ethylenediaminetetraacetic acid (EDTA), pH 8.0 (Biomol, Germany) containing buffer and subsequently washed with 2 CVs of water. The matrix was recharged with 2 CVs of 100 mM $NiSO_4$ (AppliChem, Germany) containing buffer for 10min, washed with 2 CVs of water, 2 CVs of regeneration buffer and 1 CV of water. For subsequent use, the matrix was equilibrated with protein buffer or stored in 20 % ethanol at 4 °C.

### 3.3.10. Purification of Strep-tagged proteins:

For purification, 1 ml of Strep-tactin resin (IBA, Germany) was applied to an empty column and equilibrated with 1 CV of buffer W. Cell lysate from sonified and centrifuged cell pellet was applied to the column and the matrix was washed with 2 CVs of buffer W. Protein was eluted 3x, reusing 2.5 mM desthiobiotin (IBA, Germany) containing buffer W in 2 ml and then again with 1 ml of fresh elution buffer. The protein was dialyzed against 2 L of the respective desthiobiotin-free buffer. Strep-tactin matrix was regenerated with 1 CV buffer W containing 1 mM hydroxyl-azophenyl-benzoic acid (HABA) (IBA, Germany) and washed with 3 CVs of buffer W. For subsequent use the resin was equilibrated with buffer W and stored at 4 °C in a 50 % suspension with 30 % ethanol.

### 3.3.11. Size-exclusion chromatography (SEC):

For the further purification of the proteins solution, an ÄKTA system (GE Healthcare, UK) was applied. The purification runs were carried out at 6 °C. The ÄKTA system was started and controlled using the software UNICORN 4.12 (GE Healthcare, UK). The protein concentration was determined at $A_{280}$. The peak fractions that showed a high $A_{280}$ absorption were pooled and analyzed by SDS PAGE for purity. Applied columns obtained from GE Healthcare are listed below.

• Hi-Load 16/60 Superdex 75

• Hi-Load 16/60 Superdex 200

# 3 MATERIALS AND METHODS

Lysozyme (hen egg, 14 kDa; SIGMA), thaumatin (*Thaumatococcus daniellii*; 22 kDa; SIGMA), ovalbumin (hen egg; 44 kDa), conalbumin (chicken egg white; 75 kDa), aldolase (rabbit muscle; 158 kDa), ferritin (440 kDa; horse spleen), thyroglobulin (bovine thyroid; 669 kDa) and blue dextran 2000 (all GE Healthcare) were used for calibration of SEC columns. Dailysis was performed against dialysis bufferes to change the buffer composition of the protein solution and get red of undesirable chemicals. An ultra filtration membranes (AMICON; 3 kDa/10 kDa/30 kDa/50 kDa MWCO) were used according to the company's instructions.

### 3.3.12. Protein quantification with Nanodrop 2000c:

The protein concentration was determined applying a Nanodrop 2000c instrument (Thermo Scientific, peqLab, Germany). 2 µl of protein buffer were applied to the sensor, elution buffer was used as blank. The concentration determination was repeated and the average was calculated. The protein concentration in solution was usually determined by measuring the specific absorbance of ultra violet light at a wavelength of 280 nm according to the law of Lambert and Beer $$A_{280} = \varepsilon * b * c$$

$A_{280}$ = **absorption at 280 nm**
**$\varepsilon$ = molar extinction coefficient [$M^{-1}$ $cm^{-1}$]**
**b = path length [cm]**
**c = protein concentration [mg/ml]**

The theoretical molar extinction coefficient $\varepsilon$ [$M^{-1}$ $cm^{-1}$] was calculated by ProtParam (111).

### 3.3.13. SDS-polyacrylamide gel electrophoresis (PAGE):

Electrophoresis is a widely used technique used in biochemistry, molecular biology, and biotechnology to separate biological macromolecules as proteins or nucleic acids, according to their electrophoretic mobility. The mobility is depending on the charge and molecular weight of the molecule. SDS-PAGE is a simple method used to determine size and purity of the protein sample under denaturing conditions. The separation of macromolecules occurred in an electric field, so it called electrophoresis. It considered as a very popular method for separating proteins by electrophoresis applying variant polyacrylamide gel concentration as a support medium and sodium dodecyl sulfate (SDS) as a denaturating agent of the proteins. SDS is an anionic detergent binds with peptide chains and maintains a net negative charge within a wide pH range. The negative charges on SDS damage most of the quaternary structure of proteins, and are passively attracted toward an anode (positively-charged electrode) under the effect of the electric field.

# 3 MATERIALS AND METHODS

Polyacrylamide gels act as molecular sieving and resist the larger molecules from migrating faster as the smaller particles. Protein samples were mixed with 2× concentrated sample buffer and incubated at 96 °C for 10min for denaturation. The gel was vertically placed in a gel chamber (Hoefer Inc, USA) and connected to an EV 231 power supply (Peqlab, Germany) to adjust the electric field. Electrophoresis was terminated as soon as bromophenol blue reached the low end of the gel. The gel was stained for at least 2hrs in Coomassie staining solution and subsequently destained in 20 % (v/v) acetic acid until a sufficient contrast was visualized. The gels were documented by scanning of the gel using a color scanner (Epson).

For 12.5 % polyacrylamide gel the following chemicals were pipetted:

| Components | 4 % stacking gel | 12.5 % separating |
|---|---|---|
| Stacking gel/separating gel buffer | 1 ml | 1.5 ml |
| Bi dest water | 2.59 ml | 1.9 ml |
| 40 % acrylamide | 400 μl | 1.9 ml |
| 10 % SDS | 40 μl | 60 μl |
| 10 % ammonium persulphate (APS) | 20 μl | 25 μl |

### 3.3.14. Native gel electrophoresis:

In the native gel electrophoresis, the protein keeps its folding state and net charge, therefore, the proteins are not separated according to their molecular mass, but by their isoelectric point and hydrodynamic radius. There are three native PAGE method, referred to as native clear, blue native and native quantitative preparative continuous-PAGE (QPNC-PAGE). Pre-cast polyacrylamide gels (4-16 % gradient gels) for blue or clear native PAGE were obtained from SERVA (Heidelberg) and used according to the instructions in the manual applying a mighty small II PAGE chamber (Hoefer, USA). The protein samples were diluted with sample buffer (2x without SDS) in a ratio of 1:1 mixed and applied to the native gel. Since in the electrophoresis much heat is generated, the electrophoresis was carried out on ice to avoid possible denaturation of the native proteins. The electrophoresis was started with a constant voltage of 50 V until the proteins had migrated through the stacking gel. For the migration of proteins through the separation gel, a voltage of 150 to 200 V was chosen. Furthermore, a standard protein marker to approximate the molecular weight, sample buffer as well as anode and cathode buffer was purchased from SERVA.

# 3 MATERIALS AND METHODS

### 3.3.15. Coomassie staining of Native and SDS-Gels:

During coomassie staining, the stacking gel was removed and the separating gel was incubated in coomassie staining solution for at least 2 h. To visualize only protein bands, the gel was incubated for 4 h in the destaining solution and stored in water. Results were documented by scanning the gel.

### 3.3.16. Western blot:

The Western blot was used for the immunological detection of the expressed proteins. Proteins were separated in a 12.5 % separation gel and blotted. After successful blotting, the nitrocellulose membrane was blocked with blocking buffer (0.5 % (w/v) BSA in TBS) overnight at 4 °C to prevent nonspecific antibody binding, then washed twice with TBS (750 mM NaCl, 1 M Tris-HCl, pH 7.5). The membrane was incubated with mouse anti-tetra-histidine IgG1 (QIAGEN) in blocking buffer for 2 h at room temperature. After washing with TBST [5× TBS buffer, 0.05 % (v/v) Tween 20] and TBS, the membrane was incubated with the secondary goat anti-mouse antibody linked with APC for 45 min. The membrane was washed again with TBS and TBST and transferred to reaction buffer supplemented with 50 g $l^{-1}$ NBT (dissolved in 70 % DMF) and 20 g $l^{-1}$ BCIP. After washing the blot color was developed. At a sufficient level of staining, the membrane was rinsed with deionized water to stop the reaction and finally documented by a color scanner.

## 3.4. Biophysical methods:

### 3.4.1. Dynamic Light Scattering (DLS):

Dynamic light scattering (known as photon correlation spectroscopy or quasi-elastic light scattering) is a tool that can be used to define the size distribution of small particles in solution. It can additionally be used to test the behavior of protein complexes in solutions. To investigate the dispersity and the numerical hydrodynamic radius (and its time-dependent change) of macromolecules in solution a spectroLIGHT 300/500 DLS instrument (Xtal concepts) was applied. This allows using either at least 15 μl in a quartz cuvette or at least 2 μl of sample in situ, e.g. in a terazaki 96 well plate overlaid with paraffin oil to reduce evaporation.

The sample was irradiated by a red light class 3b laser (λ = 690 nm; laser power 10-50 mW) to detect isotropic scattering at an angle of 90 degree. In preparation, samples with a concentration of 1-2 mg/ml were centrifuged at 4 °C for 30 min at 12.000 rpm to remove high molecular weight aggregates and analyzed in a quartz cuvette with a path length of

# 3 MATERIALS AND METHODS

1cm at room temperature. Measurements were accumulated per sample using an autopilot function.

The hydrodynamic radius of the particles is calculated using the Stokes-Einstein equation

$$R_H = \frac{K_B T}{6\pi\eta D_T}$$

The $R_H$ is the hydrodynamic radius that is determined by its dependence on viscosity $\eta$ and the diffusion coefficient $D_T$. $K_B$ is the Boltzmann's constant, and $T$ is the absolute temperature.

### 3.4.2 Circular dichroism (CD) spectroscopy:

CD spectroscopy is a spectroscopic technique used to determine the secondary structure contents of proteins and polypeptides in solution, as well as investigating the stability of proteins by measuring the melting temperature and folding changes after addition of additives or buffers exchange. The CD-spectroscopic experiments were performed on a Jasco J-815 instrument in 1 mm quartz cuvettes with a total volume of 100 μl of 3-10 uM protein solution. For far-UV experiments 15 measurements were applied. The wavelength range of 260-190 nm carried out with an interval of 0.1 nm, which were combined by arithmetic averaging. To reduce the background signal of the identical buffer, the buffer was measured by the same procedure and subtracted from the protein signal. The protein solutions were centrifuged at 4 °C for 30min at 12.000 rpm perior to measurment. To determine the thermal stability of the protein targets, the melting curve with a heating rate of 1 °C/min was applied, the changes in CD signal at 208 nm was detected. The *elasticity* $\theta$ is defined as the difference in absorbance of clockwise and counter clockwise circular polarized light

$$\theta = \frac{180 * 1n10}{4 * \pi} * (A_L - A_R)$$

From standard curves, secondary structure elements were determined. Typical curves for α-helical folding show minima at 208 and 220 nm and a maximum at 192 nm (see yellow curve figure 7), proteins with mainly β-sheet folding show a minimum at 215 nm and a maximum at 195 nm (see blue curve figure 7), whereas random coil structures lead to a minimum between 190 and 200 nm and a maximum at 220 nm (see red curve figure 7). All these values are not absolute values, but landmarks for CD spectra analyses. The result was

displayed by the software (Jasco, UK) as CD signal [mdeg] against wavelength and was automatically converted by the software to molar ellipticity [$\Theta*1000$ (deg cm$^2$ dmol$^{-1}$)] according to.
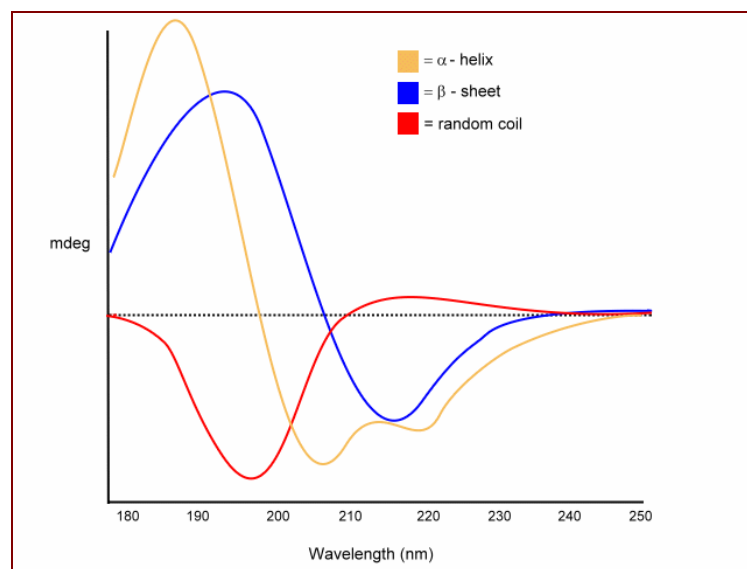


Fig.7: Standard curves for CD measurement. The CD signal for α-helices is depicted in yellow, β-sheets in blue and random coil in red. Figure maintained by (proteinchemist.com).

### 3.4.3. Mass spectrometry (MS)-based protein identification:

Digestion by trypsin and mass spectrometry data collection was performed in cooperation with the research group of Prof. H. Schlüter (UKE, Hamburg). For protein identification, protein bands were manually excised from an SDS-PAGE gel and digested with trypsin. The samples were loaded onto a zip tip by pipetting 20 times and washed three times with 10 µl 0.2 % formic acid. The tryptic fragments were eluted with 60 % acetonitrile. Tryptic peptides of the digested protein were spotted onto an Anchor Chip target (Bruker Daltonik GmbH, Bremen, Germany) with α-Cyano-4-hydroxycinnamic acid as matrix. The analysis was performed by MALDI-TOF/TOF mass analysis using an ultrafleXtreme mass spectrometer (Bruker Daltonik). An electrospray ionization (ESI) ion trap instrument (LC/MSD Trap XCT Ultra II) was used to detect peptides and results were analyzed using the Mascot search engine (Matrix Science).

### 3.5. Methods to analyze the 3D structure:

### 3.5.1. Sample preparation for protein crystallization:

# 3 MATERIALS AND METHODS

Crystallography is one of the most widely used technologies for the determination of a high-resolution 3D-structure of a protein. Crystals of the protein of interest were grown and subsequently analyzed using X-ray radiation. Before crystallization experiments, the protein concentration needs to be adjusted. Therefore, a pre-crystallization test (Hampton Research, USA) was performed to obtain the most suitable start conditions for the crystallization experiments. After protein purification, samples were concentrated and centrifuged in an ultracentrifuge at 50.000 g for 1 hour. The disparity was monitored by DLS measurements, and samples were concentrated into their final concentration. Different sets of crystallization solutions were carried out using the Zinsser Pipetting Robot (Genomic Solutions, UK) and the Oryx 8 (Douglas, UK) for the screening of various buffer conditions. When using the pipetting robot Honeybee 961 (Genomic Solutions, USA), protein and precipitant solutions were transferred to 96 well Nextal Qia1 plates (Qiagen, Germany). The commercially available screens JCSG+, Classic, Cryos, ComPAS, and AmSO4 Suite (all Qiagen, Germany), also Morpheus, Stura, and PACT (all Molecular Dimensions, UK) were used. 400 nl of protein solution were mixed with 400 nl of the respective precipitant solution in one well. The reservoir was filled with 55 μl of precipitant solution. The plates were sealed and stored at 16 °C.

## 3.5.2. Optimization of initial crystallization conditions:

After obtaining initial crystallization conditions, the conditions were optimized by varying the precipitant concentration around the condition found in the initial screen. In the presence of polyethylene glycols (PEGs) as precipitant, the concentration was increased and decreased by 10 % increments. Conditions were optimized in 24-well Limbro plates (Hampton Research, USA) using hanging drop vapour diffusion, consisting of 2 μl protein solution and 2 μl precipitant solution and the reservoir filled with 500 μl of precipitant solution. Conditions were also optimized in 48-well MRC sitting drop plates (Molecular Dimensions, UK). Droplets were made of 2 μl protein solution and 2 μl precipitant solution, the reservoir was filled with 50 μl of precipitant. Droplets were pipetted manually. The temperature was also varied, and plates were either stored at 4 °C or 16 °C in an air conditioned laboratory. Crystallization under oil and or seeding techniques has also been considered. For crystallization under oil, a Terazaki plate (Nunc, Denmark) was previously treated with paraffin oil (Applichem, Germany) to fill all wells with oil. Microseeding is a powerful tool to bypass spontaneous nucleation and to optimize the quality of the crystal. A seedstock was prepared with a crystal in a stabilizing solution in a

# 3 MATERIALS AND METHODS

tube containing a glass ball according to the supplier's protocol (Jena BioScience, Germany). The seed stock was either diluted 1:10-1:100 with protein solution containing the protein at which the crushed crystal was obtained with the same concentration, or was used to draw streak lines using with a horse-tail hair within droplets containing 2 μl precipitant and 2 μl the protein solution. Micro batch crystallization and Microseeding carried out applying the pipetting robot Oryx 4 (Douglas, UK).

### 3.5.3. Evaluation of the protein crystals:

To distinguish between protein and salt crystals, droplets containing observed crystals were monitored using the CrystalScore (Diversified Scientific Inc., USA) and SONICC (Benchtop, Formulatrix.Inc, USA), according to the manufacturer's specifications.

### 3.5.4. Radiation sources, Data processing and model building:

X-ray crystallography is a technique widely used in determination of the three dimensional structures of macromolecules. It is the most successful when applied on the third-generation synchrotron sources that allow rapid and efficient collection of X-ray data from macromolecular single crystals. At present there are over l00 synchrotron radiation facilities all over the world and most of them have an advanced dedicated line suitable for X-ray macromolecular diffraction experiments. In this study the following X-ray sources were applied.

**In-house Rigaku RU-200 rotating-anode X-ray generator**

      Operated at 50 kV and 100 mA

      Equipped with a MAR Research image-plate detector.

**P11 and P13 [Beamlines at Petra III] DESY (Hamburg)**

      Synchrotron Source: PETRA III storage ring

      Focal Spot: 10x10 $\mu m^2$

      Detector: PILATUS 6M

The crystal is illuminated with a finely focused monochromatic beam of X-rays, producing a diffraction pattern of regularly spaced spots known as reflections. The two-dimensional images taken at different orientations are converted into a three-dimensional model of the density of electrons within the crystal using the mathematical method of Fourier

# 3 MATERIALS AND METHODS

transforms. To analyze single crystal diffraction data, several computer programs have been applied. The data were processed with iMOSFLM (87), XDS (88) or HKL2000/ Denzo (89). For scaling, either SCALA (90) or SCALEPACK (89) was used. Solving of phase problem was done applying single wavelength anomalous dispersion/diffraction (SAD) method (91). For further data analysis and refinement programs from the CCP4 (92) or PHENIX (93) suite were applied. Model building was carried out with COOT (94).

### 3.5.5. Small-angle X-ray scattering (SAXS):

Biological SAXS becomes a popular technique to rapidly characterize overall structure and shape and molecular weight of proteins, nucleic acids and macromolecular complexes in solution (95, 96). The elastic scattering of X-rays (wavelength 0.1-0.2 nm) by the sample, is recorded at very low angles (typically 0.1 - 10°) and analyzed. To obtain good SAXS results; the soluble aggregates, contaminants, partially dissociating complexes or oligomers should be removed. The 3 D structural analysis require homogenous and monodisperse solutions (97). Linear Guinier plots indicated monodispersity. The sample is filtered and checked by DLS prior to analysis. The low angle data were extrapolated to infinite dilution to yield the final scattering curves. The low angle data were extrapolated to give the final scattering curves. The pair-distribution function was calculated by GNOM (98). Ab initio models were calculated and superimposed using DAMMIN/DAMMIF (99) and the automated mode of DAMAVER (100). CRYSOL (101) allows comparison/superposition of the scattering data to known high-resolution structures. CRYSOL is aware of the hydration shell and calculates a spherically averaged scattering pattern using multipole expansion of the scattering amplitudes. Varying the background, hydration shell contrast and the relative amount of displaced solvent, the $\chi$ -value is minimized. This value comprises a quality measure of the deviation of the fitted function and the scattering data. SASREF/GASBOR (102, 103) and PRIMUS (104) were applied both for rigid-body modeling. Bovine serum albumin (66 kDa; in 50 mM HEPES, pH 7.5) and Lysozyme served as a molecular weight reference.
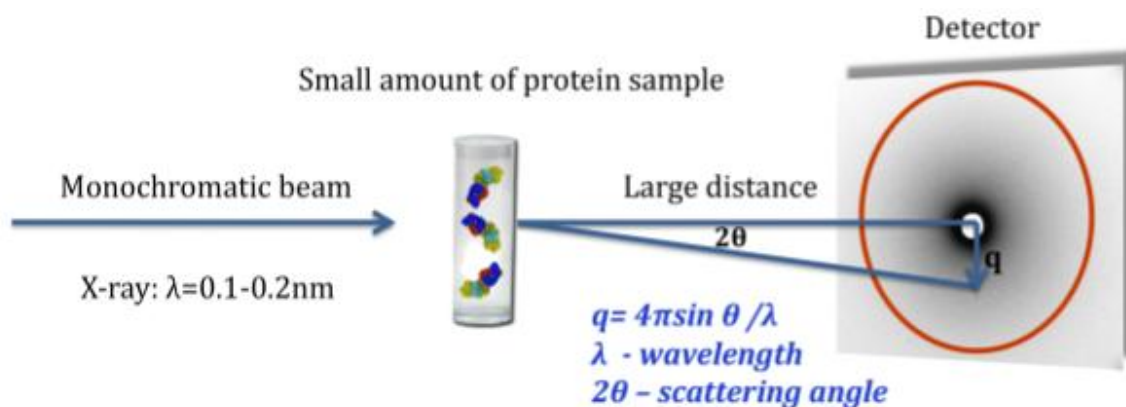
# 3 MATERIALS AND METHODS



Fig.8: Schematic diagram of a SAXS experiment (Resource for Biocomputing, Visualization, and Informatics-www.rbvi.ucsf.edu).

## 3.6. Bioinformatic tools and software packages:

### 3.6.1. Docking and prediction of binding sites:

The SITEHOUND algorithm was used to identify the potential ligand binding sites by recognizing regions possess favorable non-bonded interactions with a different chemical probe (105). SITEHOUND-web (http://sitehound.sanchezlab.org) was supplemented with a python-CGI and JavaScript based platform. The integrate program MODELLER (106) and AutoGrid (107) have beeb applied to identify new binding site. Binding site related to active site were excluded. For docking analyses AutoDock 4.2.3 (108) was used. AutoDock is a suite of automated docking software used to predict how small ligands, such as substrates or drug candidates, bind to an epitope in the known 3D structure. The Lamarckian genetic algorithm (LGA) was applied in the docking study. The active site was defined and discarded. The grid size was set to $60 \times 60 \times 60$ points with a grid spacing of 0.675 Å centered on the original ligand in the crystal structure complex. The grid box includes the entire surface of the enzyme and provides enough space for the ligand translational and rotational walk. Step sizes of 1 Å for translation and $60^{\circ}$ for rotation were chosen and the maximum number of energy evaluations was set to 150,000 and 150 runs were performed.

### 3.6.2. Sequence and Structural Analysis:

Various programs were used for sequence and structure analyzes are briefly described below:

**3.6.2.1. I-Tasser homology modeling server:** Was used to calculate/obtain homology models (http://zhanglab.ccmb.med.umich.edu/I-TASSER) (109). The FTmap server

# 3 MATERIALS AND METHODS

provided grid-based small-molecule docking (protein "hot spot" mapping) applying a fast Fourier transform correlation algorithm (110).

**3.6.2.2. ProtParam:** Is a tool which allows computation of various physical and chemical parameters of proteins such as molecular weight, amino acid composition, atomic composition, estimated half-life, molar extinction coefficient, aliphatic index, instability index, and grand average of hydropathicity from the given protein sequence (111).

**3.6.2.3. Clustal Omega:** Is a multiple sequence alignment program used to generate alignments between three or more protein or nucleotide sequences (112, 113 and 114)

**3.6.2.4. BLAST (Basic Local Alignment search Tool):** Is an algorithm used for comparing biological sequence information in proteins and nucleic acids (115). There are different tools available on the BLAST server for comparing protein or nucleotide sequences against protein or nucleotide databases (blastp and blastn).

**3.6.2.5. Easy Sequencing in PostScript (ESPRIPI):** Allows easy visualization of aligned sequences via a PostScriptk/PDF output (116, 117). It can read secondary structure files to produce an illustration having both sequence structural and information and offers several options to highlight important regions in the alignment.

**3.6.2.6. PDBsum:** Is a web-based database providing a graphical summary of the key information on the macromolecular structure and includes images of the structure, annotated plots of each protein chain's secondary structure and detailed structural analysis (118).

**3.6.2.7. DALI:** Is a network service for comparing three-dimensional protein structures (119). Dali compares the coordinates of an input query structure against those in the Protein Data Bank (PDB). The output file contains PDB IDs of structures arranged according to Z-score as well as their RMSD with the input structure.

**3.6.2.8. TM-align:** Is an algorithm for carrying out sequence independent protein structure comparisons (120). The TM-score has values between 0 and 1, where 1 indicates a perfect match between two structures. Based on statistics of structures in the PDF, scores below 0.2 corresponds to randomly chosen unrelated proteins whereas a score higher than 0.5 indicates that the proteins have the same fold in structural classification of proteins (SCOP) (121) or class, architecture, topology or fold and homologous family (CATH) (122) databases.

**3.6.2.9. Multiple structural AligNment AlGorithm (MUSTANG):** Is used for the alignment of multiple protein structures (123). The program reads a set of input protein coordinates and constructs a multiple alignment using the spatial information of the Ca

atoms in the set based on a progressive heuristic algorithm. It further gains accuracy through novel and effective refinement phases.

**3.6.2.10. BAVERAGE:** Is a simple program in CCP4 suite (92) to read a PDB file, tabulate the average B values residue by residue (main chain and side chain separately) and the rms deviation of the B values from the mean. It also determines B values of individual chains (comprising of protein or water or ligand) in the PDB file.

**3.6.2.11. BDBePISA:** Is an interactive tool for the exploration of macromolecular molecules and database searches of structurally similar interfaces and assemblies, as well as searches on various assembly and PDB entry parameters (124). The program calculates the binding strength of all contacts between macromolecules and also indicates the most likely biologically relevant assembly.

**3.6.2.12. PDB2PQR server:** Used to add charge and radius parameters in place of occupancy and B factor columns in a PDB file such that it can still be used to visualize the structure with standard molecular graphics programs (125). The modified file can be used as an input for APBS in PyMol.

**3.6.2.13. STRIDE (Structural identification):** An algorithm for the assignment of secondary structural elements given the atomic coordinates of the protein as determined by X-ray NMR or any other method (126, 127).

**3.6.2.14. ConSurf Database:** A tool for estimating the evolutionary conservation of amino/nucleic acid positions in a protein/DNA/RNA molecule. The Rate4Site algorithm implemented in the ConSurf web server (128, 129) is used to calculate the conservation of amino acids at each position of the alignment using an empirical Bayesian inference.

**3.6.2.15. Protein (Structure) Comparison, Knowledge, Similarity and Information (ProCKSI):** Is a system for extensive comparison of multiple protein structures (130).

**3.6.2.16. APBS (Adaptive Poisson-Boltzmann Solver):** A freely available macromolecular electrostatics calculation program (131). It is now available as an APBS plugin written by Michael Lerner, which can be used with PyMol. The results of the calculations can be displayed as an electrostatic potential molecular surface in PyMol.

# 4 RESULTS

## 4. RESULTS:

### 4.1. Extracellular Superoxide Dismutase from *O.Volvulus* (OvEC-SOD):

4.1.1. Cloning, protein expression and purification of OvEC-SOD:

The OvEC-SOD open reading frame starting from amino acid Gly 43 (amino acids 1-42 are the signal peptide) was amplified by PCR from *O. volvulus* genomic DNA using sequence-specific antisense oligonucleotides primers. The BamH1 and EcoR1 were used as restriction-enzyme cleavage sites. The generated PCR product was cloned into the *E. coli* expression vector pASK-IBA16 (IBA, Göttingen, Germany). The N-terminus of the native protein is supplemented with a TEV protease cleavage site and an N-terminal Strep-tag II (ENLYFQG).
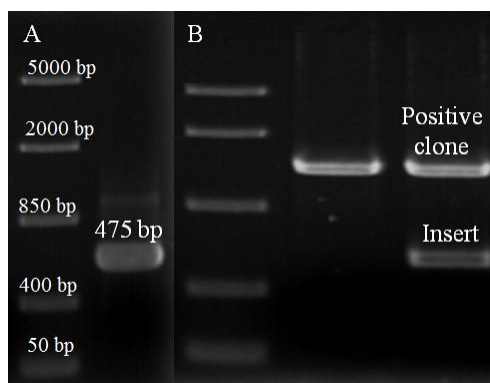


Fig.9: A) Agarose gel electrophoresis of the PCR product of the OvEC-SOD gene, the figure shows the right size band of the insert. B) The digestion products of the negative (right) and the positive (left) clone after digestion with BamH1 and EcoR1, the positive clone shows two bands of empty vector and the right size insert, while the negative clone shows only one band.

The recombinant plasmid was verified by DNA sequencing, as shown in figure 10. The competent cells (*E.Coli* BL21 DE3) were transformed with the recombinant vector.



Fig. 10: The nucleotides sequence of the OvEC-SOD, the yellow sequence is the Strep-tag, the blue sequence is the TEV protease binding site, the red sequence referes to the EcoR1 restriction site, and the green color is the BamH1 restriction site.

The protein was expressed and purified by affinity chromatography using the Strep-tag as described in the chapter 3.3.10. Maximum expression level was obtained after 4 h. The level of expression of the target protein was optimized by varying the time and/or

# 4 RESULTS

temperature of induction and the concentration of the inducer to obtain a maximum yield. The expected protein band occurred at the level of 19 kDa. The band size was predicted by the program ProtParam (111) on the ExPASy webpage, see figure 11 left side. Samples from sequential steps in the purification were resolved by 12 % SDS-PAGE and stained with Coomassie blue as shown in figure 11 on the right side.
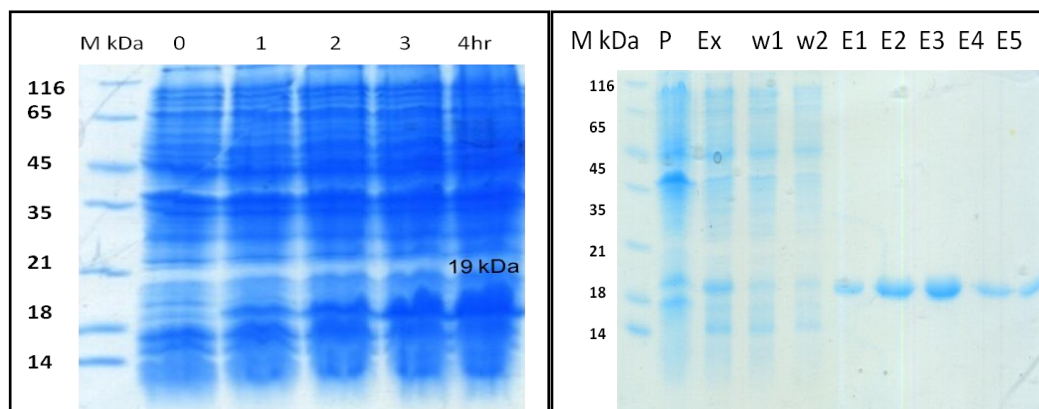


Fig.11: Left) The expression profile of the OvEC-SOD with an N-terminal Strep-tag inside the bacterial strain *E. Coli* BL21 DE3. Right) Purification of N-terminal tagged OvEC-SOD, P= pellet, Ex= supernatant, W $_{1-2}$ = washing fractions, E$_{1-5}$= elution fractions.

Complete cleavage of the tag was achieved by an enzyme concentration of a 1:10 molar ratio. The TEV protease is active only in reduced environment; this was achieved by addition of a weak reducing glutathione buffer to avoid the breakdown of disulfide bonds by DDT. The tag and the TEV protease enzyme were removed from the protein solution by affinity chromatography, leaving the untagged protein in the flow through. The final yield of the protein was ~20 mg per 4 liters of *E. coli* culture and contained more than 95 % purity, as shown in figure 12.
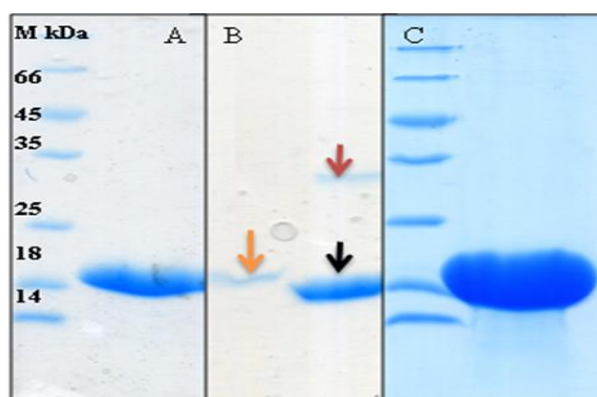


Fig.12: A) SDS-PAGE of N-terminal tagged OvEC-SOD. B) The process of the tag removal; the orange arrow refers to the tagged protein, the red arrow refers to the TEV protease, the black arrow refers to the untagged protein. C) The untagged pure protein.

# 4 RESULTS

## 4.1.2. Biochemical characterization of OvEC-SOD:

The molecular weight and oligomeric state of OvEC-SOD were estimated by gel filtration chromatography. A calibration curve was prepared using the standards carbonic anhydrase (29 kDa), albumin (66 kDa), alcohol dehydrogenase (150 kDa), amylase (200 kDa), apoferritin (443 kDa) and thyroglobulin (669 kDa). Subsequently, the relative molecular weight of the eluted protein was interpolated from a linear calibration plot of elution volume versus log (molecular weight). The calculated molecular weight indicates that the OvEC-SOD forms a stable dimer is solution, see figure 13.
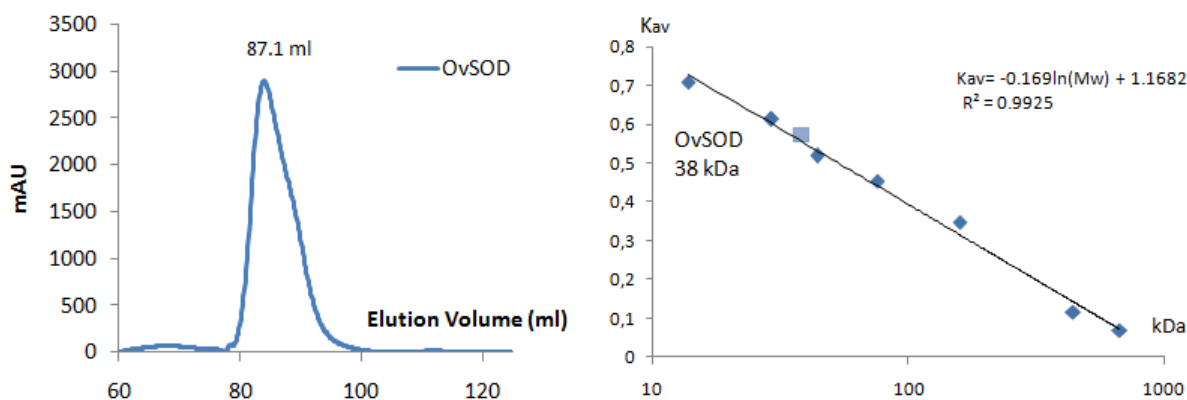


Fig.13: Superdex 200 SEC profile of OvEC-SOD, on the right side the calibration plot showing the calculated molecular weight.

The sequence integrity of the purified protein was confirmed as OvEC-SOD by mass-spectrometry. The mass spectra of the identified peptides are shown in figure 14.
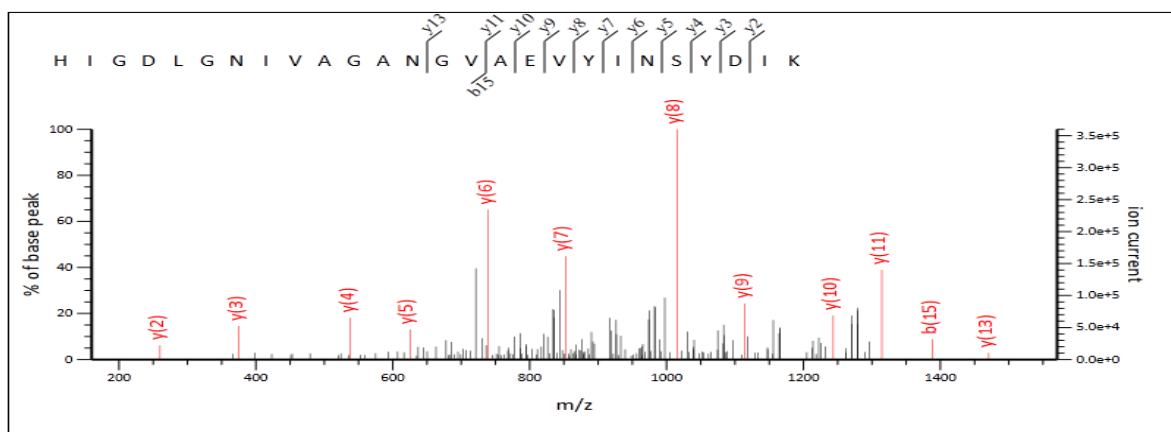


Fig.14: MS/MS sequence analysis of the identified peptide from OvEC-SOD.

CD spectroscopy verified the overall folding of the protein applying Yang`s algorithm (109). The folding has an overall approximation of the composition of helices, sheets, turns and random coils with 2.9 %, 57.5 %, 14.2 % and 24.4 %, respectively (Figure 15).
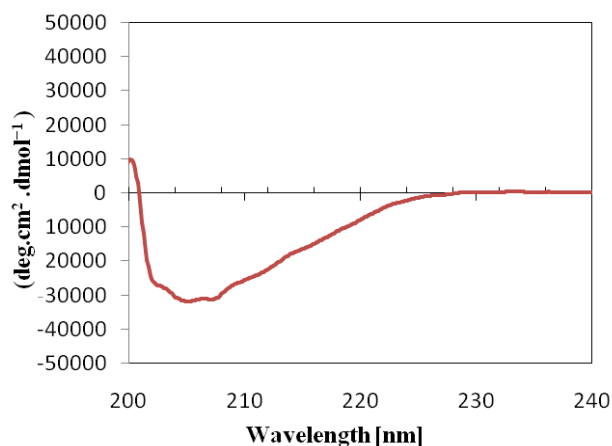
# 4 RESULTS



Fig.15: The adjusted CD spectrum of the native OvEC-SOD with a negative band at approximately 208 nm and a positive one at 200 nm. The protein was well folded and non aggregating.

The homogeneity of the purified protein was checked applying the dynamic light scattering (DLS). Different buffers and additives have been used to optimize the conditions and to obtain the most homogeneous and monodisperse protein solution as an optimal starting point for initial crystallization experiments. Centrifugation at 70.000 g and filtration of the solution were helpful to remove all larger aggregates and impurities. Measurements were performed over a time interval up to 15 days to check the protein stability. A final hydrodynamic radius of approximately 3.2 nm was observed, as shown in figure 16.
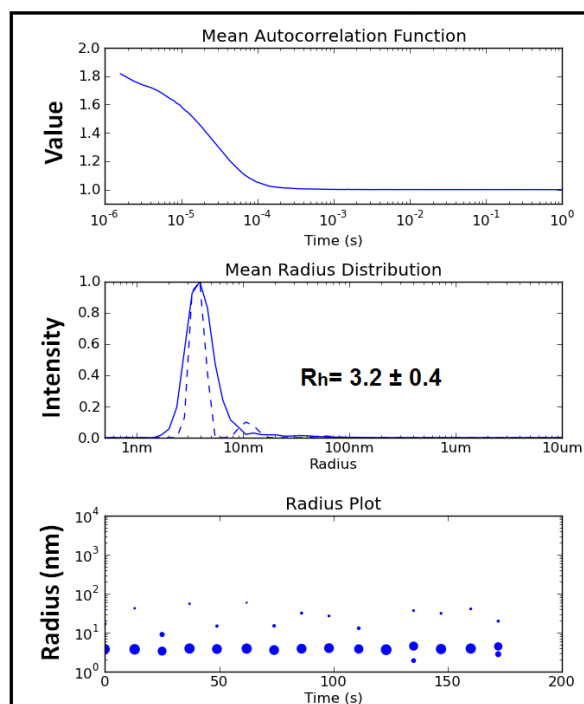


Fig.16: The DLS measurement of the OvEC-SOD in solution. The particle size was plotted against the time of the measurement. The radius distribution of protein particles is 3.2 nm, indicating a monodisperse and a non-aggregated protein in the solution.

# 4 RESULTS

### 4.1.3. OvEC-SOD crystallization in presence of an N-terminal Strep-tag:

The concentrated protein was initially screened against a diversity of available screening conditions. A 450 nl droplet of 15 mg ml$^{-1}$ protein dissolved in a crystallization buffer (described in chapter 3.1.6) was mixed with the same volume of precipitant solution and equilibrated against 45 μl reservoir solution. Brick like crystals appeared after one week in condition C6 of the ComPAS suite. Unfortunately these crystals diffract only up to 6 Å resolution, which is unsuitable for structure determination and may be explained by a presence of an internal disorder in the crystals, these crystals shown in figure 17.



Fig. 17: Low diffracting crystals appeared after one week in condition C6 of the ComPAS suite, consisting of 0.01 calcium chloride, 0.1 sodium acetate pH 4 and 60 % (w/v) MPD in the reservoir. The crystals dimensions varied from 0.18 mm to 0.24 mm.

### 4.1.4. Crystal quality optimization by removal of the affinity tag:

To optimize and improve the crystal size and quality, the Strep-tag free OvEC-SOD was used. The same concentration and buffer conditions were applied as before. After five months, varied shaped crystals were grown in three different crystallization conditions, listed in table 1. Crystals allowing high resolution diffraction were used for data collection as shown in figure 18.

Table 1: Promising screening conditions of the OvEC-SOD.

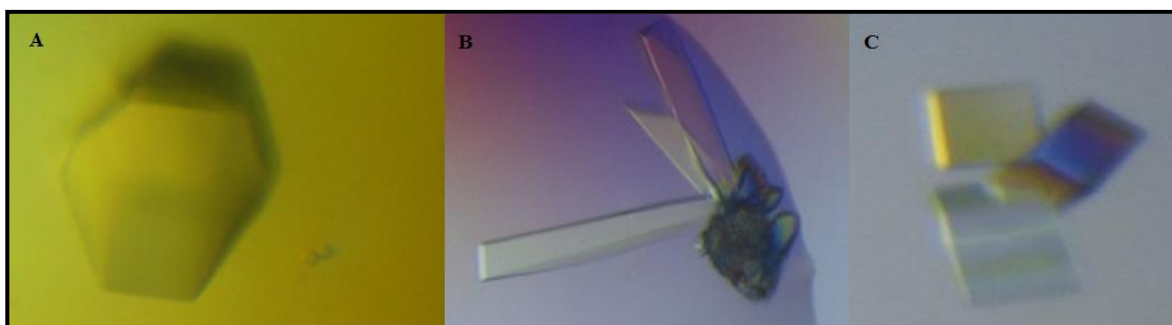| Fig.19 | Name | Condition constituents | Dimensions |
|---|---|---|---|
| A | Stura-C12 | 0.1 M MES pH 6.5, 45 % w/v PEG 2000MME | 0.25×0.36 mm |
| B | PACT-11 | 0.2 M calcium chloride, 0.1 M MES pH 6, 20 % (w/v) PEG 6000 | 0.1×0.6 mm |
| C | Morpheus-E5 | 10 % w/v PEG 20 000, 20 % v/v PEG MME 550, 0.03 M of each ethylene glycol and 0.1 M MOPS/HEPES-Na pH 7.5 | 0.25×0.25 mm |



Fig.18: The OvEC-SOD crystals grew within five months after applying three different screen conditions.

# 4 RESULTS

### 4.1.5. Data collection, processing and refinement:

Diffraction data up to 2.3 Å resolution were collected at room temperature using an in-house Rigaku RU-200 rotating anode X-ray generator. These crystals contained one monomeric molecule in the asymmetric unit and belonged to the space group $P3_121$, with the unit cell parameters a = 58.9, b = 58.9, c = 78.7 Å and 54 % solvent content. The crystal structure was solved by molecular replacement using the crystal structure of Superoxide dismutase from *C. elegans* (PDB entry 3kbe; sequence identity 60 %) as a search model. Another data set was collected up to 1.5 Å resolution at the PETRA III beamline P13 (EMBL Hamburg) at DESY (Hamburg, Germany). The crystal was cryo-protected by using the reservoir solution supplemented by 10 % glycerol (v/v %). Based on the model created from the first data set, data collection and refinement statistics shown in table 2 were calculated. Subsequent refinement and model building of strands and surface loops, adjustment of side chains and torsion angles was performed resulting in an $R_{factor}$ of 15.9 % and an $R_{free}$ of 18.1 % using all data in the range of 25.8–1.5 Å. RSM deviations for bond lengths and angles are 0.009 Å and 1.24 Å, respectively. The electron density maps are contiguous from residue 1 to residue 156, except the region extended from Gly 23 to Ser 27, which has a poor electron density; this part is the connection between β 2 and β 3 and highly accessible and flexible. The Ramachandran plot (Figure 19) indicates that 98.3 % of the main-chain dihedral angles are located in the most favourable regions, and 1.7 % was found in additional allowed regions.
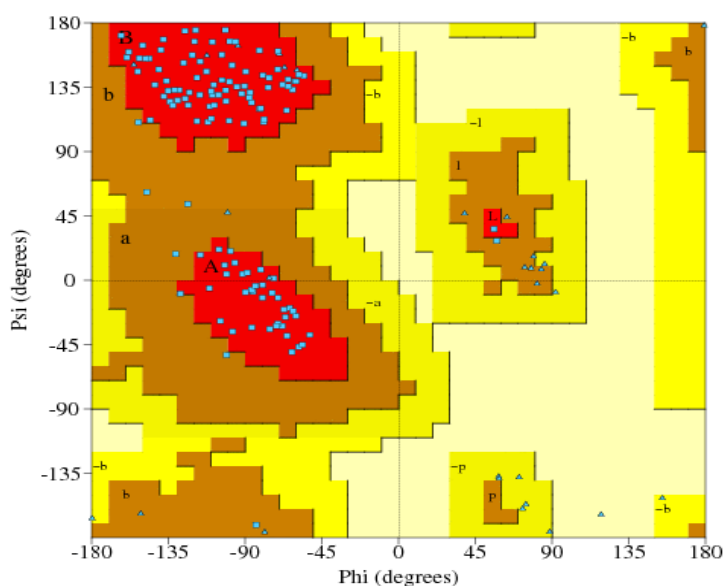


Fig.19: A Ramachandran plot generated from OvEC-SOD coordinates. The red, brown and yellow regions represent the favored, allowed, and "generously allowed" regions as defined by ProCheck program.

# 4 RESULTS

**Table 2: Data collection and refinement statistics of OvEC-SOD structure (PDB code: 5IN2):**

| Data collection parameters | Data 1 | Data 2 |
|---|---|---|
| Source | PETRA III P13 | In-Hause rotating anode |
| Wave length ($\lambda$ Å) | 1.033 | 1.542 |
| Temperature | 100 K | 291 K |
| Space group | P3$_1$21 | P3$_1$21 |
| Cell dimensions | | |
| $a$, $b$, $c$ (Å) | 58.4, 58.4, 77.6 | 58.9, 58.9, 78.7 |
| $\alpha$, $\beta$, $\gamma$ (°) | 90, 90, 120 | 90, 90, 120 |
| Resolution (Å) | 25.8 – 1.55 ( 1.6 -1.55) | 31.0 – 2.3 (2.5-2.3) |
| $R_{merge}$ (%) | 7.5 (42.3) | 8.4 (30.5) |
| <I / s(I)> | 14.1 | 28.9 |
| Completeness (%) | 100 | 98.5 |
| Redundancy | 8.7 | 18.6 |
| Average mosaicity | 0.22 | 0.75 |
| **Refinement** | | |
| No. of unique reflections | 22806 | 6210 |
| Rwork / Rfree (%) | 15.9 /18.1 | 19/24 |
| Average B value (Å$^2$) | 22.8 | 19.1 |
| **R.m.s. deviations** | | |
| Bond lengths (Å) | 0.009 | 0.038 |
| Bond angles (°) | 1.240 | 0.860 |
| **Ramachandran plot (%)** | | |
| Most favoured | 98.3 | 98.7 |
| Additionally allowed | 1.7 | 1.3 |
| **Nº of molecules** | | |
| Water | 92 | 57 |
| Zinc | 1 | 1 |
| Copper | 1 | 1 |
| Sulphate | 1 | 1 |
| Chloride | 2 | 0 |
| PEG | 1 | 0 |

†$R_{merge} = \Sigma_{hkl} \Sigma_i \mid I_i(hkl) – (I(hkl)) \mid /\Sigma_{hkl} \Sigma_i I_i$ (hkl), where $(I(hkl))$ is the mean intensity of the observations $I_i(hkl)$ of reflection *hkl*. Values in parentheses are for the highest resolution shell. Skew - Deviation from a Gaussian distribution; CRR$_{RMS}$ - correlation of a local r.m.s density. †† $R_{work} = \Sigma \mid\mid F\ o\mid$ - $\mid F\ c\mid\mid/\Sigma \mid F\ o\mid$ where F o and F c are the observed and calculated structure factor amplitudes, respectively.††† $R_{free}$ is the R-factor calculated with 5 % of the reflections chosen at random and omitted from refinement. R.m.s is a root-mean-square deviation of bond lengths or bond angles from ideal geometry.

## 4.1.6. Sequence alignment and analysis:

On the basis of the sequence alignment (Figure 20), it is clear that the OvEC-SOD shows considerable sequence homology to the human intracellular SOD (acc. no.: P00441, 56 %), rather than to the extracellular SOD (acc. no.: P08294, 45 %), as shown in Figure 15. The C-terminal region is mostly conserved between the OvEC-SOD and the human intracellular SOD. This may be a clue that OvEC-SOD has presumably evolved from the cytoplasmatic OvSOD1. The amino acids of the active site are conserved and marked with red color, see figure 20. Three functionally important residues around the active site are also conserved between the OvEC-SOD and human SODs (Arg146, Gly141, and Asp124).

```
SODEC_ONCVO    -42 MINSFIVIFL-------------------SFLIFINYANLVCVEATHVYGRRSHSNGMHGN
SODEC_HUMAN        MLALLCSCLLLAAGASDAWTGEDSAEPNSDSAEWIRDMYAKVTEIWQEVMQRRDDDGALHAA
SODEN _HUMAN       ------------------------------------------------------------
SOD_T.Solium       ------------------------------------------------------------


SODEC_ONCVO     +1 GARRAVAVLRGDAG-VSGIIYFQQGSGGSITTISGSVSGL----TPGLHGFHVHQYGDQTNG
SODEC_HUMAN        CQVQPSATLDAAQPRVTGVVLFRQLAPRAKLDAFFALEGFPTEPNSSSRAIHVHQFGDLSQG
SODEN _HUMAN       MATKAVCVLKGDGP-VQGIINFEQKESNGPVKVWGSIKGL----TEGLHGFHVHEFGDNTAG
SOD_T.Solium      --MKAVCVMRGEEG-VKGVVHFTQ--AGDAVKVHAEFEGL----KPGKHGFHVHEFGDTTQG
                   . ..: .     * *:: * *            ..*:     . . . ..:***::**   *


SODEC_ONCVO        CTSAGDHYNPFGKTHGGPNDRIKHIGDLGNIVAGANGVAEVYINSYDIKLRGPLSVIGHSLV
SODEC_HUMAN        CESTGPHYNPLAVPH------PQHPGDFGNFAV-RDGSLWRYRAGLAASLAGPHSIVGRAVV
SODEN _HUMAN       CTSAGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGTRLV
SOD_T.Solium      CTSAGAHFNPHGKNHGAPDAAERHVGDLGNVTAGADGKATLDLTDKMISLTGEHSVIGRSLV
                   * *    * **    *.        * **.**                  *     *    *


SODEC_ONCVO        VHANTDDLGQGTGNMREESLKTGNAGSRLACGVIGIAAVS--------------------
SODEC_HUMAN        VHAGEDDLGRG-GN--QASVENGNAGRRLACCVVGVCGPGLWERQAREHSERKKRRRESECK
SODEN _HUMAN       VHEKADDLGKG-GN--EESTKTGNAGSRLACGVIGIA-----------------------
SOD_T.Solium      IHVDPDDLGLG-GH--ELSLITGNAGGRVACGIIGIAKSE-------------------
                   :*    **** * *:   : *  .**** *:** ::*:.
```

Fig. 20: Sequence alignment for OvEC-SOD with the other homologue structures from *H.sapiens* and *T.solium*. The blue color refers to the signal peptide of the OvEC-SOD.

## 4.1.7. Overall crystal structure:

The structure comprises one monomeric molecule in the asymmetric unit; the biological unit is a dimer with two independent active sites (Figure 23). Every monomer contains 3 α-helices accounting for approximately 8.9 % of the secondary structure and 9 β-sheets represent 36.5 % of the secondary structure (as shown in figure 21).
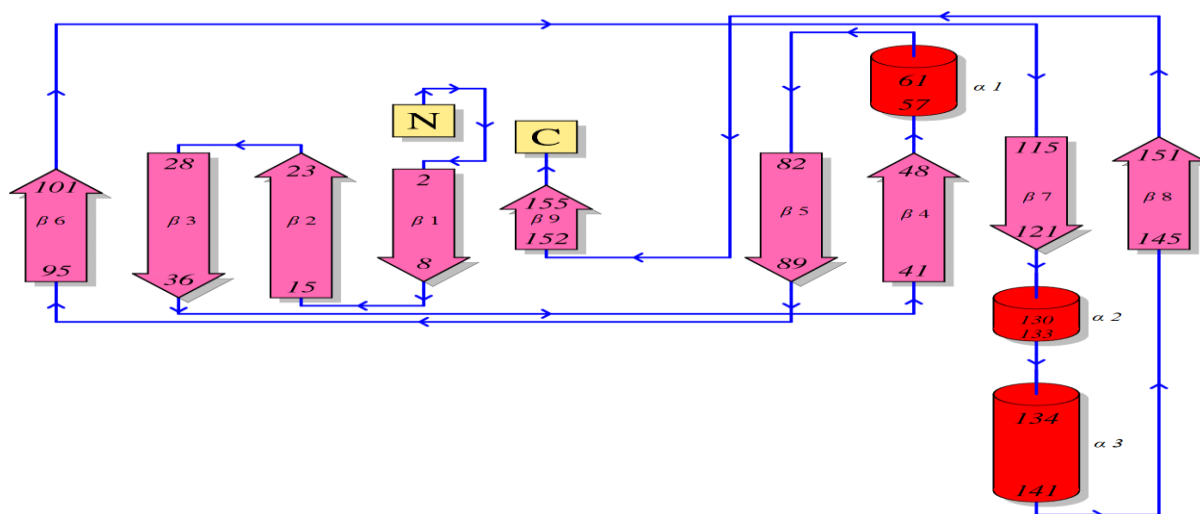


Fig.21: Topology diagram of the OvEC-SOD crystal structure. The α-helices are represented by red cylinders, the β-sheets as pink arrows, connecting loops in dark blue. The figure was prepared with PDBsum (118).

# 4 RESULTS

The OvEC-SOD structure shows the expected features conserved throughout the organisms, including the Greek-key β-barrel motif (Figure 22). Of particular interest is the comparison with the dimeric intracellular and the tetrameric extracellular human Zn/Cu SOD. Defined differences in the structures may allow the design of inhibitors of Zn/Cu SOD selectively without affecting the human Zn/Cu SOD. The common feature is a large cylindrical barrel made up of nine extended sheets of an entirely anti-parallel structure; it contains about 57 residues (36.5 % of the backbone). The barrel is somehow flattened in cross section and consists of two parts; one part toward the outside of the subunit is very regular. The other part is less regular, toward to the intern of the subunit; those sides subsidize four ligands to the metals. The rest of the subunit structure is combined of two loops of a non-repetitive type. The first loop (residues 47-84) has two distinct parts. The first part called the disulfide loop; this loop is connected to the barrel by the disulfide bridge. The second part is quite hydrophilic and contributes to the $Zn^{2+}$ ion with three ligands. The second long loop (residues 122-146) is very hydrophilic and called electrostatic loop. This loop drives the superoxide onto the active side.
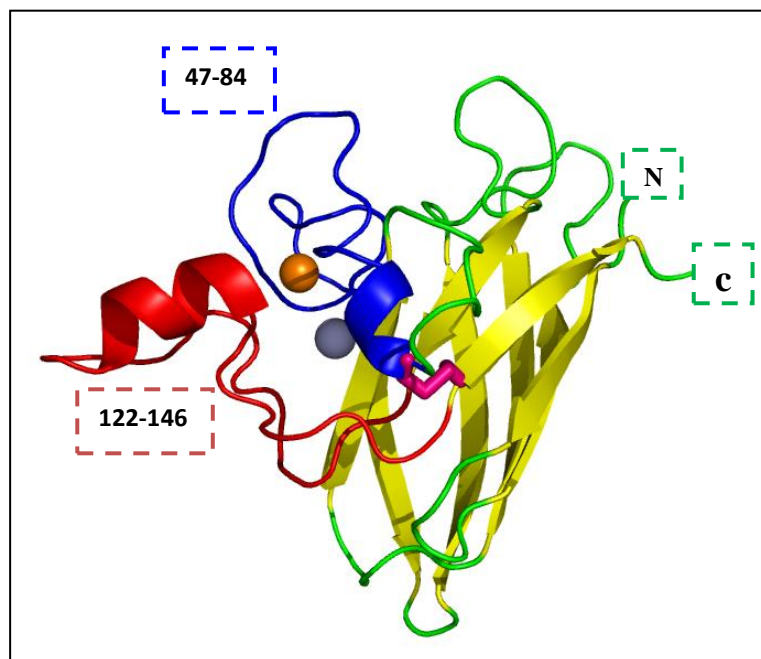


Fig.22: Cartoon representation of the monomeric OvEC-SOD structure. The red color is the electrostatic loop (122-146), the blue color is the first long loop (47-84) and the purple color is the disulfide bond, while the yellow color refers to the Greek-key β-barrel motif. The orange sphere is the $Cu^{+2}$ atom and the blue sphere is the $Zn^{2+}$ ion.

The OvEC-SOD superoxide dismutase possesses a high glycine content (30 out of the 156 residues). Fourteen of these glycines are at sharp corners and the tight turns, at these positions glycine is essential. The single disulfide bridge consists of Cys57 (α1) to Cys146

(β8). This disulfide bond may stabilize this longest loop. The bond length between the sulfur atoms of each residue is 3.03 Å, which is the ideal value. Aspartate residues in proteins are generally solvent accessible with ionized carboxyl groups. In the OvEC-SOD, two of the eight aspartate residues are completely solvent inaccessible. They are Asp83 and Asp124. Asp83 is one ligand of the $Zn^{2+}$ ion and is almost buried in the molecule. Asp124 makes a H-bond formation with the main chain nitrogen atom of Leu126.
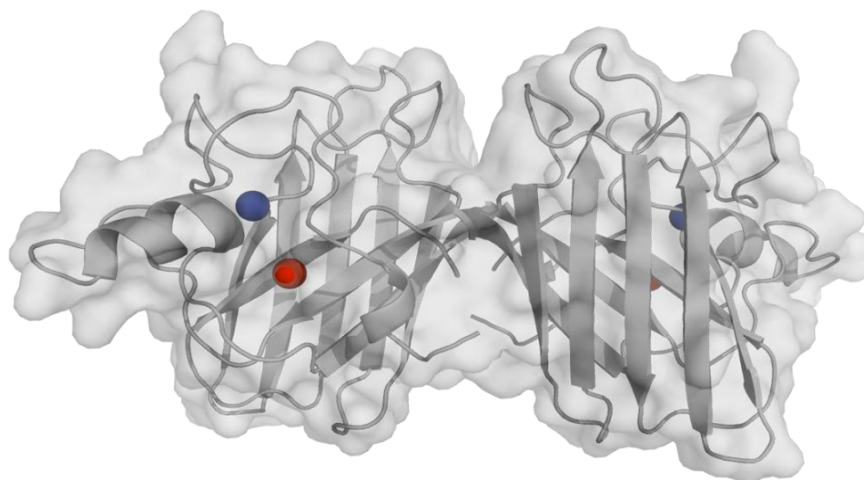


Fig.23: Ribbon representation of the dimeric OvEC-SOD structure, which is solved to 1.5 Å resolution. The structure comprises two identical monomers; the active site contains one Zn atom (blue) and one Cu atom (red).

## 4.1.8. The OvEC-SOD active site:

The active site contains one catalytically $Cu^{2+}$ ion and one structural $Zn^{2+}$ ion. The coordination sphere of the $Cu^{2+}$ and $Zn^{2+}$ ions is defined by the invariant residues His46, His48, His63, His120 and His63, His71, His80 and Asp83, respectively (Figure 24). The $Cu^{2+}$ ion is coordinated by four Histidine residues arranged in a distorted tetrahedral geometry. The $Zn^{2+}$ ion is coordinated by three histidine and one aspartate residues (Figure 24) in a tetragonal geometry. The water molecules occupy the similar positions around the $Cu^{2+}$ ion when compared with the homo sapiens SOD [PDB entry: 2V0A], except in one position where the sulfate ion, which is observed 2.06 Å away from the $Cu^{2+}$ ion. The sulfate ion binds to the $Cu^{2+}$ ion by the O2-C atom. The catalytic activity of OvEC-SOD seems to be independent of the presence of water (132). In the OvEC-SOD, His-63 forms a bridge between $Cu^{2+}$ and $Zn^{2+}$ ions that passes over 3.18 and 2.01 Å, respectively. The Cu–His63–Zn imidazolate bridge is intact in the oxidized form of the enzyme. In the reduced form of the enzyme, the imidazolate bridge is broken and the catalytic metal becomes three coordinated (133). In the oxidized form, a typical distance between Cu–Zn should be around 6.0 Å, while, in the reduced form of the enzyme, this distance should be more than

# 4 RESULTS

6.5 Å. From previous informations it is clear that OcEC-SOD crystallized in the reduced form, as the bridge was broken and the distance between $Zn^{2+}$ and $Cu^{2+}$ ions was found to be 6.85 Å.
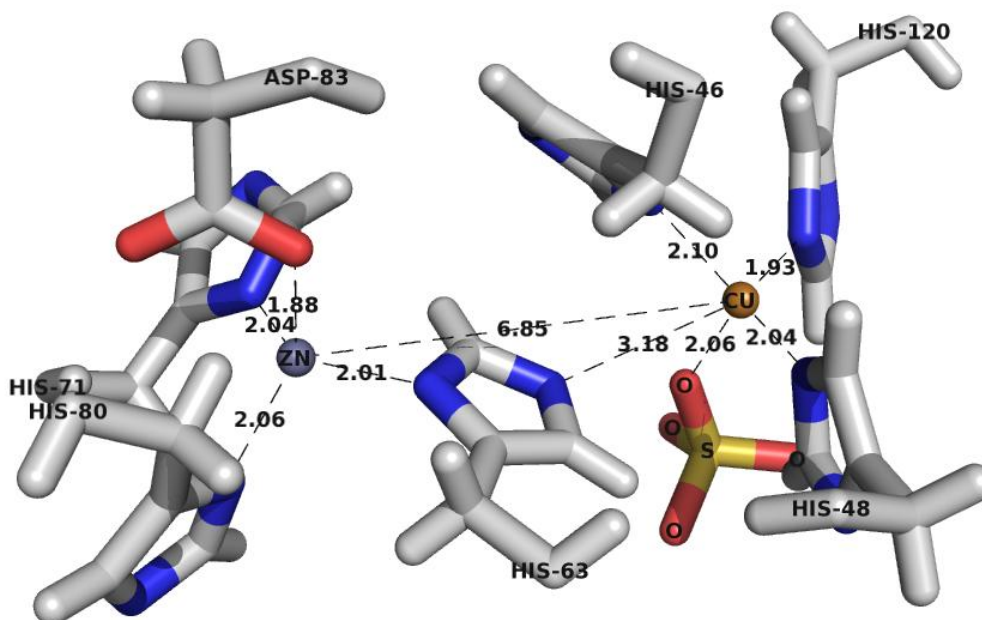


Fig.24: The OvEC-SOD active site. The $Cu^{2+}$ ligands are ND1 of His46 and NE2 of His48, His63, and His120 and show an uneven distorted tetrahedral geometry.

### 4.1.9. The active site channel:

The active site channel is formed in part by a disulfide loop (residues 50-62) and in part of an electrostatic loop (residues 130-146). The latter loop contains several charged, well-conserved residues, which provide the electrostatic potential to drive the substrate to the reaction site. This loop is well ordered and the residues show well-defined electron density in the Fo-Fc map and side-chain conformations. The active-site channel of OvEC-SOD is maintained by an intramolecular disulfide bond between the highly conserved Cys57 and Cys149 residues. The side-chain of Arg146, the residue responsible for the correct orientation of superoxide in the catalytic cavity, which is completely conserved, is well formed and stabilized by hydrogen bonding with the oxygen carbon of Cys57. The entrance of the channel to the active site shows that *O. Volvulus* Cu∕Zn-SODs vary from the human enzyme not only in amino acid sequence, but although in length and structure, as shown in figure 25. This area is called the electrostatic loop, and it is highly conserved among different species. This loop has been suggested to be responsible for directing the superoxide anions to the catalytic site (134). In human Cu∕Zn-SOD, this loop is positively

# 4 RESULTS

charged, while in the OvEC-SOD, non-polar and negatively charged residues predominate. The conformation and charge distribution is responsible for substrate attraction.
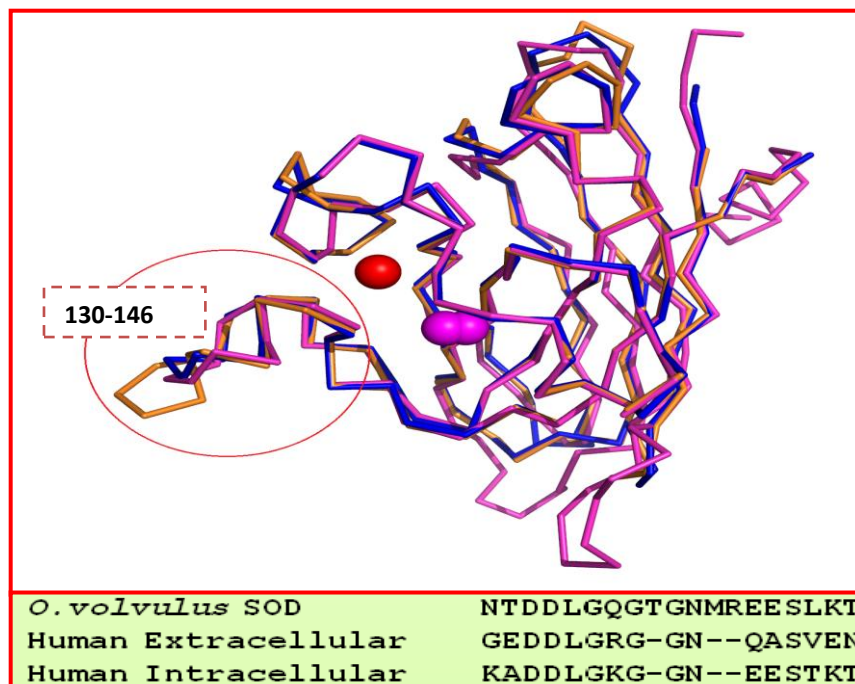


| O.volvulus SOD | NTDDLGQGTGNMREESLKT |
| Human Extracellular | GEDDLGRG-GN--QASVEN |
| Human Intracellular | KADDLGKG-GN--EESTKT |

Fig.25: Superimposition in line representation of OvEC-SOD (orange) with human intra- (blue) and extracellular SOD (purple), the electrostatic loop is surrounded by a red circle. The red sphere is the copper ion and the purple sphere represents the zink ion.

## 4.1.10. The OvEC-SOD Dimer interface:

The dimerization interface is relatively large, consisting mainly of hydrophobic interactions and four hydrogen-bond interactions. The dimer interface has an overall area of 726 $\mathring{A}^2$ per subunit. The side-chain contributions is ($\approx$ 90 %) of whole chain contacts. The dimer interface is stabilized by 83 non-bonded contacts and only 4 hydrogen bonds. The summary of binding residues and interactions donors and acceptors is listed in table 3 and 4.

Table 3: Illustrating the No. of interface residues and hydrogen bond contacts within the dimer interface.

| | Residue name | Atom name | Chain | | Residue name | Atom name | Chain | Distance |
|---|---|---|---|---|---|---|---|---|
| 1 | $GLY_{49}$ | N | A | ←→ | ILE152 | O | B | 2.81 |
| 2 | $GLY_{112}$ | O | A | ←→ | $ILE_{152}$ | N | B | 2.89 |
| 3 | $ILE_{152}$ | N | A | ←→ | $GLY_{112}$ | O | B | 2.89 |
| 4 | ILE152 | O | A | ←→ | $GLY_{49}$ | N | B | 2.81 |

# 4 RESULTS

Table 4: Illustrating of the No. of interface residues and nonbonded contacts of dimer interface.

| | Residue name | Atom name | Chain | | Residue name | Atom name | Chain | Distance |
|---|---|---|---|---|---|---|---|---|
| 1 | $VAL_3$ | CG1 | A | ←→ | $ASP_{50}$ | CA | B | 3.79 |
| 2 | $VAL_3$ | CG1 | A | ←→ | $ASP_{50}$ | CB | B | 3.80 |
| 3 | $VAL_3$ | CG1 | A | ←→ | $ASP_{50}$ | CG | B | 3.69 |
| 4 | $VAL_5$ | CG2 | A | ←→ | $GLN_{51}$ | C | B | 3.80 |
| 5 | $VAL_5$ | CG2 | A | ←→ | $THR_{52}$ | N | B | 3.81 |
| 6 | $VAL_5$ | CG2 | A | ←→ | $THR_{52}$ | CG2 | B | 3.89 |
| 7 | $ARG_7$ | NH1 | A | ←→ | $GLN_{51}$ | OE1 | B | 3.56 |
| 8 | $ARG_7$ | NH2 | A | ←→ | $GLN_{51}$ | O | B | 3.75 |
| 9 | $ARG_7$ | NH2 | A | ←→ | $GLN_{51}$ | OE1 | B | 3.84 |
| 10 | $ILE_{15}$ | CG2 | A | ←→ | $THR_{52}$ | CG2 | B | 3.82 |
| 11 | $TYR_{17}$ | OH | A | ←→ | $THR_{52}$ | CG2 | B | 3.58 |
| 12 | $TYR_{48}$ | CA | A | ←→ | $ILE_{152}$ | O | B | 3.44 |
| 13 | $TYR_{48}$ | C | A | ←→ | $ILE_{152}$ | O | B | 3.56 |
| 14 | $TYR_{48}$ | CB | A | ←→ | $ILE_{152}$ | O | B | 3.40 |
| 15 | $TYR_{48}$ | CD1 | A | ←→ | $ALA_{153}$ | C | B | 3.84 |
| 16 | $TYR_{48}$ | CD1 | A | ←→ | $ALA_{154}$ | N | B | 3.73 |
| 17 | $TYR_{48}$ | CE1 | A | ←→ | $ALA_{154}$ | N | B | 3.74 |
| 18 | $TYR_{48}$ | CE1 | A | ←→ | $ALA_{154}$ | CA | B | 3.60 |
| 19 | $TYR_{48}$ | CE1 | A | ←→ | $ALA_{154}$ | CB | B | 3.88 |
| 20 | $TYR_{48}$ | CZ | A | ←→ | $ALA_{154}$ | CA | B | 3.86 |
| 21 | $TYR_{48}$ | CZ | A | ←→ | $ALA_{154}$ | CB | B | 3.62 |
| 22 | $TYR_{48}$ | OH | A | ←→ | $ALA_{154}$ | CB | B | 3.60 |
| 23 | $GLY_{49}$ | N | A | ←→ | $ILE_{152}$ | N | B | 3.83 |
| 24 | $GLY_{49}$ | N | A | ←→ | $ILE_{152}$ | C | B | 3.81 |
| 25 | $GLY_{49}$ | N | A | ←→ | $ILE_{152}$ | O | B | 2.81 |
| 26 | $GLY_{49}$ | CA | A | ←→ | $ILE_{152}$ | O | B | 3.65 |
| 27 | $GLY_{49}$ | C | A | ←→ | $ILE_{152}$ | O | B | 3.65 |
| 28 | $GLY_{49}$ | O | A | ←→ | $GLY_{151}$ | CA | B | 3.27 |
| 29 | $GLY_{49}$ | O | A | ←→ | $GLY_{151}$ | C | B | 3.73 |
| 30 | $ASP_{50}$ | N | A | ←→ | $ILE_{152}$ | O | B | 3.75 |
| 31 | $ASP_{50}$ | CA | A | ←→ | $VAL_3$ | CG1 | B | 3.79 |
| 32 | $ASP_{50}$ | CB | A | ←→ | $VAL_3$ | CG1 | B | 3.80 |
| 33 | $ASP_{50}$ | CG | A | ←→ | $VAL_3$ | CG1 | B | 3.69 |
| 34 | $GLN_{51}$ | C | A | ←→ | $VAL_5$ | CG2 | B | 3.80 |
| 35 | $GLN_{51}$ | O | A | ←→ | $ARG_7$ | NH2 | B | 3.75 |
| 36 | $GLN_{51}$ | OE1 | A | ←→ | $ARG_7$ | NH1 | B | 3.56 |
| 37 | $GLN_{51}$ | OE1 | A | ←→ | $ARG_7$ | NH2 | B | 3.84 |
| 38 | $THR_{52}$ | N | A | ←→ | $VAL_5$ | CG2 | B | 3.81 |
| 39 | $THR_{52}$ | CG2 | A | ←→ | $VAL_5$ | CG2 | B | 3.89 |
| 40 | $THR_{52}$ | CG2 | A | ←→ | $ILE_{15}$ | CG2 | B | 3.82 |
| 41 | $THR_{52}$ | CG2 | A | ←→ | $TYR_{17}$ | OH | B | 3.58 |
| 42 | $ILE_{111}$ | CG2 | A | ←→ | $GLY_{112}$ | CA | B | 3.50 |

# 4 RESULTS

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 43 | GLY$_{112}$ | CA | A | ←→ | ILE$_{111}$ | CG2 | B | 3.50 |
| 44 | GLY$_{112}$ | O | A | ←→ | GLY$_{151}$ | CA | B | 3.44 |
| 45 | GLY$_{112}$ | O | A | ←→ | GLY$_{151}$ | C | B | 3.65 |
| 46 | GLY$_{112}$ | O | A | ←→ | ILE$_{152}$ | N | B | 3.89 |
| 47 | GLY$_{112}$ | O | A | ←→ | ILE$_{152}$ | CA | B | 3.88 |
| 48 | GLY$_{112}$ | O | A | ←→ | ILE$_{152}$ | CB | B | 3.70 |
| 49 | GLY$_{112}$ | O | A | ←→ | ILE$_{152}$ | CG1 | B | 3.76 |
| 50 | GLY$_{112}$ | O | A | ←→ | ILE$_{152}$ | CD1 | B | 3.88 |
| 51 | HIS$_{113}$ | CD2 | A | ←→ | ILE$_{152}$ | CD1 | B | 3.57 |
| 52 | HIS$_{113}$ | CE1 | A | ←→ | ILE$_{152}$ | CD1 | B | 3.31 |
| 53 | HIS$_{113}$ | NE2 | A | ←→ | ILE$_{152}$ | CD1 | B | 3.15 |
| 54 | VAL$_{149}$ | CG1 | A | ←→ | VAL$_{149}$ | CD1 | B | 3.62 |
| 55 | GLY$_{151}$ | CA | A | ←→ | GLY$_{49}$ | O | B | 3.27 |
| 56 | GLY$_{151}$ | CA | A | ←→ | GLY$_{112}$ | O | B | 3.44 |
| 57 | GLY$_{151}$ | C | A | ←→ | GLY$_{49}$ | O | B | 3.73 |
| 58 | ILE$_{152}$ | C | A | ←→ | GLY$_{112}$ | O | B | 3.65 |
| 59 | ILE$_{152}$ | N | A | ←→ | GLY$_{49}$ | N | B | 3.83 |
| 60 | ILE$_{152}$ | N | A | ←→ | GLU$_{112}$ | O | B | 2.98 |
| 61 | ILE$_{152}$ | CA | A | ←→ | GLY$_{112}$ | O | B | 3.88 |
| 62 | ILE$_{152}$ | C | A | ←→ | GLY$_{49}$ | N | B | 3.81 |
| 63 | ILE$_{152}$ | O | A | ←→ | TYR$_{48}$ | CA | B | 3.44 |
| 64 | ILE$_{152}$ | O | A | ←→ | TYR$_{48}$ | C | B | 3.56 |
| 65 | ILE$_{152}$ | O | A | ←→ | TYR$_{48}$ | CB | B | 3.40 |
| 66 | ILE$_{152}$ | O | A | ←→ | GLY$_{49}$ | N | B | 3.81 |
| 67 | ILE$_{152}$ | O | A | ←→ | GLY$_{49}$ | CA | B | 3.65 |
| 68 | ILE$_{152}$ | O | A | ←→ | GLY$_{49}$ | C | B | 3.65 |
| 69 | ILE$_{152}$ | O | A | ←→ | ASP$_{50}$ | N | B | 3.75 |
| 70 | ILE$_{152}$ | CB | A | ←→ | GLY$_{112}$ | O | B | 3.70 |
| 71 | ILE$_{152}$ | CG1 | A | ←→ | GLY$_{112}$ | O | B | 3.76 |
| 72 | ILE$_{152}$ | CD1 | A | ←→ | GLY$_{112}$ | O | B | 3.88 |
| 73 | ILE$_{152}$ | CD1 | A | ←→ | HIS$_{113}$ | CD2 | B | 3.57 |
| 74 | ILE$_{152}$ | CD1 | A | ←→ | HIS$_{113}$ | NE1 | B | 3.31 |
| 75 | ILE$_{152}$ | CD1 | A | ←→ | HIS$_{113}$ | NE2 | B | 3.15 |
| 76 | ALA$_{153}$ | C | A | ←→ | TYR$_{48}$ | CD1 | B | 3.84 |
| 77 | ALA$_{154}$ | N | A | ←→ | TYR$_{48}$ | CD1 | B | 3.73 |
| 78 | ALA$_{154}$ | N | A | ←→ | TYR$_{48}$ | CE1 | B | 3.74 |
| 79 | ALA$_{154}$ | CA | A | ←→ | TYR$_{48}$ | CE1 | B | 3.60 |
| 80 | ALA$_{154}$ | CA | A | ←→ | TYR$_{48}$ | CZ | B | 3.86 |
| 81 | ALA$_{154}$ | CB | A | ←→ | TYR$_{48}$ | CE1 | B | 3.88 |
| 82 | ALA$_{154}$ | CB | A | ←→ | TYR$_{48}$ | CZ | B | 3.62 |
| 83 | ALA$_{154}$ | CB | A | ←→ | TYR$_{48}$ | OH | B | 3.60 |

# 4 RESULTS

**4.1.11. Small-angle X-ray scattering measurements:**

A monodisperse solution containing purified protein OvEC-SOD was applied to SAXS. To monitor the radiation damage, five successive 15 s exposures of all protein solutions were matched to each other and no significant damage were observed. To confirm that the crystallographic dimer corresponds to the solution structure, we performed SAXS analyses on the OvEC-SOD protein (Figure 26). There was no concentration dependent evident of aggregation in the scattering data. The MW of 38 $\pm$2 kDa was estimated from forward scattering and suggests that the sample is dimeric in solution. An experimental $R_g$ of 2.8 $\pm$1 nm and $D_{max}$ of 9 $\pm$1 nm suggested a rhomboidal structure. The *ab initio* structure is calculated using the range of scattering vectors up to s = 0.40 Å$^{-1}$. Comparison of the experimental SAXS scattering data with the theoretical scattering curves from the crystal structures of the OvEC-SOD dimer with chi$^2$ value = 0.7 confirms that the OvEC-SOD dimer seen in the crystal is the same as that in solution.
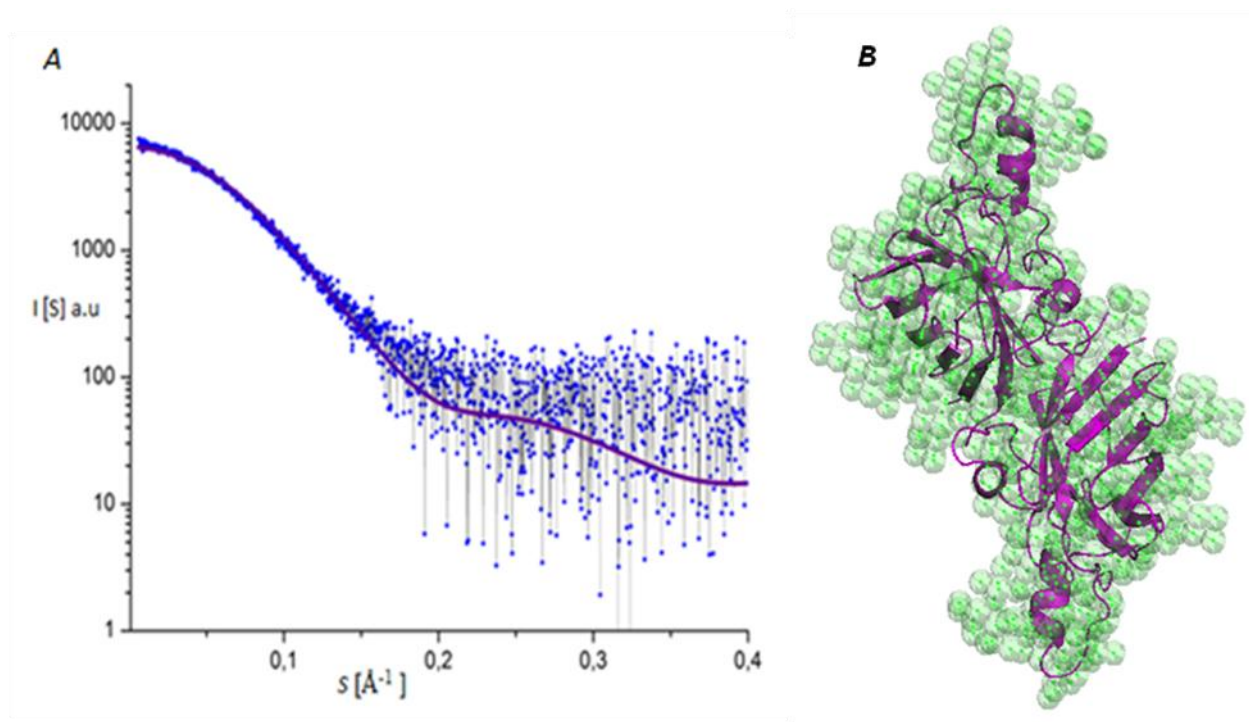


Fig.26:Processed solution scattering pattern and *ab initio* model of OvEC-SOD obtained by SAXS. A) Experimental data are displayed in blue with the theoretical scattering fit in red for crystallographic data generated with CRYSOL and the purple fit for the DAMMIF *ab initio* model. B) Side view of the dimeric OvEC-SOD *ab initio* model in green in superimposition with the crystal structure in a purple ribbon representation. The figure was generated with Pymol.

# 4 RESULTS

### 4.1.12. Novel Inhibitors and new binding site:

Before starting docking calculations, the potential sites for ligand binding from the 3D atomic coordinates of the OvEC-SOD were determined. Active sites are usually hydrophobic pockets that involve side chain atoms tightly packed (135). The residues directly or indirectly involved in metal binding are conserved among the Cu/Zn–SODs from all species so far examined, from bacteria to mammals; therefore in order to attain species selectivity, the cavities from these sites were also discarded. On this basis, ten potential sites were detected and distributed over the entire surface of the dimer, shown in table 5. One site is the more prominent one in size; this site is located inside the interface region located 19 Å away from the active site with an extended shape and additionally rich in hydrophobic residues (Figure 27). About 20 amino acids cooperate to form this binding site with a volume of 143 Å$^3$, see table 5. The amino acids are not strictly conserved between OvEC-SOD and human Cu/Zn–SOD sequences; this drives this site to be an attractive target for designing species-specific inhibitors. This binding site is similar to another binding site detected in *Taenia solium* (round worm), as shown in figure 28.
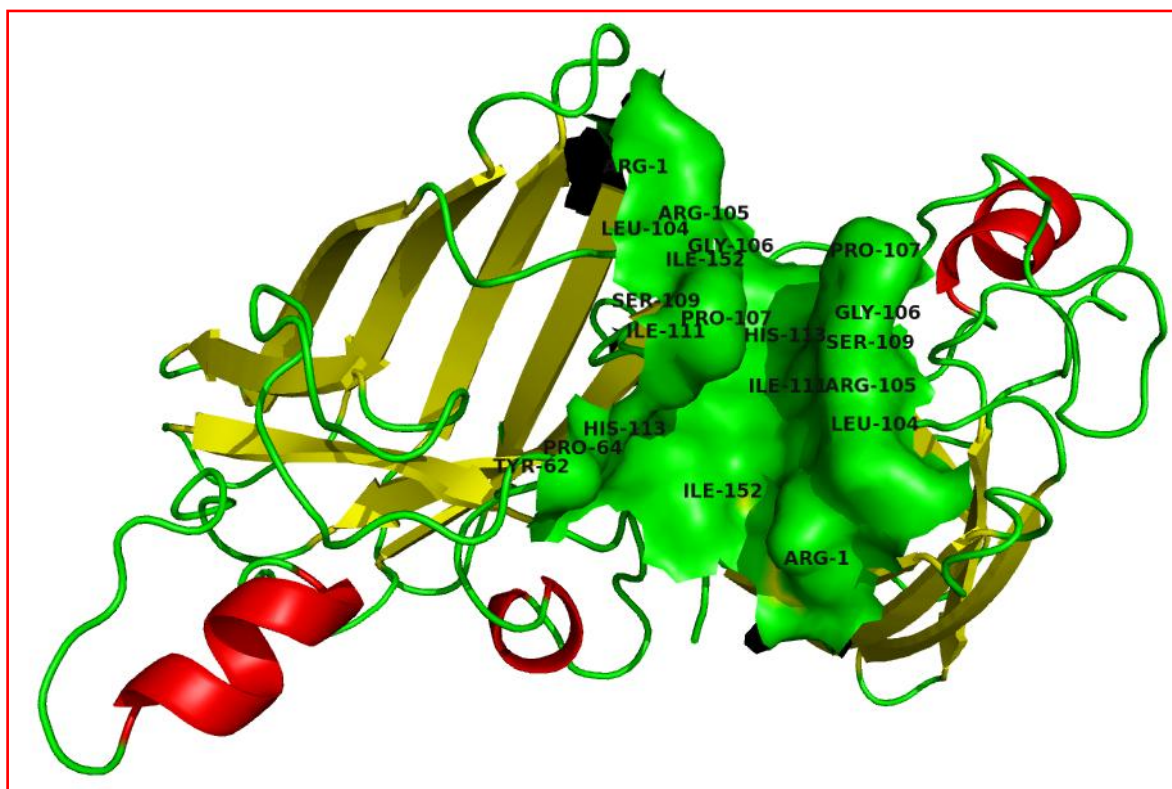


Fig.27: A) The novel binding site lying at the groove of the dimer interface region. The green color is the surface representation of the OvEC-SOD binding site.

# 4 RESULTS

Fifty compounds were applied in this study based on previous docking studies with Cu/Zn–SOD from *Taenia solium*. These compounds are confirmed as Cu/Zn–SOD ligands and contribute selectively to a binding site similar to the binding site of OvEC-SOD.
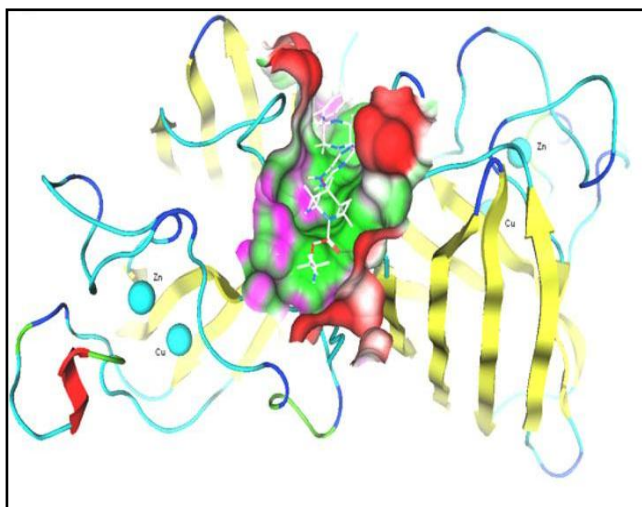


Fig.28: The similar binding site of the Cu/Zn–SOD from *Taenia solium.*

Three inhibitors were selected depending on higher score (Table 6), interestingly; these compounds bind to human Cu/Zn–SOD in unspesifically sites with low docking score. The results obtained here are very promising, and further studies should be carried out to determine the more specific inhibitors and to investigate the mechanism involved in the inactivation observed against OvEC-SOD and Hs Cu/Zn–SOD.

The mechanism of binding is mainly assured through the hydrophobic interactions, with every compound showing only one hydrogen bond. Four amino acids are conserved and seem to be essential binding partners; these amino acids are Ile104, Ile152, His113 and Ser109. Figure 29 illustrates the binding partners of these compounds.
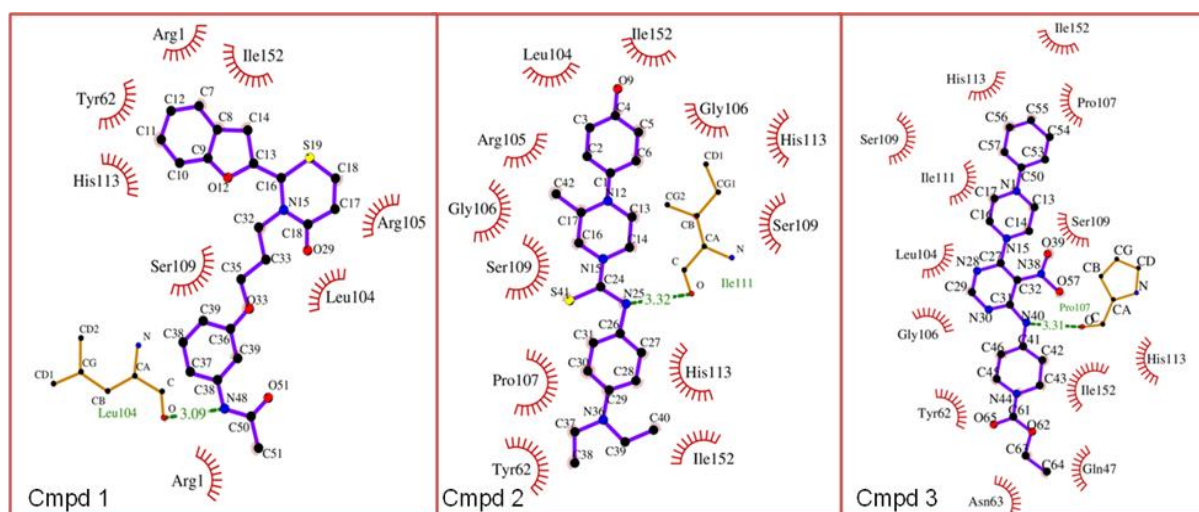


Fig.29: Virtual binding sites of the novel inhibitors with the corresponding OvEC-SOD binding sites. This figure has been created by the Ligplot+ program.

# 4 RESULTS

Table 5: Characterization of the binding sites applying the methyl carbon probe*

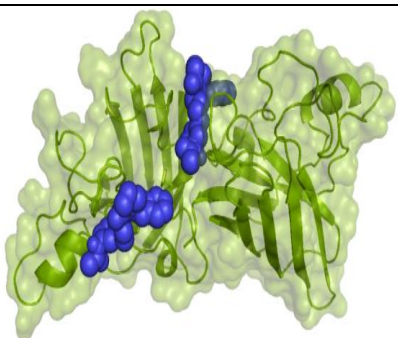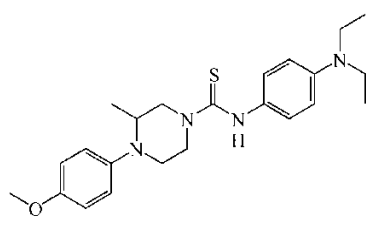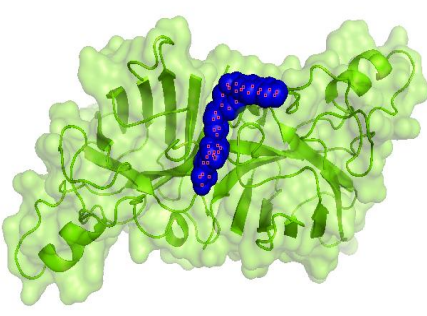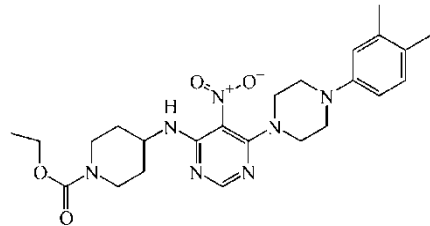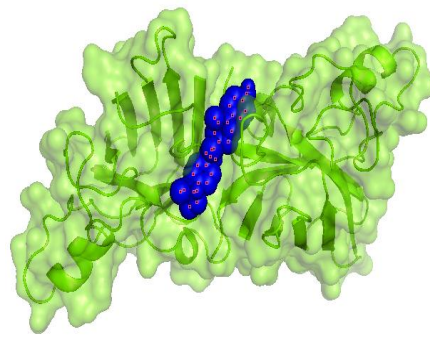| Rank | Energy[kcal/mol]* | Volume*/Å³ | Residues for Cluster |
|------|------|------|------|
| 1 | -1545 | 143.00 | Arg 1 A, Tyr 62 A, Pro 64 A, Leu 104 A, Arg 105 A, Gly 106 A, Pro 107 A, Ser 109 A, Ile 111 A, His 113 A, Ile 152 A, Arg 1 A, Leu 104 B, Arg 105 B, Gly 106 B, Pro 107 B, Ser 109 B, Ile 111 B, His 113 B, Ile 152 B |
| 2 | -674.10 | 59.00 | Asn 63 A, Phe 65 A, Lys 67 A, Pro 72 A, Asn 73 A, Asp 74 A, Arg 75 A, Ile 76 A, Lys 77 A, His 78 A, Ile 79 A, Tyr 100 A, Asp 101 A, Lys 103 A, Leu 108 A |
| 3 | -661.64 | 58.00 | Asn 63 B, Phe 65 B, Lys 67 B, Pro 72 B, Asn 73 B, Asp 74 B, Arg 75 B, Ile 76 B, Lys 77 B, His 78 B, Ile 79 B, Tyr 100 B, Asp 101 B, Lys 103 B, Leu 108 B |
| 4 | -441.76 | 27.00 | Val 3 A, Ala 4 A, Val 5 A, Arg 7 A, Gly 49 A, Gln 81 A, Cys 147 A, Gly 148 A, Val 149 A, Val 3 B, Ala 4 B, Val 5 B, Arg 7 B, Gly 49 B, Gln 81 B, Cys 147 B, Gly 148 B, Val 149 B |
| 5 | 416.33 | 36.00 | Ile 26 B, Pro 72 B, Asn 73, Lys 77 B, Tyr 96 B, Ile 97 B, Asn 98 B, Ser 99 B, Tyr 100 B, Asp 101 B |
| 6 | -406.38 | 36.00 | Val 5 A, Arg 7 A,Gly 8 A, Gln 51 A, Thr 52 A, Gly 54 A, Cys 55 A, Cys 147 A, Val 5 B, Arg 7 B, Ile 15B |
| 7 | -403.87 | 35.00 | Ile 26 A, Pro 72 A, Asn 73 A, Lys 77 A, Tyr 96 A, Ile 97 A, Asn 98 A, Ser 99 A, Tyr 100 A, Asp 101 A |
| 8 | -382.35 | 33.00 | Thr 68 A, His 69 A, Gly 70 A, Asp 74 A, Ile 76 A, Leu 124 A, Gly 125 A, Gly 126 A, Arg 132 A, Leu 136 A |
| 9 | -382.35 | 33.00 | Thr 68 B, His 69 B, Gly 70 B, Asp 74 B, Ile 76 B, Leu 124 B, Gly 125 B, Gly 126 B, Arg 132 B, Leu 136 B |
| 10 | -357.82 | 34.00 | Asp 9 A, Ala 10 A, Asn 53 A, Gly 54 A, Cys 55 A, Thr 56 A, Ser 143 A, Arg 144 A, Leu 145 A |
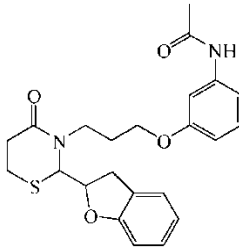
*__Total Energy:__ is the sum of interaction energy of all points contributing to this cluster.
*__Carbon probe__: is suitable for detection of small molecules binding sites nominated by Van der Waals interactions.
*__Volume__: size of cluster in Å³.

# 4 RESULTS

Table 6: Structures of new compounds, as identified by virtual screening and their binding sites.

| No | Cmpd | LeadQuest ID | Best score | Binding Site |
|---|---|---|---|---|
| 1 |  | 1545-7806 | −7.6 |  |
| 2 |  | 1526-6128 | −8.1 |  |
| 3 |  | 1502-3317 | − 7.9 |  |

# 4 RESULTS

## 4.2. Immunodominant hypodermal antigen (Ov Ral-2):

### 4.2.1. Sequence alignment and comparison:

The Ov Ral-2 precursor consisted of 164 residues. The full Ov Ral-2 (acc. no: P36991) contains an N-terminal signal peptide (16 residues) followed by a poly-Q region consisting of 11 glutamine residues, then a conserved domain follows with unknown function (101 residues). To get a better idea about the putative functional important regions of the Ral-2/SPX family, the the Ral-2 family was compared regarding sequence conservation (figure 30). In general, the intrahelical coil regions are more conserved. Based on this analysis, 40 fully reserved amino acids were identified, of which five were aromatic amino acids and six alanine and four proline residues. It is obvious that Ov Ral-2 is the only family member which has this poly-Q region, but it is unclear if this protein performs different functions. not.

```
RAL2 Onchocerca_volvulus   ------------MKFVILLTI-GLLVVAAIPQRRQQQQQQQQQQQRDERE----IPPFLEGAPPSVIDEF
Ani S 5 Anisakis Simplex   ------------MKTLIVAALFCTIGMAL-----------ADDTPPP--------PPFLAGAPQDVVKAF
OV17 Toxocara Canis        ------------MKSLIAAAI-CLFAFAV-----------AQNPPP--------PPFLVGAPENIIKEF
SXP-1 protein Loa Loa      MPSITLSVNPTKIYCLILLSI-GLITETL----------AQRRVEPPQQDALPGIPPFLDGAPPDVVQRF
SPX-1 antigen B Malayi     ---VTSSLNLTKMKYFIFLSI-GLIAAAS---------AQREAQLPQPE----IPPFLSGVPSHVVKQF
SPX antigen W Bancrofti    ---VTSSLNLTKMKYFIFLSI-GLIAGAL---------AQREAQIPQSD----IPPFLSGAPNHVVKQF
                                      :  .*  ::   :  :                          **** *.*  ::. *

RAL2 Onchocerca_volvulus   YNLLKTDENKTDQQTEADVEAFINRLGGSYKVRFTQFMEEVKKARADYERIHQQAVARFSPAAKDADARM
Ani S 5 Anisakis Simplex   FELLKKDETKTDPEIEKDLDAWVDTLGGDYKAKFETFKKEMKAKEAELAKAHEEAVAKMTPEAKKADAEL
OV17 Toxocara Canis        NDLIQRDESKTDSQIEADLNAWVAKLGGDYKAKFEAFKNELKQHQAEFDKAHAAAVAKFSPAAKEADAKL
SXP-1 protein Loa Loa      YDLVN-NESKTTPEIEADVEKFIHKLGGNYTTRFEKFKEEVKRLNAEFETQRSKIVSNLSPEARKADAML
SPX-1 antigen B Malayi     FDLLKADESKTDPQTEADIEAFIRRLGGDYQTRFEQFKQEIKKEKAQYEKIHQAALLKFSPAAREADAKM
SPX antigen W Bancrofti    FDLLRADESKTDPQTEADIEAFMRRLGGVYQARFEQFKQEMKKQFAQYDKVHQAALSRFSPAARQADARM
                             *:. :*.**  : * *:: ::  *** *  ..*  * :*:*   *:    .   : .::* *..***

RAL2 Onchocerca_volvulus   SAIADSPHLTTRQKSQQIQAIMDSLSESVRREIINALSPQE-------------------------
Ani S 5 Anisakis Simplex   SKIAEDDSLNGIQKAQKIQAIYKTLPQSVKDELEKGIGPAVPQ----------------------
OV17 Toxocara Canis        SAIAADSKLNGIQKRKQIKAIFDGLPKQVREEIEKAAGGGK------------------------
SXP-1 protein Loa Loa      SDIAKDMHLTNRQKTERILTIMNSLSETVRNEILEALRPK-------------------------
SPX-1 antigen B Malayi     SAIADSTQLTNHQKTEQIKAIMDSLSEAVRKEILEGFNSQ------------------------
SPX antigen W Bancrofti    SAIAESKQLTVKQKTEQIKAIMDSLSESVRKEILEGFNSKFCVISVLLNVTIKIFSKWRKNHMRQKSNK
                           * **  .  *.  ** :.* :* . *.: *. *: :.
```
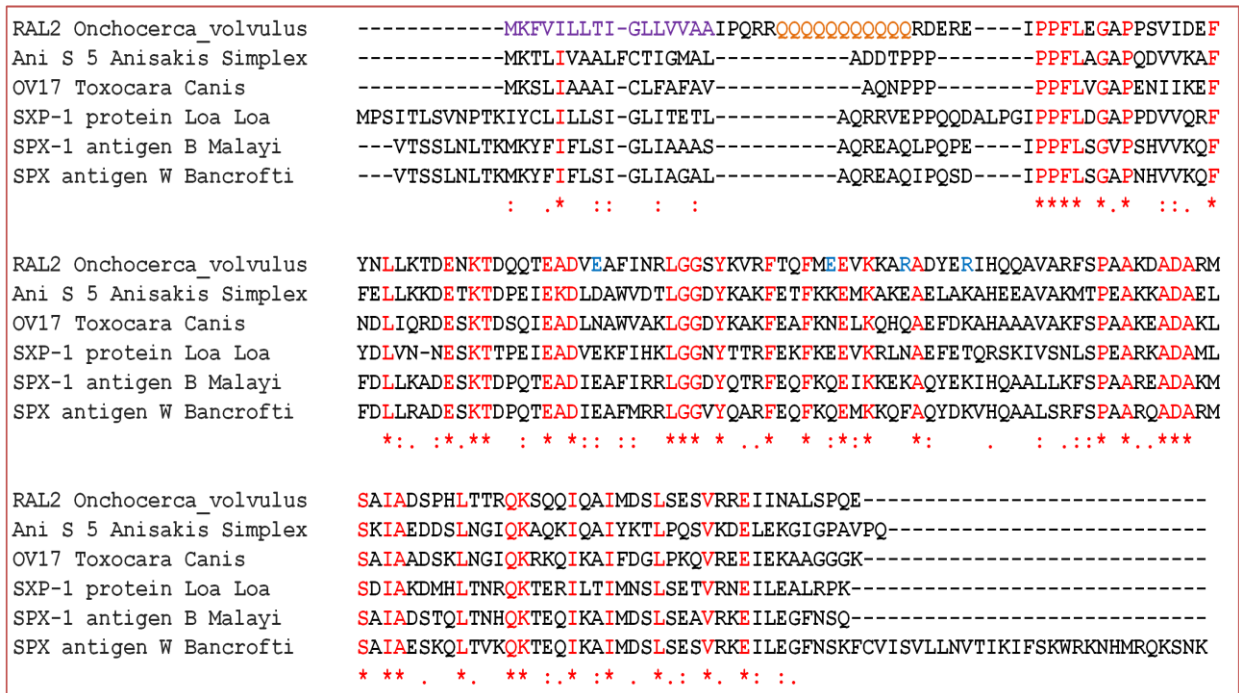
Fig.30: Sequence alignment of the Ov Ral-2 with the other homologue proteins, the brown color at the beginning is the signal peptide and the violet color refers to the poly-Q tract. The red amino acids are conserved, while the blue residues form salt bridges.

### 4.2.2 Expression and purification of the full length Ov Ral-2:

The full length Ov Ral-2 was expressed in BL21 Star (DE3) *E. coli* cells (Stratagene, USA) heterologous. Expression was performed in LB medium containing 50 ug/ml ampicillin. Once the bacterial culture OD 600 has reached a value of 0.6, expression was induced with 1 mM IPTG and incubated for 4 h at 37 °C. After the cell disruption, the full

# 4 RESULTS

Ov Ral-2 was purified by Ni-NTA affinity chromatography. To investigate the purity of the protein fractions, they were analyzed by SDS PAGE and native gel electrophoresis. The band size showed 19 kDa, however the native gel electrophoresis provided a double value (38 kDa). The final yield of the protein was ~10 mg per 4 liters of *E. coli* culture medium with more than 95 % purity, shown in figure 31.
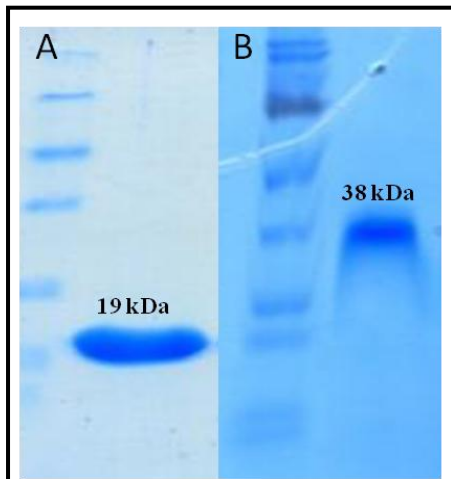


Fig. 31: A) SDS-PAGE of the full length Ov Ral-2, B) Native PAGE of the full length Ov Ral-2.

### 4.2.3. Confirming the dimeric state of the full length Ov Ral-2:

The purified full length Ov Ral-2 was analyzed for its molecular weight by Superdex 200 chromatography including a calibration curve (Figure 32a, b). The Superdex 75 calibration curve indicated a higher mass (38 kDa) for the full-length protein, which corresponds to the dimer state. This finding was similar to the result obtained from native gel electrophoresis, see figure 31b.
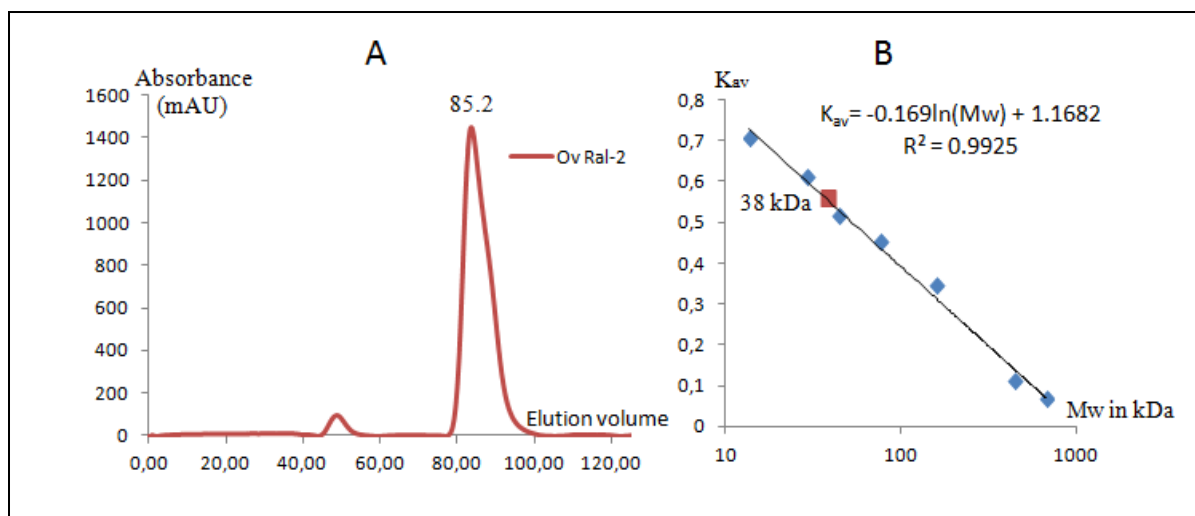


Fig.32: A) Superdex 200 SEC profile of the full length Ov Ral-2. B) The calibration plot showing the calculated molecular weight.

# 4 RESULTS

## 4.2.4. Sub cloning, protein expression and purification of Ov Ral-2$^{17-148}$:

Since the trials to produce good diffracting crystals from the full-length Ov Ral-2 construct were successful, we sub cloned the full-length clone to exclude the poly-Q region as possible source of flexibility. The ORF of the sequence encoding the Ov Ral-2$^{17-148}$ was amplified by polymerase chain reaction (PCR) using the pJC40 vector containing the full length as a template, where the NcoI and BamH1 restriction-enzyme cleavage sites were used as cloning sites. The generated PCR product was cloned into the *E. Coli* expression vector pJC40 (supplied by Prof. Eva Liebau-Muenster University) that provides the recombinant protein with an N-terminal hexahistidine (His6) tag that facilitates protein purification. Figure 33 illustrates the steps of cloning.
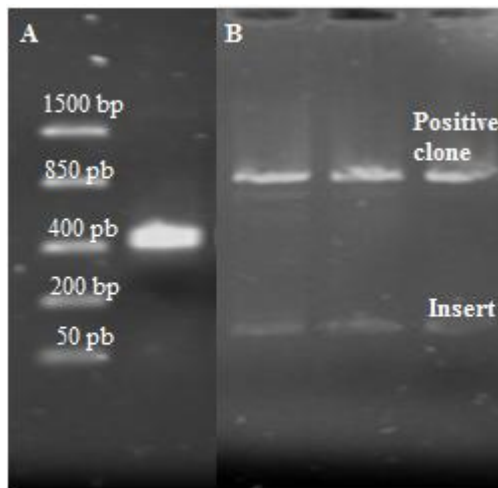


Fig.33: A) Agarose gel electrophoresis of the PCR product showing the Ov Ral-2 gene after removal of the poly-Q region. B) The digestion products of the positive clone after digestion with NcoI and BamH1, the positive clone possesses two bands of empty vector and the right size insert.

The recombinant plasmid was verified by DNA sequencing, as shown in figure 34. The competent cells (*E.Coli* BL21 star) were transformed with the recombinant vector.



Fig. 34: The sequencing result of the Ov Ral-2$^{17-148}$, the blue sequence is the NcoI restriction site, the red sequence is the His-tag, while the green color points out the BamH1 restriction site and the stop codon is shown in grey color.

# 4 RESULTS

Protein was expressed and purified by affinity chromatography using the His-tag as described in the chapter 3.3.6. Maximum expression level was obtained after 4 h. The level of expression of the target protein was optimized by varying the time and/or temperature of induction and the concentration of the inducer to obtain a maximum expression level. The expected protein band appeared at the level of 16 kDa, see figure 35. The protein yield was 6 mg/L *E. coli* growing culture and judged to be > 95 % pure by SDS–PAGE.
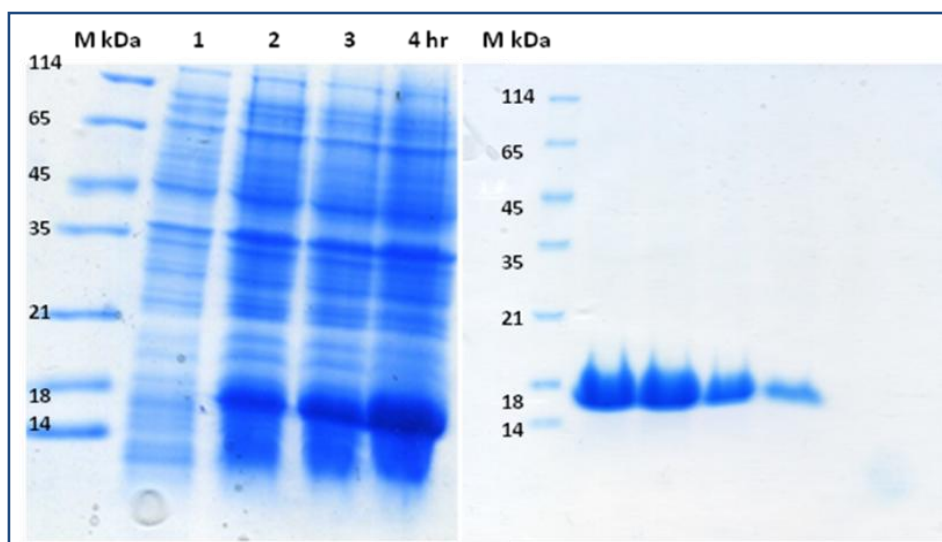


Fig.35: Left: The expression profile of the Ov Ral-2$^{17\text{-}148}$ with an N-terminal His-tag inside the bacterial strain *E.Coli* BL21 DE3. Right: The purified Ov Ral-2$^{17\text{-}148}$ protein after one step purification.

## 4.2.5. Expression of SeMet-substituted Ov Ral-2$^{17\text{-}148}$ protein:

To confirm the selenomethionine incorporation, the native and SeMet-substituted Ov Ral-2$^{17\text{-}148}$ proteins were analyzed by matrix-assisted laser desorption/ionization reflection time-of-flight (MALDI-TOF) mass spectrometry. Three different modified peptides were indentified; the difference of molecular mass between the native and SeMet-substituted peptides was 47 Da, which corresponds to the difference in molecular mass between the methionine and selenomethionine residues.

## 4.2.6. Biophysical characterization of the full Ov Ral-2 and Ov Ral-2$^{17\text{-}148}$:

The folding states of purified full Ov Ral-2 and Ov Ral-2$^{17\text{-}148}$ were checked by applying circular dichroism (CD). The far-UV CD spectra of both proteins showed α helical profiles with a double minimum at 208 nm and 222 nm along with the maxima near 190 nm. The $MRE_{222}/MRE_{208}$ ratio of 1.084 indicated the coiled coil confirmation. The folding of Ov Ral-2$^{17\text{-}148}$ has an overall approximation of the composition of helix, sheet, turn and random coils with 73.7 %, 0 %, 11.5 % and 14.7 %, respectively, while the full Ov Ral-2

# 4 RESULTS

has approximately an overall conformation consisting of helix, sheet, turn and random coils with 61.2 %, 0 %, 12.5 % and 26.3 %, respectively, as shown in figure 36.
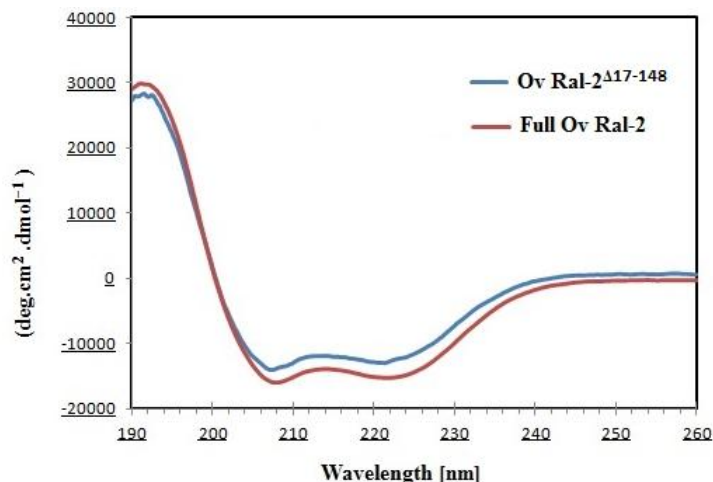


Fig.36: The adjusted CD spectrum of the native full Ov Ral-2 (red) and Ov Ral-2[17-148] (blue) with two negative peaks at approximately 208, 222 nm.

The pure solutions of the full Ov Ral-2 and Ov Ral-2[17-148] were monitored by dynamic light scattering (DLS), by this an hydrodynamic radius of approximately 3.6 and 3.1 nm was observed. The radius distribution of the protein obtained by DLS indicates a monodisperse and non-aggregated protein in the solution, as illustrated in figure 37.
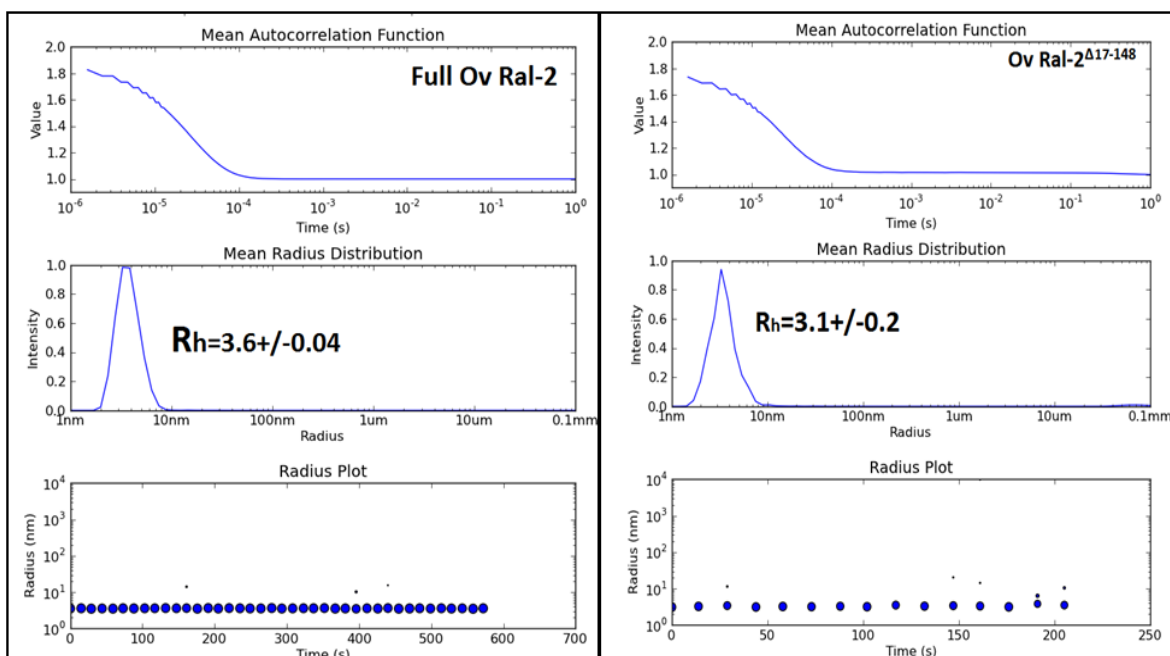


Fig.37: DLS measurement of the full Ov Ral-2 and Ov Ral-2[17-148], peaks were observed at 3.6 and 3.1 nm hydrodynamic radius, respectively.

# 4 RESULTS

## 4.2.7. Crystallization of the native and SeMet-substituted Ov Ral-2$^{17-148}$ protein:

The concentrated native and selenomethionine substituted Ov Ral-2$^{17-148}$ protein were initially screened against a diversity of available screening conditions using a Honeybee 961 dispensing robot (Genomic solutions, UK) at 293 K, combined with either a 2-well MRC plate (NUNC) or a Nextal Qia1 plate (Qiagen, Germany), applying the sitting-drop vapour-diffusion method. Based on the commercially available JCSG+, ComPAS, Classics and Cryos Suites (NeXtal, Qiagen), 450 nl droplets of 20 mg ml$^{-1}$ protein were mixed with the same volume of precipitant solutions and equilibrated against 45 μl reservoir solutions. Needles and crystalline states appeared after one week in condition D4 from the Index Suite, consisting of 0.1 M citric acid, pH 3.5, 25 % polyethylene glycol (PEG) 3350 in the reservoir (shown in figure 38a). To optimize and improve the crystal size and quality, Microseed Matrix Screening was applied using the Oryx 4 (Douglas Instruments Ltd.). The initial screening condition was optimized by changing the precipitant/protein/seed stock ratio. The final protein/precipitant ratio was 3:2:1 precipitant/protein/seed stock, respectively. After five days, rod-shaped crystals were obtained using 0.1 M citric acid, pH 4.5, 18 % polyethylene glycol (PEG) 3350 in 24-well Linbro plates (figure 38 b). The native and the selenomethionine labeled crystals were fished using nylon loops with different sizes, then briefly immersed in mother liquour supplemented with 10 % glycerol as a cryoprotectant. Prior to data collection, crystals were flash cooled in a nitrogen-gas stream at 100 K. The data from the peak wavelength displayed a very clear anomalous signal and the SAD technique was used. The data collection statistics are presented in Table 7.

| A | B | C |
|---|---|---|
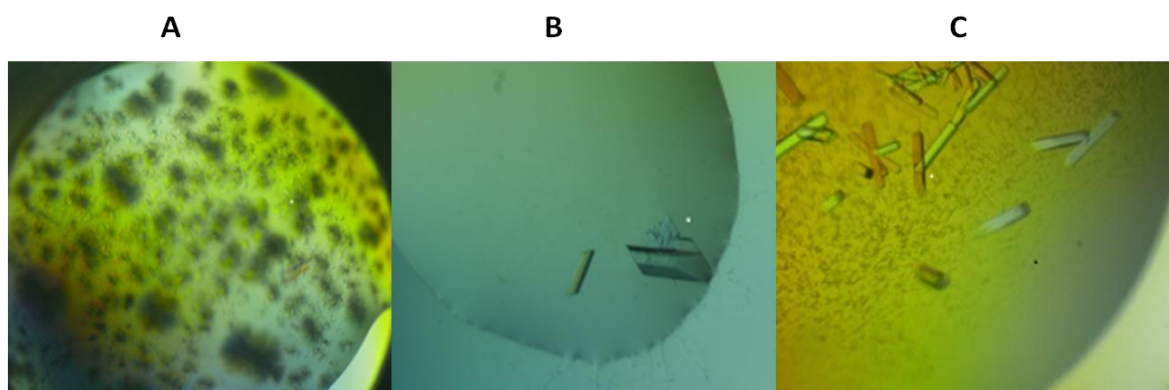


Fig.38: Ov Ral-2$^{17-148}$ crystals grew within one week applying the Index Suite D4 (Hampton Research). A) Initial needles, B) The crystals after microseeding optimization, C) The selenomethionine labeled crystals.

# 4 RESULTS

## 4.2.8. Data collection, model building and refinement:

Since attempts to solve the structure of the poly-Q missing Ov Ral-2 by molecular replacement using the NMR derived coordinates of Ani s 5 from *Anisakis simplex* (PDB codes:2MAR) were unsuccessful, SAD (*Single Anomalous Dispersion*) was applied to solve the phase problem. SeMet and native data sets were collected at the beamline P11, PETRA III (DESY/Hamburg) and the crystals diffracted up to 2.3 and 1.6 Å resolution, respectively. Phases were determined by SAD using the program AutoSol from the PHENIX software suite (136) by exploiting the anomalous signal. The initial electron density map was of sufficient quality to build approximately 90 % of one polypeptide chain. The first 14 amino acids had no electron density. The first amino acid that showed electron density was Pro31, therefore a new name was given: Ov Ral-2$^{31-148}$. The initial models were subjected to crystallographic refinement, initially using REFMAC5 (137) and further with PHENIX (138) in combination with the inspection of the electron density maps by using the program Coot (139). These programs were used to complete and refine the model to 1.6 Å resolution and to an $R_{value}$ of 15.5 % and $R_{free}$ of 19.8 %. The final model also contains 140 solvent water molecules, 3 glycerol molecules and 1 chloride ion. The Ramachandran plot (Figure 39) indicates that 100 % of the main-chain dihedral angles are located in the most favorable regions, as shown in table 7.
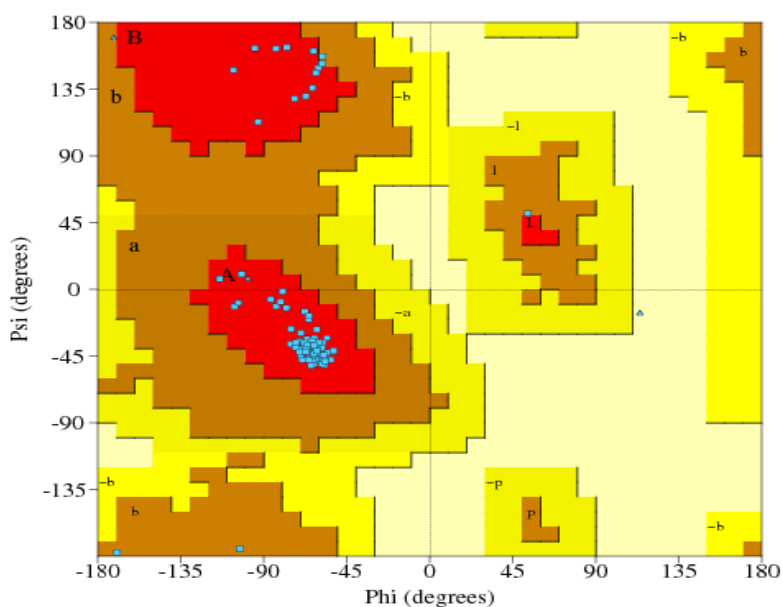


Fig.39: A Ramachandran plot generated from Ov Ral-2$^{31-148}$ coordinates. The red, brown and yellow regions represent the favored, allowed, and "generously allowed" regions, as defined by ProCheck program.

# 4 RESULTS

**Table 7: Data collection and refinement statistics of the Ov Ral-2[31-148] structure (PDB code: 5IN0):**

| Data collection | Se-derivative | Native |
|---|---|---|
| Wave length (λ Å) | 0.979 | 1.033 |
| Source | PETRA III P11 | PETRA III P11 |
| Space group | P $4_1 2_1 2$ | P $4_1 2_1 2$ |
| Cell dimensions | | |
| $a$, $b$, $c$ (Å) | 103.6, 103.6, 36.9 | 103.1, 103.1, 36.7 |
| Resolution | 100.0 - 2.3 | 19.5 - 1.6 |
| Unique reflections | 116679 | 559545 |
| Completeness (%) | 99.9 (100.0) | 99.9 (99.9) |
| Averaged redundancy | 6.82 (6.4) | 25.0 (25.4) |
| Average I/σ(I) | 13.01 (3.1) | 28.11 (4.0) |
| $R_{merge}$ (%) | 10.9 (56.5) | 6.8 (95.9) |
| **Structure solution (autosol)** | | |
| No. of sites | 3 | |
| Skew | 0.22 | |
| CORRRMS | 0.85 | |
| Figure of merit (FOM) | 0.37 | |
| Estimated map CC | 0.54 | |
| **Structure building (autobuild)** | | |
| Residues built | 106 | |
| $R_{work}$, $R_{free}$ | 23.1, 26.5 | |
| Map CC | 0.85 | |
| **Refinement** | | |
| Resolution (Å) | | 19.5-1.6 |
| $R$ factor (%) | | 15.2 (17.6) |
| Free $R$ factor (%) | | 19.8 (25.8) |
| Overall B factor (Å2) | | 38.3 |
| **R.m.s. deviations** | | |
| Bond lengths (Å) | | 0.008 |
| Bond angle (º) | | 0.96 |
| **Ramachandran plot (%)** | | |
| Most favoured | | 100.0 |
| Additionally allowed | | 0.0 |
| **Nº of molecules** | | |
| Protein | | 118 |
| Water | | 140 |
| Chloride ions | | 1 |

†$R_{merge} = \Sigma_{hkl} \Sigma_i | I_i(hkl) - (I(hkl)) | / \Sigma_{hkl} \Sigma_i I_i$ (hkl), where $(I(hkl))$ is the mean intensity of the observations $I_i(hkl)$ of reflection $hkl$. Values in parentheses are for the highest resolution shell. Skew - Deviation from a Gaussian distribution; $CRR_{RMS}$ - correlation of a local r.m.s density. †† $R$-work $= \Sigma \, ||F \, o \, | - |F \, c \, ||/\Sigma \, |F \, o \, |$ where F o and F c are the observed and calculated structure factor amplitudes, respectively.††† $R$-free is the $R$-factor calculated with 5 % of the reflections chosen at random and omitted from refinement. R.m.s is a root-mean-square deviation of bond lengths or bond angles from ideal geometry.

# 4 RESULTS

## 4.2.9. The overall structure of the Ov Ral-2$^{31-148}$:

The structure comprises one monomeric molecule in the asymmetric unit; the biological unit is built from a dimer with two identical subunits that are related by a non crystallographic 2 fold axis that is oriented as an interlocking two ring symmetry (figure 40b). Each monomer is formed of six α helices, which resemble 100 % of the secondary structure (Figure 40a). The angels between the helices are very sharp. The N-terminal region after the end of the poly-Q tract until Pro31 shows no electron density and is fully missing in the structure. The folding mechanisms and functions are not well understood, but possibly the flexible properties of the N-terminal region play role in such type of folding and dimer formation. The function of knots has been discussed that they might increase thermal and kinetic stability or prevent proteins from repaid degradation by proteolytic enzymes.
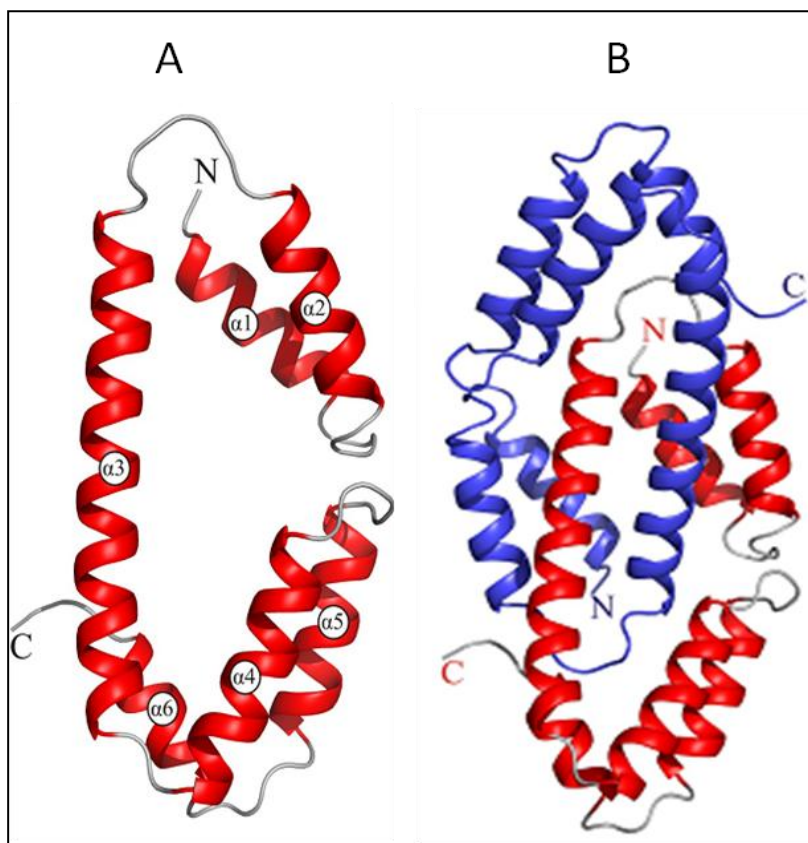


Fig.40: The overall structure of the Ov Ral-2$^{31-148}$. A) A ribbon representation of the asymmetric unit molecule. B) Biological assembly created by the PISA software (124).

The dimerization interface is relative large, consisting of a variety of salt bridges, hydrophobic interactions and hydrogen-bonds interactions. The dimer interface has an overall area of 2223 Å$^2$ per subunit. The main-chain contributions (backbone) consist of (≈ 100 %) whole chain contacts. 43 amino acids contribute to form 171 non bonded contacts,

# 4 RESULTS

14 hydrogen bonds and 4 salt bridges. The summary of binding residues and interaction donors and acceptors is listed in tables 8 and 9.

Table 8: Illustration of the interface residues and salt bridge contacts within the dimer interface:

| No. | Residue name | Atom name | Chain | | Residue name | Atom name | Chain | Distance |
|---|---|---|---|---|---|---|---|---|
| 1 | GLU$_{57}$ | OE2 | A | ←→ | ARG$_{87}$ | NH1 | B | 3.9 |
| 2 | GLU$_{76}$ | OE1 | A | ←→ | ARG$_{82}$ | NH2 | B | 3.1 |
| 3 | ARG$_{82}$ | NH2 | A | ←→ | GLU$_{76}$ | OE1 | B | 3.1 |
| 4 | ARG$_{87}$ | NH1 | A | ←→ | GLU$_{57}$ | NE2 | B | 3.9 |

Table 9: Illustrating of the interface residues and hydrogen bond contacts within the dimer interface:

| No. | Residue name | Atom name | Chain | | Residue name | Atom name | Chain | Distance |
|---|---|---|---|---|---|---|---|---|
| 1 | ASP$_{35}$ | OD1 | A | ←→ | TYR$_{38}$ | OH | B | 2.76 |
| 2 | TRY$_{38}$ | OH | A | ←→ | ASP$_{35}$ | OD1 | B | 2.76 |
| 3 | ASN$_{61}$ | O | A | ←→ | GLN$_{147}$ | O | B | 3.04 |
| 4 | GLY$_{65}$ | N | A | ←→ | LEU$_{144}$ | N | B | 3.33 |
| 5 | TYR$_{67}$ | OH | A | ←→ | ASP$_{104}$ | OD1 | B | 2.61 |
| 6 | GLN$_{73}$ | OE1 | A | ←→ | LYS$_{121}$ | NZ | B | 2.79 |
| 7 | GLN$_{73}$ | NE2 | A | ←→ | THR$_{118}$ | OG1 | B | 3.09 |
| 8 | GLU$_{76}$ | OE1 | A | ←→ | ARG$_{82}$ | NH2 | B | 3.13 |
| 9 | ARG$_{82}$ | NH2 | A | ←→ | GLU$_{76}$ | OE1 | B | 3.13 |
| 10 | ASP$_{104}$ | OD1 | A | ←→ | TYR$_{67}$ | OH | B | 2.61 |
| 11 | THR$_{118}$ | OG1 | A | ←→ | GLN$_{73}$ | NE2 | B | 3.09 |
| 12 | LYS$_{121}$ | NZ | A | ←→ | GLN$_{73}$ | OE1 | B | 2.79 |
| 13 | LEU$_{144}$ | O | A | ←→ | GLY$_{65}$ | N | B | 3.33 |
| 14 | GLN$_{147}$ | N | A | ←→ | ASN$_{61}$ | O | B | 3.04 |

## 4.2.10. The Ov Ral-2$^{31\text{-}148}$ surface properties:

To obtain an insight into the possible functionally important sites on the Ov Ral-2$^{31\text{-}148}$ domain surface, its surface properties were analyzed. The top surface contains a cluster of aromatic amino acids. The distribution of aromatic residues (see figure 41) on the upper surface of the Ov Ral-2$^{31\text{-}148}$ indicates a possible surface for protein interaction with several aromatic rings . For example, carbohydrate ligands in antibodies, for instance, would be candidates for an interaction with the Ov Ral-2 antigen. The apparently concerted movement of large side chains at the top face between the two crystal conformations was observed, this indicates that the model shows high flexibility. This also shows some

# 4 RESULTS

flexibility in the structure, which could be necessary for interactions. A plasticity of the aromatic surface could allow binding of diverse proteins and/or carbohydrate ligands.
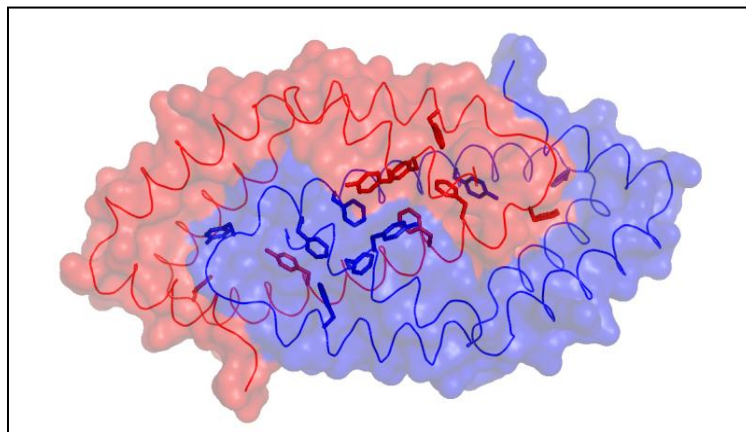


Fig.41: Surface representation of the dimeric Ov Ral-2$^{31-148}$ that shows the aromatic amino acids distribution on the upper surface of the protein.

The electrostatic surface of the Ov Ral-2$^{31-148}$ domain showed an unique distribution of the positive und the negative charges. The positive charges may be necessary for attracting binding partners of the Ov Ral-2$^{31-148}$. The electrostatic interactions play a role in protein folding by contributing to conformational stability. Typical charge-charge interactions that favor protein folding are those between oppositely charged residues, such as ARG or LYS and ASP or GLU. The electrostatic potentials were calculated by applying the Adaptive Poisson-Boltzmann Solver (APBS) [a software package for modeling bimolecular solvation by solving the Poisson-Boltzmann equation (PBE)]. The outer surface of the monomeric Ov Ral-2$^{31-148}$ is more negative, while inner surface has predominant positive charges, as shown in figure 42. Such charge distributions can contribute to knot formation during protein folding.
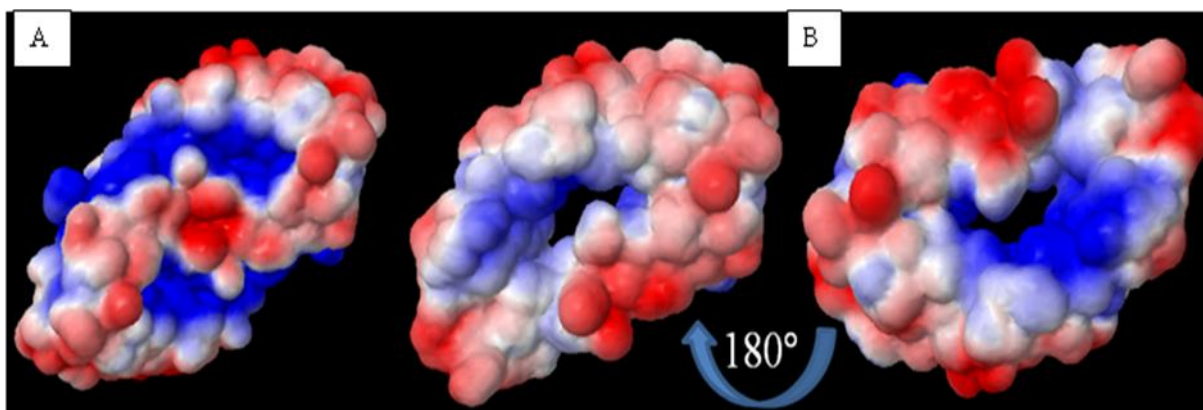


Fig.42: A) APBS-generated electrostatic surface of the dimeric Ov Ral-2$^{31-148}$ B) the electrostatic surface of monomeric Ov Ral-2$^{31-148}$ as displayed by PyMOL. The blue color represents positive charges and the red color negative charges.

# 4 RESULTS

### 4.2.11. Predection of the protein-ligand binding sites:

In the trial to identify the ligand binding sites on the the Ov Ral-2$^{31-148}$ surface, a binding-site identification server powered by the SITEHOUND program was used (140). This structure-based method recognizes geometrical features, such as a presence of clefts, grooves or energetic features that define the interaction properties of the protein surface. Structure-based methods can be observed as complementary to sequence-based methods that use evolutionary information. They apply an energy-based approach for recognizing regions that harbor a high potential for interaction with ligands. Binding site identification programs can help in finding and characterizing the parts of the protein to which a ligand may bind, but there is no guarantee that a detected site will bind a ligand or not. Two types of probes are mostly available:[1] A carbon probe for the identification of binding localities for molecules that interact principally through van der Waals contacts.[2] The phosphate probe is exploited to characterize the sites that bind to phosphorylated ligands. The carbon probe supports the drug-like molecules screening (141) and the phosphate probe detects binding sites with phosphosugars, phosphopeptides and ATP. As a result of using this program, three binding sites were identified applying the carbon probe* and the phosphate probe*, respectively, as shown in figure 43. Scoring of predicted binding sites is size and energy dependant. More details about binding sites clusters are shown in Table 10 and 11.
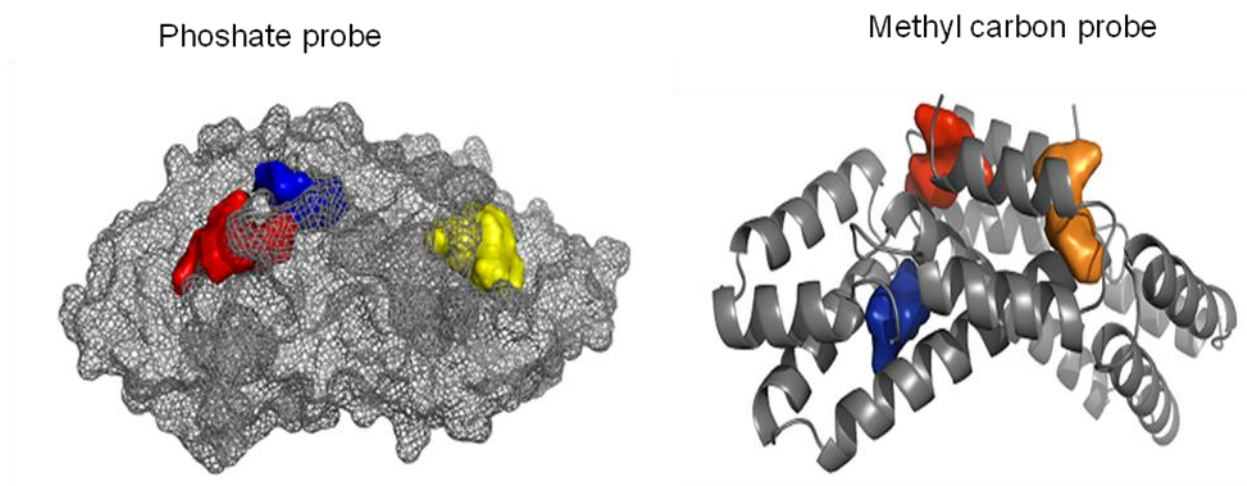


Fig. 43: Detection of proposed binding sites applying SITEHOUND-web: a server for ligand binding site identification in protein structures.

# 4 RESULTS

Table 10: Characterization of binding sites applying the methyl carbon probe *

| Rank | color | Energy[kcal /mol]* | Volume/Å³ | Residues for Cluster |
|---|---|---|---|---|
| 1 | Orange | -853 | 63.00 | Tyr 54 A, Leu 57 A, Lys 58 A, Asp 60A, Glu 60 A, Lls 63 A, Thr 64 A, Asp 65 A, THR 68 A, Thr 134 A,Ser 48 B, Ile 50 B, Asp 51 B, Arg 86 B, Gln 89 B, Phe 90 , Glu 93 B |
| 2 | Red | -853 | 63.00 | Ser 48 A, Ile 50 A, Asp 51 B, Arg 86 B, Gln 89 A, Phr90 A, Glu 93A, Tyr 54 B, Leu 57 B, Lys 58 B, Asp 60 B, Glu 61 B, Lys 63 B, Thr 64 B, Asp65 B, Thr 68 B, Thr 134 |
| 3 | Blue | -700 | 61.00 | Asp 65 A, Arg 98 A, Tyr 101 A, Glu 102, His 105 A, Ser 124 A, Ala 127 A, Gly 128 A, Tyr 83 B, Thr 88 B, Gln89 B, Met 91 B, Glu 92B |

***Total Energy:** is the sum of interaction energy of all points contributing to this cluster.
***Carbon probe**: is suitable for detection of small molecules binding site nominated by Van der Waals interactions.

Table 11: Characterization of binding sites applying the phosphate probe *

| Rank | Color | Energy[kcal /mol] | Volume/Å³ | Residues for Cluster |
|---|---|---|---|---|
| 1 | Yellow | -2807 | 176.00 | Thr 134 A, Ser 138 A, Ile 141 A, Met 145 A,Leu 160 A, Pro 47 B, Ser 48 B, Val 49 B, Ile 50 ,Gly 80 B, Gly81 B, Ser 82 B, Tyr 83 B, Lys 84 B, Val 85 B, Arg 86 B, Phe 87 B |
| 2 | Red | -2751 | 171.00 | Pro 47 A, Ser 48 A, Val 49 A, Ile 50 A, Gly 80 A, Ser 82 A, Try 83A, Lys 84 A, Val 85 A, Arg 86 A, Phe 87 A, Thr 134 B, Ser 138, Ile 141 B |
| 3 | Blue | -1669 | 135.00 | Tyr 45 A, Leu 57 A, Lys 58 A, Asp 60 A, Glu 61 A, Lys 63 A, Thr 64 A, Asp 65 A, Thr 68 A , Thr 134 A , Ser 48 B, Ile 50 B, Arg 86 B, Gln 89 B, Phe 90 B, Glu 93 B |

*Phosphate probe: identifies binding sites of phosphorylated ligands

## 4.2.12. Small-angle X-ray scattering measurements of the Ov Ral-2[17-148]:

We further characterized the protein in solution using SAXS, and it behaves as a dimer in solution. The SAXS data indicate a radius of gyration ($R_g$) of 2.8 nm and a maximal dimension ($D_{max}$) of 9 nm. These values are a little bit larger than those calculated from the crystal structure ($R_g$ = 1.75 nm, $D_{max}$ = 7.3 nm), which is most likely caused by the flexible N-terminus that was not observed in the crystal structure. The molecular weight, based on the excluded volume of the *ab initio* model, is 34 kDa, which is very closely to that expected for a dimer (32.0 kDa). The monomeric molecular weight was calculated from the sequence (16 kDa). This model is in good agreement with the crystal structure. Finally, the crystal structure yields a good fit (chi$^2$ value = 1.3, figure 40) after correction of the absence of the flexible N-terminal part (14 aa).
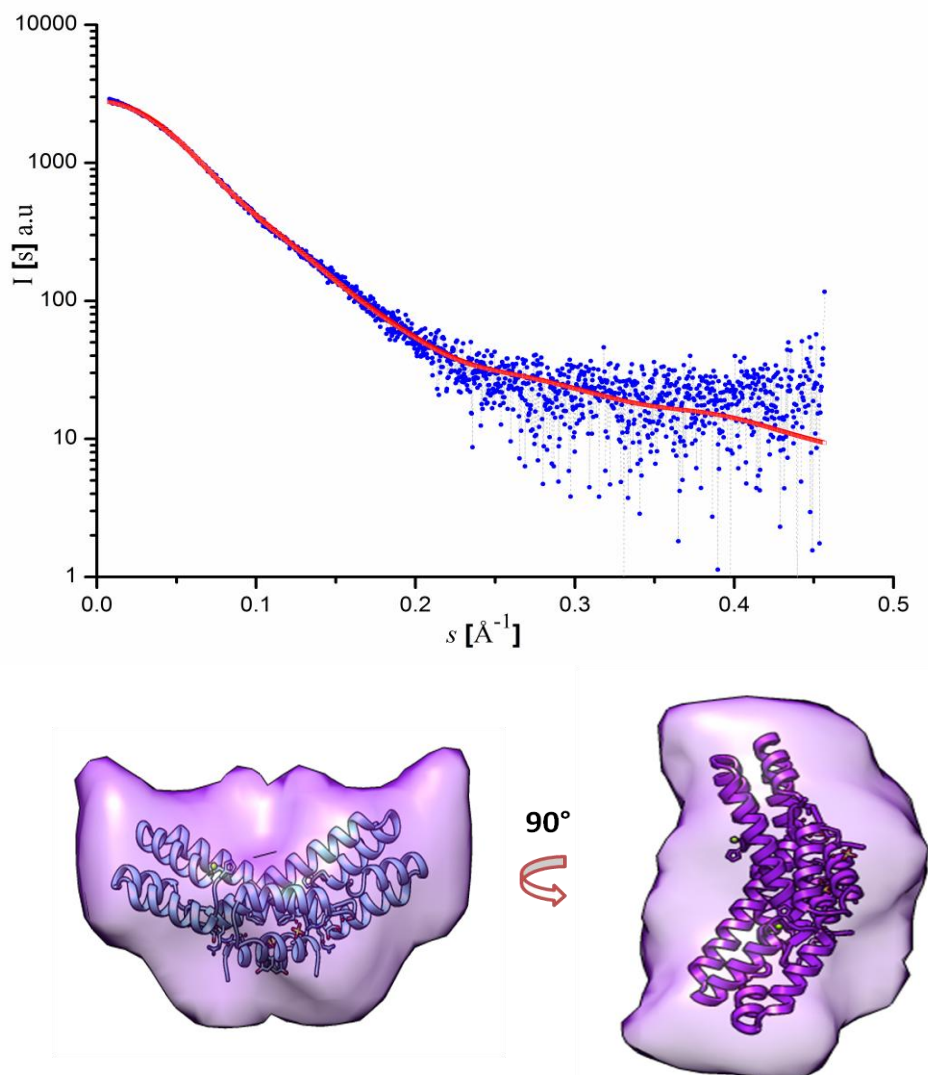
Fig.40: Processed solution scattering pattern and *ab initio* shape of dimeric Ov Ral-2[17-148], as obtained by SAXS. A) Experimental data are displayed in blue, with the theoretical scattering fit in red for the crystallographic data generated with CRYSOL (101) and the purple fit for the DAMMIF (99) *ab initio* model. B) Side and vertical view of the dimeric Ov Ral-2[17-148] *ab initio* model in blue, in superimposition with the crystal structure of Ov Ral-2[31-148] shown in a violet ribbon representation. The figure was generated with CHIMERA (142).

# 5 DISCUSSION

## 5. DISCUSSION:

### 5.1. Extracellular superoxide dismutase (OvEC-SOD):

**5.1.1. OvEC-SOD as a drug design target:**

Superoxide dismutases (SODs) are metalloproteins which are responsible for the dismutation of superoxides and are protective against the biological damage of cellular membranes and cell components during oxidative stress. The filarial parasite *O. volvulus* has a well-developed defense system which includes superoxide dismutases, to scavenge free radicals and combat oxidative stress during migration and sojourning of the microfilariae and adult worms in the human body. The search for new targets as well as drugs is essential due to the phenomenon of drug resistance. Although various approaches have been developed against onchocerciasis, ivermectin is still the drug of choice (143, 144). However, ivermectin is effective only against microfilariae (145). Furthermore, adverse effects and an emerging resistance have been reported due to an extensive treatment of ivermectin (66, 146). Hence, there is an urgent need for drugs that would also affect adult worms since they resume the production of microfilariae (147). Alternative treatments with medicinal plants have also been set in place using *Onchocerca ochengi* as a model (148, 149 and 150).

Recently, investigations have been focused on targeting the bacteria *Wolbachia bancrofti*, which co-exist in mutualistic symbiosis with *O. volvulus* (151, 152). Although these attempts have been successful to decrease the total number of infections, onchocerciasis has not been eliminated and persists to be a public health problem. Therefore, there is an increasing demand to develop effective drugs and vaccines. One approach is to identify excretory/secretory products (ESPs) that play an essential role in parasite/host interactions. Due to their extracellular accessibility, ESPs represent potential targets for the development of new control strategies (74, 151 and 153). The secretion of antioxidant enzymes by parasites is believed to protect the parasite against reactive oxygen species (ROS), produced by immune effector cells as a host defense mechanism to eradicate invading parasites (153). *O. volvulus* responds to these toxic oxidants by producing antioxidant enzymes such as thioredoxin peroxidase (154), glutathione S-transferase (155) and superoxide dismutase (SODs) (82, 156). Alignment studies of the primary structure and structural comparison of *O. volvulus* SOD (OvEC-SOD) revealed that the active site residues are highly conserved in comparison with the human cytosolic and extracellular counterparts, but the overall sequence identity of 45-55 % is relatively low. Structural differences can be expected and exploited for structure-based drug design approaches.

# 5 DISCUSSION

## 5.1.2. The N-terminal signal peptide:

The first superoxide dismutase (OvSOD1) was characterized by Henkle *et al.* (157). However SOD activity was also found in *in vitro* culture supernatants of microfilariae and adult worms. This finding presumed the existence of another secretory form of this enzyme and it was believed to be secreted and membrane-associated (158). Later on, another study reported about the identification of a secreted OvEC-SOD in larval and adult stages (159). The N-terminal signal peptide (SP) is eventually cleaved off by signal peptidases. The prediction of SPs and their cleavage sites in nematodal proteins is a great challenge i. It was first predicted that the N-terminal SP of OvECSOD is cleaved between Asn42 and Gly43 (158). Over the last decades, the computational tools for the detection of N-terminal signal peptides were greatly improved, but their abilities to predict the cleavage sites varied significantly and prediction differs from software to another. According to a review by Choo *et. al*. (160), using the SignalP program version 4.0, the cleavage site of the OvEC-SOD SP is between Leu 21 and Val 22. This is especially critical when making constructs for prokaryotic expression of recombinant ESPs for functional and structural studies.

## 5.1.3. OvEC-SOD cloning, expression and purification:

To obtain structural information for drug discovery approaches, the genes of the involved enzymes starting from Val22 were cloned and expressed as Step-tag fusion proteins with a TEV protease cleavage site. The protein is expressed in an insoluble form and purification reveals many impurities. The CD measurements showed high noises, while dynamic light scattering results indicate severe aggregation. Trials to crystallize the enzyme including the N-terminal signal peptide failed.

Therefore in the present study, I designed a construct starting from Gly43. The purification procedure took advantage of the affinity of the Strep-tag provided by the expression vector to isolate the protein by immobilized metal affinity chromatography. This successful purification procedure resulted in an average yield of 20 mg per four liters of culture.SDS-PAGE analysis revealed a pure protein with a molecular mass of about 19 kDa. The correctness of the folding state was examined by CD spectroscopy before crystallization trialsand revealed an overall composition  of helices (2.9 %), sheets (57.5 %), turns (14.2 %) and random coil regions (24.4 %) according to Yang`s reference (161).


## 5.1.4. Crystallization, data collection and structure determination:

The protein solution was analyzed by dynamic light scattering (DLS) over 15 days and showed a stable particle with an hydrodynamic radius of approximately 3.2 nm. A

# 5 DISCUSSION

concentrated protein solution was initially analyzed against a variety of available screening conditions. Low diffracting crystals appeared after one week in condition C6 from the ComPAS suite, which consisted of a 0.01 calcium chloride, 0.1 sodium acetate, pH 4.6 and 60 % (w/v) MPD in the reservoir. Tag removal was the key step to improve the crystal quality and to decrease the mosaicity. After 5 months, hexagonal shaped crystals were obtained in three conditions. The best crystals grew in a condition containing 0.1 M MES, pH 6.5, 45 % w/v PEG 2000 MME. Two data sets were collected from two different crystallization conditions; the first data set was collected at 293 K up to 2.3 Å resolution using an in-house Rigaku RU-200 rotating anode X-ray generator. The second data set was collected at the PETRA III beamline P13 (EMBL Hamburg) at DESY (Hamburg, Germany). The crystal structure was solved by molecular replacement with MOLREP using the crystal structure of superoxide dismutase from *C. elegans* (PDB entry 3kbe; sequence identity 60 %) as a homology model. The initial $R_{factor}$ for all data was 28 %. Subsequent refinement and model building of strands and surface loops, adjustment of side chains and torsion angles was performed resulting in an $R_{factor}$ of 15.9 % and a $R_{free}$ of 18.1 % using all data in the range between 25.8–1.55 Å. RMS deviations for bond lengths and angles are 0.009 Å and 1.240 Å, respectively. The Ramachandran plot (162) indicates that 98.3 % of the main-chain dihedral angles are located in the most favorable regions, and 1.7 % are found in additional allowed regions. The overall structure resembles a dimer, which was already described for the OvEC-SOD. The OvEC-SOD structure shows the expected features conserved throughout the organisms, including the Greek-key β-barrel motif. The active site includes a one catalytically active $Cu^{2+}$ ion and a one structural $Zn^{2+}$ ion. The coordination sphere of the $Cu^{2+}$ and $Zn^{2+}$ ions is defined by the invariant residues His46, His48, His63, His120 and His63, His71, His80, Asp83, respectively. The dimeric state of the OvEC-SOD in solution was also confirmed applying SAXS measurements. The MW of (38 ± 2 kDa) was estimated from forward scattering, which is the same than the calculated molecular weight for the dimer with an experimental Rg of 2.8 ±1 Å and Dmax of 9 ±1 Å.

### 5.1.5. Novel binding site and new inhibitors:

Before starting docking calculations, the potential sites for ligand binding from the 3D atomic coordinates of the OvCu/ZnSOD were determined. The active sites are normally hydrophobic pockets that involve side chain atoms and prosthetic groups that are tightly packed (135). Thus, the search for these sites on the molecule surface was carried out by filtering out sites with significant convex surfaces and the residues that were directly or indirectly involved in

# 5 DISCUSSION

metal binding. These residues are conserved among the Cu/Zn–SODs from all species so far examined, from bacteria to mammals; therefore to attain species selectivity, the cavities from these sites were also discarded. One site is the more prominent in size and located in a deep cavity close to the interface region away from the active site with an extended shape and rich in hydrophobic residues. The amino acids are not strictly conserved between OvCu/Zn–SOD and human Cu/Zn–SOD sequences; this drives this site to be an attractive target for designing species-specific inhibitors. A library of superoxide dismutases is already established for other SODs. It concludes a subset of uncharged polyaromatic compounds, its molecular weights range from 200 to 700 Da; some of them have selective *in vitro* anti-Cu/Zn activity, with $IC_{50}$ values in the micromolar range. In terms of docking studies, 50 compounds were tested, some inhibitors were selected depending on the higher score, interestingly; these compounds bind to human Cu/Zn–SOD in unspecific sites with a low docking score.

At first sight, it was challenging to inhibit OvCu/ZnSOD using classical 'active-site directed' approaches. The substrate-binding site is relatively very narrow and accommodate for just two atoms. Another problem in designing synthetic or natural selective inhibitors is the structural similarity and sequence identity to the other superoxide dismutase. The substrate is attracted to the active site of dismutase through a funnel shape channel. This channel is charged and starts as a shallow groove about 24 Å towards to a deeper tunnel that is about 10 Å wide and ends with a slit-like opening, which is not more than 4 Å close to the copper ion. The negatively charged small molecules are attracted away from the copper ion by the guanidinium group of Arg146, which is positioned approximately 5.11 Å away from the copper ion, while the Thr140 discards the large non substrate anions (163). However, some small anions such as azide, cyanide and fluoride do enter the active site cavity and competitively inhibit the enzyme, probably by irreversible binding with the copper ion (164). Nuclear magnetic resonance studies have shown that some anions such as chloride and phosphate can enter the active site tunnel, but do not bind directly to the copper. Instead, they remain associated with Arg146 in the active site channel in an "anion binding site" about 5 Å away from the copper ion (165, 166).

The principle of inhibiting any enzyme is that potential compounds bind directly or block the superoxide substrate to access the active site (167). Another alternative inhibiting strategy is to target the inhibitors towards the dimer interface to disorganize the quaternary structure. It is believed that the dimer interface in SODs is vital in the SOD reaction. For example a dimer interface mutation of human Mn-containing SOD leads to reduced steady-state catalytic constants (168). Dimer destabilizing mutations of Cu/Zn SOD have been related to many

# 5 DISCUSSION

degenerative motor neuron diseases (169). A parasite-selective inhibitor of triosephosphate isomerase from the *Trypanosoma cruzi* acts by blocking dimerization and disruption of the quaternary structure (170), so it is clear that the strategy could be useful and efficient, particularly for a protein that is in an oligomeric state in nature.

In a recent study it has been reported that the Cu/Zn-SOD enzyme from *T. solium* is almost entirely inhibited by 300 mM albendazole, whereas this compound has no effect on the bovine SOD (171). This could be interpreted by the hydrophobic nature of the albendazole, which binds more favorably with the hydrophobic amino acid residues at the entry of the electrostatic loop. These studies propose that the Cu/Zn-SOD could be selectively inhibited and can be exploited as a target protein for a rational drug design to control the river blindness.

Taking this into consideration, it would be desirable to perform more screening and to further check the $IC_{50}$ value of these lead molecules with recombinant OvEC-SOD. In addition, *in vitro* evaluation of the possible ability of these leads molecules to inhibit the host/parasite SODs is a great challenge. The undesirable interaction of the lead molecule with OvEC-SOD would decrease its activity *in vivo*, and this could result in oxidative stress and death of the parasite. In conclusion, this study shows that these compounds could be quite useful inhibitors.

## 5.2. Immunodominant hypodermal antigen (Ov Ral-2):

### 5.2.1. Ov Ral-2 as a novel nematode specific antigen:

There is an increasing demand to gain detailed information about proteins involved in pathogenesis and to emerge the diagnostic tools to tackle this worldwide health problem. Ov Ral-2, an immunodominant hypodermal antigen, belongs to the SXP family. This active protein is found exclusively in the hypodermal region in the adult parasite and uterine region in microfilaria. Although SXP has been studied in many nematodes, its structure and function are still unknown. Studies in *B. malayi* indicated a rapid turnover of SXP in the parasite, shown by *in vitro* labeling with [$S^{35}$] methionine (172). This means these proteins play an active role in the parasite life cycle. Earlier studies have found the SXP/Ral-2 proteins to be localized in the filarial hypodermis (172) or as secreted proteins (173, 47). SXP/Ral-2 proteins were exploited as antigens for the development of diagnostic serological assays (174, 175 and 176). In a recent study, immunization of mice with *E. coli* expressed Ov Ral-2, induced a statistically significant 64 % level of host protection and 39 % reduction in the larval survival.

# 5 DISCUSSION

Whereas, the immunized mice with the *P. pastoris* expressed Ov Ral-2 induced a 55 % level of host protection and a 24 % reducing of parasite survival (177).

One feature of the Ov Ral-2 protein sequence is the presence of a putative leader segment followed by a poly-Q region; other known homologs from other nematodes do not possess this poly-Q region. The antibody reactivity against recombinant proteins containing and missing the poly-Q region showed that this segment does not play a significant role as a B cell epitope (172). However, the poly-Q region is a common feature of many human proteins, it may play a crucial role in the cellular functions, like the role in stabilizing protein-protein interactions (178).

## 5.2.2. Cloning, expression and purification of the full Ov Ral-2 and Ov Ral-2$^{17-148}$:

Since the trials to produce good diffracting crystals from full-length Ov-Ral-2 construct were unsuccessful, the full-length constructwas subcloned to exclude the poly-Q region. Reasons might be that either this poly-Q can form a highly flexible loop, or this region was a structurally disordered part of the protein. The ORF of the sequence encoding the Ov Ral-2$^{17-148}$ was amplified by the polymerase chain reaction. The *Escherichia coli* strain BL21 (DE3) competent cells were transformed with the recombinant plasmid. Protein expression was induced by adding isopropyl β-d-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. The cells were collected by centrifugation and resuspended in lysis buffer. The supernatant was loaded onto a HisTrap FF column (GE Healthcare, USA). The bound protein was eluted using an elution buffer consisting of 500 mM imidazole. The SeMet labeled protein was expressed in the methionine auxotrophic *E. coli* strain B834 (DE3). The cells were grown in the M9 minimal media containing 30 mg/l L-SeMet and purified as described for the native protein, except that 2 mM TCEP were added to the buffer after the eluting step from the Ni-NTA affinity column to prevent oxidation of SeMet protein. The molecular weight of the purified protein was analyzed by native gel electrophoresis, Superdex 200 calibration curve and was judged to be at least 95 % pure by SDS–PAGE. The sequence integrity was experimentally determined by mass spectroscopy de. The peptide masses were in good agreement with the theoretical molecular weights calculated from the sequences. Analytical size exclusion chromatography and SAXS were used to confirm that the designed proteins exhibited the correct quaternary structures in solution. The full protein was shown to be a dimer in solution. As expected, Ov Ral-2$^{17-148}$ eluted slightly later, as expected due to its marginally smaller molecular weight. The spectroscopic properties of the full Ov Ral-2 and Ov Ral-2$^{17-148}$ were compared using CD, the results showed high coils percentage in full Ov-

# 5 DISCUSSION

Ral-2. Before setting crystallization trials, the pure solution was examined by dynamic light scattering over a suitable period. A hydrodynamic radius of approximately 3.1 nm was observed, while the folding state of the purified Ov Ral-2$^{17-148}$ was checked applying circular dichroism (CD) spectroscopy. The far-UV CD spectra of both full and Ov Ral-2$^{17-148}$ proteins showed α helical profile with double minimum negative peaks at 208 nm and 222 nm along with the maxima near 190 nm. The $MRE_{222}/MRE_{208}$ ratio of 1.084 indicated the coiled coil (179).

## 5.2.3. The Ov Ral-2$^{17-148}$ overall structure:

The Ov Ral-2$^{17-148}$ structure is composed of two identical subunits. The Ov Ral-2$^{17-148}$ dimer is related by a non-crystallographic two-fold axis that is oriented as interlocking 2-ring architecture. Each monomer is formed of 6 α helical structures. The first 14 amino acids of Ov Ral-2$^{17-148}$ were disordered. This disordered part did not appear in the crystal structure and showed no electron density. To verify that the part was still intact in the crystallized protein, I subjected washed crystals of to non reducing SDS-PAGE alongside solutions of the purified proteins. I did not observe cleavage products of the proteins in any lane of the gel, confirming that the crystal structure of Ov Ral-2$^{17-148}$ is intact. I have no accurate data indicating how this knotted dimer is formed, but it provides interesting insights into the mechanisms of subunits assembly and protein topology. Ov Ral-2$^{17-148}$ may be expressed in an open, monomeric form and may pass through transient state under the effect of both hydrophobic and electrostatic forces, resulting in a stable dimer with knotted conformation. The crystal structure harbors many hydrophobic patches on the surface of the protein. Furthermore, the instability index (II) of this protein is estimated to be 57.3; this classifies this protein as unstable (180). It is known that protein stability is necessary for the understanding of the fundamental of thermodynamics in the folding process. The OvRal-2$^{17-148}$ dimer is maintained by four salt bridges, formed between residues Glu76, Glu57, Arg82 and Arg87 of chain A, corresponding to Arg82, Arg87, Glu76 and Glu57 of chain B. The distances are 3.98, 3.13, 3.13 and 3.98 Å, respectively. The dimer is also stabilized by 14 hydrogen bonds and 171 non bonded contacts (Van der Waals interactions) belonging to 43 residues.

Ov Ral-2$^{17-148}$ was further characterized in solution using SAXS and it behaves as a dimer in solution. The SAXS data show a radius of gyration (Rg) of 2.8 nm and a maximal dimension (Dmax) of 9 nm. The molecular weight, based on the excluded volume of the *ab initio* model, is 34 kDa, which is very closely to that expected for a dimer (32.0 kDa). This *ab initio* model fits well with the crystal structure, after considering the absence of the missing 14 N-terminal residues.

# 5 DISCUSSION

## 5.2.4. Insights into protein topology and mechanisms of subunits assembly:

After solving the protein structure, the structure showed unusual topology, which can be described by two identical rings interacting together like Olympic rings or a knot-like structure. Scientists from different disciplines have increasingly recognized the prevalence of molecular knots. A knotted topology could increase the ligand binding affinity or even increase the catalytic activity through enhancing the stability (kinetic, thermodynamic and mechanical) of the protein, thus giving it more resistance to degradation pathways and cellular translocation (181,182 and 183). The mechanism of the folding process that leads to this unusual topology is still unknown. Nowadays it is widely believed that hydrophobicity is the main force of protein folding. These forces may influence not only the surface parts of the structure, but also the central core (184, 185). Till now, more than 750 knotted proteins have been deposited in the PDB, equal to approximately 1 % of all entries (186). Nevertheless the inter-molecular non-covalent interactions can result in an interlocked topology, where the two chains can form an inseparable oligomeric structure (187). The hydrophobic influence is thought to be the significant driving force for the folding of globular proteins (188). The crystal structure of bovine mitochondrial peroxiredoxin III (PDB code: 1ZYE), that was solved recently, is the clearest example. In addition, there is a linear relationship between the peripheral parts of hydrophobic amino acids and the changes of free energy of the amino acids resulted from movement away from water or organic solvent (189). The high-resolution structure of Ov Ral-2 revealed a novel protein topology with greater backbone flexibility at the N-terminal region. This flexibility might allow Ov Ral-2 to bind its ligands or receptor(s). Folded proteins are highly dynamic; structural rearrangements are frequent. In most instances the structural rearrangements are minor and restricted to loop regions or some particular secondary structure element (strand or helix) that switches between ordered/disordered states (for example, upon ligand binding). Nevertheless, sometimes major structural rearrangements are detected.

## 4.2.5 Impact of knotted topology on the protein properties:

Knotted structures present a state of challenge to who interested in protein folding, and have developed their theory based on the simple folding kinetics (190). For such proteins, it has been proposed that the folding energy outlook looks like a funnel (191), implying that the native state can be reached by moving from high toward the lower energy from any point of denatured confirmations. Substantial and recent research has been carried out on knotted proteins, both experimentally and computationally. The researchers investigate how this

# 5 DISCUSSION

unusual topology occurs during folding. Jackson *et al*. have performed a series of experiments to characterize the folding pathways of two structurally related, knotted methyltransferases (192). Work on the methyltransferase model system led to the proposal that knots can occur in the early stage of folding (193). On the other hand, computational simulations propose the various scenarios for knotting mechanisms (194).

Other important questions related to the existence of knotted proteins have not yet been fully addressed experimentally. For example, does a knotted topology have any impact on the protein stability, folding, or even rigidity? The difficulty in answering such a question lies in a lack of suitable control models. And based on the analysis of specific knotted proteins, it has been suggested that knots could play a important role in protecting the protein against degradationby the proteasome pore (195), furnish structural stability in transporter proteins (196), support the structural rigidity of the native state (197), shape and form the binding site of some enzymes (198), enhance thermal and mechanical stability (199), or even enzymatic activity (197). Furthermore, kinetic stability increases with the complexity of knot type.Kinetic stability is a property of the native state, that is necessary to maintain the biological function of the protein during a physiologically conditions (200). However, in the majority of cases, it was not possible to detect the structural and functional advantages of knotted folds (193). On the other side, several mutational studies on SpoU- and TrmD-like proteins have illuminated the functional role of the knotted region in their structure, and have shown that it forms the cofactor-binding pocket and/or the active site (201, 202). Despite the topological difficulties, some knotted structures can fold quickly and cooperatively enough to prevent any deleterious misfolding or aggregation events.

## 4.2.6. Insight into structure-function relationship:

Here the first crystal structure of the Ov Ral-2$^{17-148}$ is presented. Ov Ral-2 has been used to develop a successful vaccine candidate with a potential clinical application. Ov Ral-2 reveals a very low sequence homology when compared with human proteins. The protein consists of an N-terminal polyglutamine region and a conserved domain of unknown function (DUF148-based on Conserved Domain Architecture Retrieval Tool, NCBI), falling into the Pfam 02520, that is restricted only to nematodes. Several keratin-like proteins bear this DUF148 domainwithin various nematodes. Therefore, Ov Ral-2 might be assigned as keratin- like proteins, which might be directly or indirectly involved in the worm development (203). However, RNAi studies with gene orthologue C06A8.3 or ZK970.7 that tends to pair up with T23G11.3 (GLD-1), an RNA binding protein regulating germ line development did not result

# 5 DISCUSSION

in any lethal effect (204). In another study, involved in the characterization of Mi-SXP-1, a homolog from plant parasitic root-knot nematode *Meloidogyne incognita*, the scientists excluded any role in induction of the feeing cells (205).

Ov Ral-2 reveals a unique distribution of aromatic amino acids at the surface and this aromatic maid be relevant for the stability and protein folding. The shielding of the hydrophobic side chains against the solvent results in increased entropy, These residues have therefore important structural roles, both through the hydrophobic effect and by permitting stacking interactions with other aromatic moieties in other molecules bound to the protein.

The positive patches on the surface of the protein consist of arginine and lysine side chains, which can interact with the negatively charged phosphate groups on membranes or other protein surfaces. Also, they can bind to metals and other counter ions, as well as to other proteins and cell structures or to histidines.

Up to date, the structural basis and the function of the poly-Q region is not clear and the precise structure of poly-Q itself is unknown (206). The suggested structural conformation of either synthetic poly-Q or naturally occurring proteins with poly-Q region involve helices, random coils or extended loops (207, 208). Statistical studies showed that many proteins possess a poly-Q region are functioning in the transcriptional regulation and are therefore related to nuclear mechanisms (209). Nevertheless, the different localization and excretory properties of Ov Ral-2 plus absence of DNA binding motifs in the Ov Ral-2 ORF suggest that Ov Ral-2 function are more related to the protein-protein interactions. Some experimental evidence hints to the fact that poly-Q region play a role in signal transduction or activation of gene transcription (210).

More functional studies on the Ov Ral-2 protein need to be performed in order to understand the real function and its pathway within the human body. This requires knock out experiments to examine the variation in functionality after removing the entire gene or particular polyglutamine stretches.

# 6 SUMMARY AND ZUSAMMENFASSUNG

## SUMMARY:

Approximately 18 million people around the world (99 % living in Africa) are infected with onchocerciasis and 270,000 become blind, 500,000 visually impaired upon infection. *O.volvulus*, a filarial parasite, is a master regulator of host immune responses, developing complex mechanisms to depress host protective Th2-type responses and promotes long-term persistence. Although various treatments have been developed against onchocerciasis, ivermectin turned out to be most effective choice. However, ivermectin is only effective against microfilariae and adverse effects, as well as an emerging resistance has been reported due to extensive treatment with ivermectin. Helminth-derived molecules can be used as tools to identify the underlying mechanisms of immune regulation or to determine new anti-inflammatory and filaricide therapeutics. One approach is to identify the excretory/secretory products (ESPs) that play an essential role in parasite/host interaction. Due to their extracellular accessibility, ESPs represent suitable targets for the development of new control strategies. Inactivation of these products may help to weaken the defense mechanisms of the parasite and may enable the host to eliminate the parasite. Among many important possible drug targets or vaccine candidates, two proteins, known as extra cellular superoxide dismutase enzyme (OvEC-SOD) and the immunodominant hypodermal antigen 2 (Ral-2), particularly attract attention and are structurally characterized within this study. The genes encoding for OvEC-SOD and the Ov Ral-2 were successfully cloned and expressed in *E. coli*. Both proteins could be crystallized and structural models were obtained applying X-ray diffraction. The oligomeric state, observed in the crystal structures, were confirmed in solution by small-angle X-ray scattering (SAXS). For prodrug development, a subset of uncharged poly aromatic compounds, with molecular weights in the range of 200–700 Da, have been tested *in vitro* for anti Cu/Zn activity. Three inhibitors were selected upon their high binding energy. Interestingly, these compounds bind human Cu/Zn SOD unspecifically with low docking scores. Although the function of Ov Ral-2 is unknown, it is considered as one of the most promising vaccine candidates. The structure-function relationship of Ov Ral-2 was investigated. The structural model showed unusual topology, which can be described as two identical rings interlocking in a knot-like structure. The interesting distribution of charges and aromatic amino acids elucidate the possible role of this protein- protein interaction. This work contributes towards a better understanding of the immunotherapeutic strategy and can support the development of a novel therapeutic agent or a new vaccine and may support new diagnostic tests.

# 6 SUMMARY AND ZUSAMMENFASSUNG

## ZUSAMMENFASSUNG:

Weltweit sind über 18 Millionen Menschen mit dem Erreger der Flussblindheit *Onchocerca volvulus* infiziert, wobei 270.000 Menschen aufgrund der Erkrankung erblinden und bei weiteren 500.000 Menschen die Sehkraft eingeschränkt ist. *O.volvulus*, ein Filarien Parasit, wirkt als ein Master-Regulator der Wirtsimmunantwort, indem er die Entwicklung von komplexen Mechanismen zur Wirtsschutz-Reaktionen des Th2-Typs unterdrückt und damit seine Persistenz fördert. Obwohl verschiedene therapeutische Ansätze gegen Onchozerkose entwickelt wurden, hat sich insbesondere Ivermectin als am wirkungsvollsten herausgestellt. Jedoch ist Ivermectin nur gegen Mikrofilarien wirksam und ferner wurde von Nebenwirkungen und Resistenzentwicklungen durch umfangreiche Behandlung mit Ivermectin berichtet. In einem neuen, alternativen Ansatz werden Helminth-abgeleiteten Molekülen verwendet, um die zugrunde liegenden Mechanismen der Immunregulation zu identifizieren, oder um neue entzündungshemmende und filaricide Therapeutika zu entwickeln. Dabei werden die exkretorischen/sekretorischen Produkte (ESPs), die eine wesentliche Rolle bei Parasit/Wirt-Interaktion spielen, identifiziert. Aufgrund ihrer extrazellulären Zugänglichkeit stellen ESPs offensichtliche Ziele für die Entwicklung neuer Kontrollstrategien dar. Die Inaktivierung dieser Produkte kann helfen, die Abwehrmechanismen des Parasiten zu schwächen und es dadurch zu eliminieren. Unter vielen wichtigen Drogenzielen oder Impfstoffkandidaten stechen zwei Proteine besonders heraus, das Enzym Superoxid Dismutase (SOD) und das immun hypodermale Antigen 2 (Ov Ral-2). Diese sollten im Rahmen dieser Arbeit strukturell charakterisiert werden.

Die Gene die für die Proteine OvEC-SOD und Ov Ral-2 kodieren wurden kloniert und erfolgreich in *E. coli* exprimiert. Beide Proteine konnten kristallisiert werden und mit Hilfe der Röntgenstrukturanalyse konnte ein Strukturmodell erstellt werden. Der in der Kristallstruktur beobachtete oligomere Zustand wurde in Lösung durch Kleinwinkelröntgenstreuung (SAXS) bestätigt. Im Rahmen der Prodrug-Entwicklung, wurde eine Untergruppe ungeladener polyaromatischer Verbindungen mit Molekulargewichten im Bereich von 200-700 Da *in vitro* auf ihre Anti Cu/Zn Aktivität hin untersucht, wobei $IC_{50}$ Werte im mikromolaren Bereich gemessen wurden. Drei Inhibitoren wurden aufgrund ihrer hohen Bindungsenergie ausgewählt. Aufgrund von Docking-Studien wurde identifiziert, dass diese Verbindungen menschliche Cu/Zn-SOD unspezifisch binden. Obwohl die Funktion von Ov Ral-2 unbekannt ist, gilt es als einer der vielversprechendsten Impfstoffkandidaten. Aufgrund dessen wurde die Struktur-Funktions-Beziehungen von Ov Ral-2 untersucht. Die Struktur weisst eine ungewöhnliche Topologie auf, die zwei identische Ringe in einer

# 6 SUMMARY AND ZUSAMMENFASSUNG

knotenartigen Struktur darstellt. Die Protein-Protein-Interaktion kann aufgrund der interessanten Verteilung der elektrostatischen Ladungen und aromatischen Aminosäuren erklärt werden. Diese Arbeit trägt zu einem besseren Verständnis der immuntherapeutischen Strategie und der Herstellung von neuen therapeutischen Wirkstoffen bei und leistet einen wesentlichen Beitrag für die Entwicklung neuer diagnostischer Tests.

# 7 REFERENCES

## References:

1. Anderson, R. M. (1994). Mathematical studies of parasitic infection and immunity. *Science*, *264*(5167), 1884-1886.

2. Ottesen, E. A., Duke, B. O., Karam, M., & Behbehani, K. (1997). Strategies and tools for the control/elimination of lymphatic filariasis. *Bulletin of the World Health Organization*, *75*(6), 491.

3. Gardon, J., Gardon-Wendel, N., Kamgno, J., Chippaux, J. P., & Boussinesq, M. (1997). Serious reactions after mass treatment of onchocerciasis with ivermectin in an area endemic for Loa loa infection. *The Lancet*, *350*(9070), 18-22.

4. WHO (1995). Onchocerciasis and its control. Report of a WHO Expert Committee on Onchocerciasis Control. *World Health Organ Tech Rep Ser*, 852, 1-104.

5. WHO (1976). Epidemiology of onchocerciasis. Report of a WHO Expert Committee. *World Health Organ Tech Rep Ser*, 1-94.

6. WHO (1987). Protective immunity and vaccination in onchocerciasis and lymphatic filariasis. Report of the 13th Scientific working group on filariasis. *Geneva: WHO document TDR/FIL/SWG*, 13, 1-167.

7. Amazigo, U. O. (1993). Detrimental effects of onchocerciasis on marriage age and breast-feeding. *Tropical and geographical Medicine*, *46*(5), 322-325.

8. Brieger, W. R., Awedoba, A. K., Eneanya, C. I., Hagan, M., Ogbuagu, K. F., Okello, D. O., & Remme, J. H. (1998). The effects of ivermectin on onchocercal skin disease and severe itching: results of a multicentre trial. *Tropical medicine & international health: TM & IH*, *3*(12), 951-961.

9. Vlassoff, C., Weiss, M., Ovuga, E. B. L., Eneanya, C., Nwel, P. T., Babalola, S. S., & Shetabi, P. (2000). Gender and the stigma of onchocercal skin disease in Africa. *Social science & medicine*, *50*(10), 1353-1368.

10. Crosskey, R. W. (1969). A re-classification of the Simuliidae (Diptera) of Africa and its islands. *Bull. Br. Mus. Natur. Hist.(Entomol.)*, *14*, 1-195.

11. Duke, B. O. (1993). The population dynamics of Onchocerca volvulus in the human host. *Tropical medicine and parasitology: official organ of Deutsche Tropenmedizinische Gesellschaft and of Deutsche Gesellschaft fur Technische Zusammenarbeit (GTZ)*, *44*(2), 61-68.

12. Akogun, O. B., & Tembo, K. (1996). Microfilaruria in an area of Nigeria with hyperendemic onchocerciasis. *Zentralblatt für Bakteriologie*, *285*(1), 86-91.

13. Duke, B. O. (1968). Studies on factors influencing the transmisson of onchocerciasis. IV. The biting-cycles, infective biting density and transmission potential of" forest" Stimulium dannosum. *Annals of tropical medicine and parasitology*, *62*(1), 95-106.

14. Combes, C. (2005). *The art of being a parasite*. University of Chicago Press.

15. Bandi, C., Trees, A. J., & Brattig, N. W. (2001). Wolbachia in filarial nematodes: evolutionary aspects and implications for the pathogenesis and treatment of filarial diseases. *Veterinary parasitology*, *98*(1), 215-238.

16. CDC, Laboratory Identification of Parasites of Public Health Concern: Filariasis. http://www.dpd.cdc.gov/dpdx/HTML/Filariasis.htm.

17. Maizels André, A. V., Blackwell, N. M., Hall, L. R., Hoerauf, A., Brattig, N. W., Volkmann, L., & Pearlman, E. (2002). The role of endosymbiotic Wolbachia bacteria in the pathogenesis of river blindness. *Science*, *295*(5561), 1892-1895.

# 7 REFERENCES

18. Taylor, M. J., Bandi, C., & Hoerauf, A. (2005). Wolbachia. Bacterial Endosymbionts of Filarial Nematodes. *Advances in parasitology*, *60*, 245-284.

19. Duke, B. O. (1991). Observations and reflections on the immature stages of Onchocerca volvulus in the human host. *Annals of tropical Medicine and Parasitology*, *85*(1), 103-110.

20. Kluxen, G., & Hörauf, A. (2007). Okuläre Onchozerkose: Wolbachien haben eine Schlüsselrolle. *Der Ophthalmologe*, *104*(10), 860-865.

21. Tamarozzi, F., Halliday, A., Gentil, K., Hoerauf, A., Pearlman, E., & Taylor, M. J. (2011). Onchocerciasis: the role of Wolbachia bacterial endosymbionts in parasite biology, disease pathogenesis, and treatment. *Clinical microbiology reviews*, *24*(3), 459-468.

22. Hoerauf, A., Specht, S., Büttner, M., Pfarr, K., Mand, S., Fimmers, R., & Brattig, N. (2008). Wolbachia endobacteria depletion by doxycycline as antifilarial therapy has macrofilaricidal activity in onchocerciasis: a randomized placebo-controlled study. *Medical microbiology and immunology*, *197*(3), 295-311.

23. Saint André, A. V., Blackwell, N. M., Hall, L. R., Hoerauf, A., Brattig, N. W., Volkmann, & Diaconu, E. (2002). The role of endosymbiotic Wolbachia bacteria in the pathogenesis of river blindness. *Science*, *295*(5561), 1892-1895.

24. Anderson, J., & Fuglsang, H. (1977). Ocular onchocerciasis. *Tropical diseases bulletin*, *74*(4), 257-272.

25. Gillette-Ferguson, I., Hise, A. G., McGarry, H. F., Turner, J., Esposito, A., Sun, Y., & Pearlman, E. (2004). Wolbachia-induced neutrophil activation in a mouse model of ocular onchocerciasis (river blindness). *Infection and immunity*, *72*(10), 5687-5692.

26. Abraham, D., Lucius, R., & Trees, A. J. (2002). Immunity to Onchocerca spp. in animal hosts. *Trends in parasitology*, *18*(4), 164-171.

27. Nanduri, J., & Kazura, J. W. (1989). Paramyosin-enhanced clearance of Brugia malayi microfilaremia in mice. *The Journal of Immunology*, *143*(10), 3359-3363.

28. Soboslay, P. T., Dreweck, C. M., Taylor, H. R., Brotman, B., Wenk, P., & Greene, B. M. (1991). Experimental onchocerciasis in chimpanzees. Cell-mediated immune responses, and production and effects of IL-1 and IL-2 with Onchocerca volvulus infection. *The Journal of Immunology*, *147*(1), 346-353.

29. Elson, L. H., Calvopina, M., Araujo, E., Bradley, J. E., Guderian, R. H., & Nutman, T. B. (1995). Immunity to onchocerciasis: putative immune persons produce a Thl-like response to Onchocerca volvulus. *Journal of Infectious Diseases*, *171*(3), 652-658.

30. Nutman, T. B., Steel, C., Ward, D. J., Zea-Flores, G., & Ottesen, E. A. (1991). Immunity to onchocerciasis: recognition of larval antigens by humans putatively immune to Onchocerca volvulus infection. *Journal of Infectious Diseases*, *163*(5), 1128-1133.

31. Steel, C., Lujan-Trangay, A., Gonzalez-Peralta, C., Zea-Flores, G., & Nutman, T. B. (1991). Immunologic responses to repeated ivermectin treatment in patients with onchocerciasis. *Journal of Infectious Diseases*, *164*(3), 581-587.

32. Elson, L. H., Calvopina, M., Araujo, E., Bradley, J. E., Guderian, R. H., & Nutman, T. B. (1995). Immunity to onchocerciasis: putative immune persons produce a Thl-like response to Onchocerca volvulus. *Journal of Infectious Diseases*, *171*(3), 652-658.

33. Ottesen, E. A. (1995). Immune responsiveness and the pathogenesis of human onchocerciasis. *Journal of Infectious Diseases*, *171*(3), 659-671.

34. Brattig, N. W., Lepping, B., Timmann, C., Büttner, D. W., Marfo, Y., Hamelmann, C., & Horstmann, R. D. (2002). Onchocerca volvulus-exposed persons fail to produce interferon-γ in response to O. volvulus

antigen but mount proliferative responses with interleukin-5 and IL-13 production that decrease with increasing microfilarial density. *Journal of Infectious Diseases*, *185*(8), 1148-1154.

35. Hogarth, P. J., Taylor, M. J., & Bianco, A. E. (1998). IL-5-dependent immunity to microfilariae is independent of IL-4 in a mouse model of onchocerciasis. *The Journal of Immunology*, *160*(11), 5436-5440.

36. Maizels, R. M. (2005). Infections and allergy—helminths, hygiene and host immune regulation. *Current opinion in immunology*, 17(6), 656-661.

37. Boatin, B. (2008). The onchocerciasis control programme in West Africa (OCP). *Annals of Tropical Medicine & Parasitology*, *102*(Supplement-1), 13-17.

38. Kaifi, J. T., Diaconu, E., & Pearlman, E. (2001). Distinct roles for PECAM-1, ICAM-1, and VCAM-1 in recruitment of neutrophils and eosinophils to the cornea in ocular onchocerciasis (river blindness). *The Journal of Immunology*, *166*(11), 6795-6801.

39. Maizels, R. M., & Yazdanbakhsh, M. (2003). Immune regulation by helminth parasites: cellular and molecular mechanisms. *Nature Reviews Immunology*, *3*(9), 733-744.

40. Bartlett, A., Turk, J., Ngu, J., Mackenzie, C. D., Fuglsang, H., & Anderson, J. (1978). Variation in delayed hypersensitivity in onchocerciasis. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *72*(4), 372-377.

41. Ngu, J. L., Tume, C., Lando, G., Ndumbe, P., Leke, R. G., Titanji, V., & Asonganyi, T. (1989). Comparative studies of clinical groups of patients in an onchocerciasis endemic area for evidence of immune-mediated protection. *Tropical medicine and parasitology: official organ of Deutsche Tropenmedizinische Gesellschaft and of Deutsche Gesellschaft fur Technische Zusammenarbeit (GTZ)*, *40*(4), 460-463.

42. Piessens, W. F., & Mackenzie, C. D. (1982). Immunology of lymphatic filariasis and onchocerciasis. *Immunology of parasitic infections*, *2*, 622-653.

43. Buck, A. A., Anderson, R. I., Kawata, K., & Hitchcock, J. C. (1969). Onchocerciasis: Some New Epidemiologic and Clinical Findings Results of an Epidemiologic Study in the Republic of Chad. *The American journal of tropical medicine and hygiene*, *18*(2), 217-230.

44. Prost, A., Nebout, M., & Rougemont, A. (1979). Lepromatous leprosy and onchocerciasis. *Br Med J*, *1*(6163), 589-590.

45. Greene, B. M., Fanning, M. M., & Ellner, J. J. (1983). Non-specific suppression of antigen-induced lymphocyte blastogenesis in Onchocerca volvulus infection in man. *Clinical and experimental immunology*, *52*(2), 259.

46. Philipp, M., Parkhouse, R. M. E., & Ogilvie, B. M. (1980). Changing proteins on the surface of a parasitic nematode.

47. Allen, J. E., & Maizels, R. M. (2011). Diversity and dialogue in immunity to helminths. *Nature Reviews Immunology*, *11*(6), 375-388.

48. McSorley, H. J., & Maizels, R. M. (2012). Helminth infections and host immune regulation. *Clinical microbiology reviews*, *25*(4), 585-608.

49. van Riet, E., Hartgers, F. C., & Yazdanbakhsh, M. (2007). Chronic helminth infections induce immunomodulation: consequences and mechanisms. *Immunobiology*, *212*(6), 475-490.

50. Harnett, W., & Harnett, M. M. (2010). Helminth-derived immunomodulators: can understanding the worm produce the pill?. *Nature Reviews Immunology*, *10*(4), 278-284.

51. Rajamanickam, A., & Babu, S. (2013). Immunomodulation by filarial parasites. *International Trends in Immunity*, *1*(4).

# 7 REFERENCES

52. Klotz, C., Ziegler, T., Daniłowicz-Luebert, E., & Hartmann, S. (2011). Cystatins of parasitic organisms. In *Cysteine Proteases of Pathogenic Organisms* (pp. 208-221). Springer US.

53. Manoury, B., Gregory, W. F., Maizels, R. M., & Watts, C. (2001). Bm-CPI-2, a cystatin homolog secreted by the filarial parasite Brugia malayi, inhibits class II MHC-restricted antigen processing. *Current Biology*, *11*(6), 447-451.

54. Schönemeyer, A., Lucius, R., Sonnenburg, B., Brattig, N., Sabat, R., Schilling, K., & Hartmann, S. (2001). Modulation of human T cell responses and macrophage functions by onchocystatin, a secreted protein of the filarial nematode Onchocerca volvulus. *The Journal of Immunology*, *167*(6), 3207-3215.

55. Muller, F. L., Lustgarten, M. S., Jang, Y., Richardson, A., & Van Remmen, H. (2007). Trends in oxidative aging theories. *Free Radical Biology and Medicine*, *43*(4), 477-503.

56. Awadzi, K. (1980). The chemotherapy of onchocerciasis II. Quantitation of the clinical reaction to microfilaricides. *Annals of tropical medicine and parasitology*, *74*(2), 189-197.

57. Awadzi, K. (2003). Clinical picture and outcome of Serious Adverse Events in the treatment of Onchocerciasis. *Filaria Journal*, *2*(Suppl 1), S6.

58. Bird, A. C., el-Sheikh, H. A. D. I., Anderson, J., & Fuglsang, H. (1980). Changes in visual function and in the posterior segment of the eye during treatment of onchocerciasis with diethylcarbamazine citrate. *British journal of ophthalmology*, *64*(3), 191-200.

59. Campbell, W. C. (1985). Ivermectin: an update. *Parasitology Today*, *1*(1), 10-16.

60. Aziz, M., Diop, I., Diallo, S., Lariviere, M., & Porta, M. (1982). EFFICACY AND TOLERANCE OF IVERMECTININ HUMAN ONCHOCERCIASIS. *The Lancet*, *320*(8291), 171-173.

61. Baraka, O. Z., Khier, M. M., Ahmed, K. M., Ali, M. M., El Mardi, A. E., Mahmoud, B. M., & Williams, J. F. (1995). Community based distribution of ivermectin in eastern Sudan: acceptability and early post-treatment reactions. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *89*(3), 316-318.

62. Baraka, O. Z., Mahmoud, B. M., Ali, M. M., Ali, M. H., El Sheikh, E. A., Homeida, M. M., & Williams, J. F. (1995). Ivermectin treatment in severe asymmetric reactive onchodermatitis (sowda) in Sudan. *Transactions of the royal Society of Tropical Medicine and Hygiene*, *89*(3), 312-315.

63. Amazigo, U. V., Nnoruka, E., Maduka, C., Bump, J., Benton, B., & Sékétéli, A. (2004). Ivermectin improves the skin condition and self-esteem of females with onchocerciasis: a report of two cases. *Annals of tropical medicine and parasitology*, *98*(5), 533-537.

64. Pacqué, M., Elmets, C., Dukuly, Z. D., Muñoz, B., White, A. T., Taylor, H. R., & Greene, B. M. (1991). Improvement in severe onchocercal skin disease after a single dose of ivermectin. *The American journal of medicine*, *90*(5), 590-594.

65. Trees, A. J., Graham, S. P., Renz, A., Bianco, A. E., & Tanya, V. (2000). Onchocerca ochengi infections in cattle as a model for human onchocerciasis: recent developments. *Parasitology*, *120*(07), 133-142.

66. Keiser, P. B., Reynolds, S. M., Awadzi, K., Ottesen, E. A., Taylor, M. J., & Nutman, T. B. (2002). Bacterial endosymbionts of Onchocerca volvulus in the pathogenesis of posttreatment reactions. *Journal of Infectious Diseases*, *185*(6), 805-811.

67. Chavasse, D. C., Post, R. J., Davies, J. B., & Whitworth, J. A. (1993). Absence of sperm from the seminal receptacle of female Onchocerca volvulus following multiple doses of ivermectin. *Tropical medicine and parasitology: official organ of Deutsche Tropenmedizinische Gesellschaft and of Deutsche Gesellschaft fur Technische Zusammenarbeit (GTZ)*, *44*(3), 155-158.

68. Soboslay, P. T., Lüder, C. G. K., Hoffmann, W. H., Michaelis, I., Helling, G., Heuschkel, C., & Schulz-Key, H. (1994). Ivermectin-facilitated immunity in onchocerciasis; activation of parasite-specific Th1-

# 7 REFERENCES

type responses with subclinical Onchocerca volvulus infection. *Clinical and experimental immunology*, *96*(2), 238.

69. Ducorps, M., Gardon-Wendel, N., Ranque, S., Ndong, W., Boussinesq, M., Gardon, J., & Chippaux, J. P. (1994). [Secondary effects of the treatment of hypermicrofilaremic loiasis using ivermectin]. *Bulletin de la Societe de pathologie exotique (1990)*, *88*(3), 105-112.

70. Gardon, J., Kamgno, J., Folefack, G., Gardon-Wendel, N., Bouchité, B., & Boussinesq, M. (1997). Marked decrease in Loa loa microfilaraemia six and twelve months after a single dose of ivermectin. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *91*(5), 593-594.

71. Boussinesq, M., & Gardon, J. (1998). Challenges for the future: loiasis. *Annals of tropical medicine and parasitology*, *92*, S147.

72. Cotreau, M. M., Warren, S., Ryan, J. L., Fleckenstein, L., Vanapalli, S. R., Brown, K. R., & Schwertschlag, U. S. (2003). The antiparasitic moxidectin: safety, tolerability, and pharmacokinetics in humans. *The Journal of Clinical Pharmacology*, *43*(10), 1108-1115.

73. Hoerauf, A., Mand, S., Volkmann, L., Büttner, M., Marfo-Debrekyei, Y., Taylor, M., & Büttner, D. W. (2003). Doxycycline in the treatment of human onchocerciasis: kinetics of Wolbachia endobacteria reduction and of inhibition of embryogenesis in female Onchocerca worms. *Microbes and Infection*, *5*(4), 261-273.

74. Lustigman, S., James, E. R., Tawe, W., & Abraham, D. (2002). Towards a recombinant antigen vaccine against Onchocerca volvulus. *Trends in parasitology*, *18*(3), 135-141.

75. Eisenbeiss, W. F., Apfel, H., & Meyer, T. F. (1994). Protective immunity linked with a distinct developmental stage of a filarial parasite. *The Journal of Immunology*, *152*(2), 735-742.

76. Le Goff, L., Martin, C., Oswald, I. P., Vuong, P. N., Petit, G., Ungeheuer, M. N., & Bain, O. (2000). Parasitology and immunology of mice vaccinated with irradiated Litomosoides sigmodontis larvae. *Parasitology*, *120*(03), 271-280.

77. Lucius, R., Textor, G., Kern, A., & Kirsten, C. (1991). Acanthocheilonema viteae: vaccination of jirds with irradiation-attenuated stage-3 larvae and with exported larval antigens. *Experimental parasitology*, *73*(2), 184-196.

78. Hörauf, A., & Fleischer, B. (1997). Immune responses to filarial infection in laboratory mice. *Medical microbiology and immunology*, *185*(4), 207-215.

79. Abraham, D., Leon, O., Leon, S., & Lustigman, S. (2001). Development of a recombinant antigen vaccine against infection with the filarial worm Onchocerca volvulus. *Infection and immunity*, *69*(1), 262-270.

80. van der Kleij, D., Tielens, A. G., & Yazdanbakhsh, M. (1999). Recognition of schistosome glycolipids by immunoglobulin E: possible role in immunity. *Infection and immunity*, *67*(11), 5946-5950.

81. McVay, C. S., Bracken, P., Gagliardo, L. F., & Appleton, J. (2000). Antibodies to tyvelose exhibit multiple modes of interference with the epithelial niche of Trichinella spiralis. *Infection and immunity*, *68*(4), 1912-1918.

82. Henkle-Dührsen, K., & Kampkötter, A. (2001). Antioxidant enzyme families in parasitic nematodes. *Molecular and Biochemical Parasitology*, *114*(2), 129-142.

83. Selkirk, M. E., Smith, V. P., Thomas, G. R., & Gounaris, K. (1998). Resistance of filarial nematode parasites to oxidative stress. *International journal for parasitology*, *28*(9), 1315-1332.

84. Riveau, G., Poulain-Godefroy, O., Dupré, L., Remoué, F., Mielcarek, N., Locht, C., & Capron, A. (1998). Glutathione S-transferases of 28kDa as major vaccine candidates against schistosomiasis. *Memórias do Instituto Oswaldo Cruz*, *93*, 87-94.

# 7 REFERENCES

85. Zahner, H., Taubert, A., Harder, A., & von Samson-Himmelstjerna, G. (2001). Filaricidal efficacy of anthelmintically active cyclodepsipeptides. *International journal for parasitology*, *31*(13), 1515-1522.

86. Huang, P., Feng, L., Oldham, E. A., Keating, M. J., & Plunkett, W. (2000). Superoxide dismutase as a target for the selective killing of cancer cells. *Nature*, *407*(6802), 390-395.

87. Leslie, A. G. W. (1992). Recent changes to the MOSFLM package for processing film and image plate data.

88. Kabsch, W. (2010). Xds. *Acta Crystallographica Section D: Biological Crystallography*, *66*(2), 125-132.

89. Otwinowski, Z., Minor, W., & W Jr, C. C. (1997). Processing of X-ray diffraction data collected in oscillation mode.

90. Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallographica Section D: Biological Crystallography*, *62*(1), 72-82.

91. Wang, J. W., Chen, J. R., Gu, Y. X., Zheng, C. D., Jiang, F., Fan, H. F., & Hao, Q. (2004). SAD phasing by combination of direct methods with the SOLVE/RESOLVE procedure. *Acta Crystallographica Section D: Biological Crystallography*, *60*(7), 1244-1253.

92. Collaborative, C. P. (1994). The CCP4 suite: programs for protein crystallography. *Acta crystallographica. Section D, Biological crystallography*, *50*(Pt 5), 760.

93. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., & Terwilliger, T. C. (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallographica Section D: Biological Crystallography*, *58*(11), 1948-1954.

94. Emsley, P., & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography*, *60*(12), 2126-2132.

95. Hura, G. L., Budworth, H., Dyer, K. N., Rambo, R. P., Hammel, M., McMurray, C. T., & Tainer, J. A. (2013). Comprehensive macromolecular conformations mapped by quantitative SAXS analyses. *Nature methods*, *10*(6), 453-454.

96. Svergun, D. I., Koch, M. H., Timmins, P. A., & May, R. P. (2013). *Small angle X-ray and neutron scattering from solutions of biological macromolecules* (Vol. 19). Oxford University Press.

97. Jacques, D. A., Guss, J. M., & Trewhella, J. (2012). Reliable structural interpretation of small-angle scattering data from bio-molecules in solution-the importance of quality control and a standard reporting framework. *BMC structural biology*, *12*(1), 9.

98. Svergun, D. I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of applied crystallography*, *25*(4), 495-503.

99. Franke, D., & Svergun, D. I. (2009). DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *Journal of applied crystallography*, *42*(2), 342-346.

100. Volkov, V. V., & Svergun, D. I. (2003). Uniqueness of ab initio shape determination in small-angle scattering. Journal of applied crystallography, 36(3), 860-864

101. Svergun, D., Barberato, C., & Koch, M. H. J. (1995). CRYSOL–a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of applied crystallography*, *28*(6), 768-773

102. Petoukhov, M. V., & Svergun, D. I. (2005). Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophysical journal*, *89*(2), 1237-1250.

103. Svergun, D. I., Petoukhov, M. V., & Koch, M. H. (2001). Determination of domain structure of proteins from X-ray solution scattering. *Biophysical journal*, *80*(6), 2946-2953.

# 7 REFERENCES

104. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H., & Svergun, D. I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. Journal of applied crystallography, 36(5), 1277-1282.

105. Ghersi,D. and Sanchez,R. (2009) Improving accuracy and efficiency of blind protein-ligand docking by focusing on predicted binding sites. Proteins, 74, 417–424.

106. Sali,A. and Blundell,T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.*, 234, 779–815.

107. Morris,G.M., Goodsell,D.S., Halliday,R.S., Huey,R., Hart,W.E., Belew,R.K. and Olson,A.J. 1998). Automated docking using aLamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.*, 19, 1639–1662.

108. Morris, G.M., R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, et al., AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry*, 2009. 30 (16): p. 2785-2791.

109. Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. *Nature methods*, *12*(1), 7-8.

110. Brenke, R., Kozakov, D., Chuang, G. Y., Beglov, D., Hall, D., Landon, M. R., & Vajda, S. (2009). Fragment-based identification of druggable 'hot spots' of proteins using Fourier domain correlation techniques. *Bioinformatics*, *25*(5), 621-627.

111. Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). *Protein identification and analysis tools on the ExPASy server* (pp. 571-607). Humana Press.

112. Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., & Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL–EBI. *Nucleic acids research*, *38*(suppl 2), W695-W699.

113. McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., & Lopez, R. (2013). Analysis tool web services from the EMBL-EBI. *Nucleic acids research*, *41*(W1), W597-W600.

114. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., & Higgins, D. G. (2011). Fast, scalable generation of high quality protein multiple sequence alignments using Clustal Omega. Molecular systems biology, 7(1), 539.

115. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, *215*(3), 403-410.

116. Gouet, P., Courcelle, E., & Stuart, D. I. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics*, *15*(4), 305-308.

117. Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic acids research*, *42*(W1), W320-W324.

118. Laskowski, R. A., Hutchinson, E. G., Michie, A. D., Wallace, A. C., Jones, M. L., & Thornton, J. M. (1997). PDBsum: a Web-based database of summaries and analyses of all PDB structures. *Trends in biochemical sciences*, *22*(12), 488-490.

119. Holm, L., & Rosenström, P. (2010). Dali server: conservation mapping in 3D. *Nucleic acids research*, *38*(suppl 2), W545-W549.

120. Zhang, Y., & Skolnick, J. (2005). TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic acids research*, *33*(7), 2302-2309.

121. Murzin, A. G., Brenner, S. E., Hubbard, T., & Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *Journal of molecular biology*, *247*(4), 536-540.

# 7 REFERENCES

122. Orengo, C. A., Todd, A. E., & Thornton, J. M. (1999). From protein structure to function. *Current opinion in structural biology*, *9*(3), 374-382.

123. Konagurthu, A. S., Whisstock, J. C., Stuckey, P. J., & Lesk, A. M. (2006). MUSTANG: a multiple structural alignment algorithm. *Proteins: Structure, Function, and Bioinformatics*, *64*(3), 559-574.

124. Krissinel, E., & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *Journal of molecular biology*, *372*(3), 774-797.

125. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., & Baker, N. A. (2004). PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. *Nucleic acids research*, *32*(suppl 2), W665-W667.

126. Frishman, D., & Argos, P. (1995). Knowledge-based protein secondary structure assignment. *Proteins: structure, function, and genetics*, *23*(4), 566-579.

127. Kabsch, W., & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, *22*(12), 2577-2637.

128. Goldenberg, O., Erez, E., Nimrod, G., & Ben-Tal, N. (2009). The ConSurf-DB: pre-calculated evolutionary conservation profiles of protein structures. *Nucleic acids research*, *37*(suppl 1), D323-D327.

129. Celniker, G., Nimrod, G., Ashkenazy, H., Glaser, F., Martz, E., Mayrose, I., & BenTal, N. (2013). ConSurf: using evolutionary data to raise testable hypotheses about protein function. Israel Journal of Chemistry, 53(3-4), 199-206.

130. Barthel, D., Hirst, J. D., Błażewicz, J., Burke, E. K., & Krasnogor, N. (2007). ProCKSI: a decision support system for protein (structure) comparison, knowledge, similarity and information. *BMC bioinformatics*, 8(1), 416.

131. Baker, N., Holst, M., & Wang, F. (2000). Adaptive multilevel finite element solution of the Poisson–Boltzmann equation II. Refinement at solvent accessible surfaces in biomolecular systems. Journal of computational chemistry, 21(15), 1343-1352.

132. Banci L, Bertini I, Hallewell RA, Luchinat C & Viezzoli MS (1989). Water in the active cavity of copper /zinc superoxide dismutase. A water 1H-nuclear-magnetic-relaxation-ispersion study. *Eur J Biochem 184*, 125–129.

133. Ascone I, Castañer R, Tarricone C, Bolognesi M, Stroppolo ME & Desideri A (1997). Evidence of His61 imidazolate bridge rupture in reduced crystalline Cu,Zn superoxide dismutase. *Biochem Biophys Res Commun 241*, 119–121.

134. García-Gutiérrez, P., Landa-Piedra, A., Rodríguez-Romero, A., Parra-Unda, R., & Rojo-Domínguez, A. (2011). Novel inhibitors to Taenia solium Cu/Zn superoxide dismutase identified by virtual screening. *Journal of computer-aided molecular design*, 25(12), 1135-1145.

135. Laurie ATR, Jackson RM (2006). Methods for the prediction of protein–ligand binding sites for structure-based drug design and virtual ligand screnning. *Curr Protein Pept Sci* 7:395–406

136. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., & Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography*, 66(2), 213-221.

137. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., & Vagin, A. A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallographica Section D: Biological Crystallography, 67(4), 355-367.

# 7 REFERENCES

138. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., & Adams, P. D. (2012). Towards automated crystallographic structure refinement with phenix. refine. *Acta Crystallographica Section D: Biological Crystallography*, 68(4), 352-367.

139. Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. Acta Crystallographica Section D: Biological Crystallography, 66(4), 486-501.

140. Hernandez, M., Ghersi, D., & Sanchez, R. (2009). SITEHOUND-web: a server for ligand binding site identification in protein structures. *Nucleic acids research*, 37(suppl 2), W413-W416.

141. Ghersi, D., & Sanchez, R. (2009). Improving accuracy and efficiency of blind protein ligand docking by focusing on predicted binding sites. Proteins: Structure, Function, and Bioinformatics, 74(2), 417-424.

142. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13), 1605-1612.

143. Brattig, N.W., Krawietz, I., Abakar, A.Z., Erttmann, K.D., Kruppa, T.F., Massougbodji, A. (1994). Strong IgG isotypic antibody response in sowdah type onchocerciasis. *Journal of Infectious Diseases* 170, 955–961.

144. Basanez, M.G., Pion, S.D., Boakes, E., Filipe, J.A.N., Churcher, T.S., Boussinesq, M. (2008). Effect of single-dose ivermectin on Onchocerca volvulus: a systematic review and meta-analysis. *Lancet Infectious Diseases* 8, 310–322.

145. Borsboom, G.J., Boatin, B.A., Nagelkerke, N.J., Agoua, H., Akpoboua, K.L., Alley, E.W., Bissan, Y., Renz, A., Yameogo, L., Remme, J.H., Habbema, J.D. (2003). Impact of ivermectin on onchocerciasis transmission: assessing the empirical evidence that repeated ivermectin mass treatments may lead to elimination/eradication in West-Africa. *Filaria Journal* 2, 8.

146. Osei-Atweneboana, M.Y., Eng, J.K., Boakye, D.A., Gyapong, J.O., Prichard, R.K. (2007). Prevalence and intensity of Onchocerca volvulus infection and efficacy of iver-mectin in endemic communities in Ghana: a two-phase epidemiological study. *Lancet* 369, 2021–2029.

147. Churcher, T.S., Pion, S.D.S., Osei-Atweneboana, M.Y., Prichard, R.K., Awadzi, K., Boussinesq, M., Collins, R.C., Whitworth, J.A., Basanez, M.G. (2009). Identifying sub-optimal responses to ivermectin in the treatment of River Blindness. *Pro-ceedings of the National Academy of Sciences of the United States of America* 106, 16716–21621.

148. Ndjonka, D., Agyare, C., Lüersen, K., Djafsia, B., Achukwi, D., Nukenine, E.N., Hensel, A., Liebau, E. (2010). In vitro activity of Cameroonian and Ghanaian medicinal plants on parasitic (Onchocerca ochengi) and free living (Caenorhabditis elegans) nematodes. *Journal of Helminthology* 85, 304–312.

149. Cho-ngwa, F., Abongwa, M., ngemenya, M.N., Nyongbela, K.D. (2010). Selective activ-ity of extracts of Margaritaria discoidea and Homalium africanum on Onchocerca ochengi. *BMC Complementary and Alternative Medicine* 10, 62.

150. Specht, S., Mand, S., Marfo-Debrekyei, Y., Debrah, A.Y., Konadu, P., Adjei, O., Büttner, D.W., Hoerauf, A. (2008). Efficacy of 2- and 4-week rifampicin treatment on the Wolbachia of Onchocerca volvulus. *Parasitology Research* 103, 1303.

151. Harcus, Y.M., Parkinson, J., Fernandez, C., Daub, J., Selkirk, M.E., Blaxter, M.L., Maizels, R.M. (2004). Signal sequence analysis of expressed sequence tags from the nematode Nippostrongylus brasiliensis and the evolution of secreted proteins in parasites. *Genome Biology* 5 (R39), 1–15.

152. Wanji, S., Tendonfor, N., Nji, T., Esum, M., Che, J.N., Nkwescheu, A., Alassa, F., Kamnang, G., Enyong, P.A., Taylor, M., Hoerauf, A., Taylor, D.W. (2009). Community-directed delivery of doxycycline for the treatment of onchocerci-asis in areas of co-endemicity with loiasis in Cameroon. *Parasites & Vectors* 2, 1–10.

# 7 REFERENCES

153. Hewitson, J.P., Harcus, Y.M., Curwen, R.S., Dowle, A.A., Atmadja, A.K., Ashton, P.D., Wilson, A., Maizels, R.M. (2008). The secretome of the filarial parasite, Brugia malayi: proteomic profile of adult excretory–secretory products. *Molecular and Biochemical Parasitology* 160, 8–21.

154. Chandrashekar, R., Curtis, K.C., Lu, W., Weil, G.J. (1998). Molecular cloning of an enzy-matically active thioredoxin peroxidase from Onchocerca volvulus. *Molecular and Biochemical Parasitology* 93, 309–312.

155. Liebau, E., Höppner, J., Mühlmeister, M., Burmeister, C., Lüersen, K., Perbandt, M., Schmetz, C., Büttner, D., Brattig, N. (2008). The secretory omega-class glutathione transferase OvGST3 from the human pathogenic parasite Onchocerca volvulus. *FASEB Journal* 275, 3438–3453.

156. Lizotte-Waniewski, M., Tawe, W., Guiliano, D.B., Lu, W., Liu, J., Williams, S.A., Lustig-man, S. (2000). Identification of potential vaccine and drug target candidates by expressed sequence tag analysis and immunoscreening of Onchocerca volvulus larval cDNA libraries. *Infection and Immunity* 68, 3491–3501.

157. Henkle, K.J., Liebau, E., Müller, S., Bergmann, B., Walter, R.D. (1991). Characterization and molecular cloning of a Cu/Zn superoxide dismutase from the human parasite Onchocerca volvulus. *Infection and Immunity* 59, 2063–2069.

158. James, E.R., McLean, D.C., Perler, F. (1994). Molecular cloning of an Onchocerca volvulus extracellular Cu–Zn superoxide dismutase. *Infection and Immunity* 62, 713–716.

159. Henkle-Dührsen, K., Tuan, R.S., Wildenburg, G., Eschbach, M.L., Tawe, W., Zipfel, P., Walter, R.D. (1997). Localization and functional analysis of the cytosolic and extracellular CuZn superoxide dismutases in the human parasitic nematode *Onchocerca volvulus*. *Molecular and Biochemical Parasitology* 88, 187–202.

160. Choo, K. H., Tan, T. W., & Ranganathan, S. (2005). SPdb–a signal peptide database. *BMC bioinformatics*, *6*(1), 249.

161. Yang, J. T. (1986). Calculation of protein conformation from circular dichroism.*Methods Enzymol.*, *130*, 208-269.

162. Ramachandran, G. N. (1963). Aspects of protein structure. *In International Symposium on Protein Structure and Crystallography* (1963: Madras, India). Academic Press.

163. Bertini, I., Manganl, S., & Viezzoli, M. S. (1998). Structure and properties of copper-zinc superoxide dismutases. *Advances in inorganic chemistry*, 45, 127-250.

164. Rigo, A., Stevanato, R., Viglino, P., & Rotilio, G. (1977). Competitive inhibition of Cu, Zn superoxide dismutase by monovalent anions. *Biochemical and biophysical research communications*, 79(3), 776-783.

165. Mota de Freitas, D., Luchinat, C., Banci, L., Bertini, I., & Valentine, J. S. (1987). Phosphorus-31 NMR study of the interaction of inorganic phosphate with bovine copper-zinc superoxide dismutase. *Inorganic Chemistry*, 26(17), 2788-2791

166. Mota de Freitas, D., Ming, L. J., Ramasamy, R., & Valentine, J. S. (1990). Chlorine-35 and proton NMR study of anion binding to reduced bovine copper-zinc superoxide dismutase. Inorganic Chemistry, 29(18), 3512-3518.

167. Muñoz, I. G., Moran, J. F., Becana, M., & Montoya, G. (2005). The crystal structure of an eukaryotic iron superoxide dismutase suggests intersubunit cooperation during catalysis. *Protein science*, 14(2), 387-394.

168. Hearn, A. S., Fan, L., Lepock, J. R., Luba, J. P., Greenleaf, W. B., Cabelli, D. E., & Silverman, D. N. (2004). *Journal of Biological Chemistry*, 279(7), 5861-5866.

# 7 REFERENCES

169. Hough, M. A., Grossmann, J. G., Antonyuk, S. V., Strange, R. W., Doucette, P. A., Rodriguez, J. A., & Hasnain, S. S. (2004). Dimer destabilization in superoxide dismutase may result in disease-causing properties: structures of motor neuron disease mutants. *Proceedings of the National Academy of Sciences of the United States of America*, 101(16), 5976-5981.

170. Téllez-Valencia, A., Olivares-Illana, V., Hernández-Santoyo, A., Pérez-Montfort, R., Costas, M., Rodríguez-Romero, & Gómez-Puyou, A. (2004). Inactivation of triosephosphate isomerase from Trypanosoma cruzi by an agent that perturbs its dimer interface. *Journal of molecular biology*, 341(5), 1355-1365.

171. Landa, A., Vaca-Paniagua, F., Torres-Rivera, A., & Parra-Unda, R. (2008). Taenia solium: antioxidant metabolism enzymes as targets for cestocidal drugs and vaccines. *Current topics in medicinal chemistry*, 8(5), 393-399

172. Bradley, J.E., Tuan, R.S., Shepley, K.J., Tree, T.I., Maizels, R.M., Helm, R., Gregory, W.F., Unnasch, T.R., 1993. Onchocerca volvulus: characterization of an immunodominant hypodermal antigen present in adult and larval parasites. *Experimental Parasitology* 77, 414-424.

173. Jones, J. T., Smant, G., & Blok, V. C. (2000). SXP/RAL-2 proteins of the potato cyst nematode Globodera rostochiensis: secreted proteins of the hypodermis and amphids. *Nematology*, 2(8), 887-893.

174. Rao, K. V. N., Eswaran, M., Ravi, V., Gnanasekhar, B., Narayanan, R. B., Kaliraj, P., & Scott, A. L. (2000). The Wuchereria bancrofti orthologue of Brugia malayi SXP1 and the diagnosis of bancroftian filariasis. Molecular and biochemical parasitology, 107(1), 71-80.

175. Anthony, R. M., Rutitzky, L. I., Urban, J. F., Stadecker, M. J., & Gause, W. C. (2007). Protective immune mechanisms in helminth infection. Nature Reviews Immunology, 7(12), 975-987.

176. Dissanayake, S., Xu, M., & Piessens, W. F. (1992). A cloned antigen for serological diagnosis of Wuchereria bancrofti microfilaremia with daytime blood samples. Molecular and biochemical parasitology, 56(2), 269-277.

177. Harnett, W., Bradley, J. E., & Garate, T. (1999). Molecular and immunodiagnosis of human filarial nematode infections. Parasitology, 117(07), 59-71.

178. Hess, J. A., Zhan, B., Bonne-Année, S., Deckman, J. M., Bottazzi, M. E., Hotez, P. J., & Abraham, D. (2014). Vaccines to combat river blindness: expression, selection and formulation of vaccines against infection with Onchocerca volvulus in a mouse model. *International journal for parasitology*, 44(9), 637-646.

179. Kelly, S. M., Jess, T. J., & Price, N. C. (2005). How to study proteins by circular dichroism. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 1751(2), 119-139.

180. Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). *Protein identification and analysis tools on the ExPASy server* (pp. 571-607). Humana Press.

181. Sułkowska, J. I., Sułkowski, P., Szymczak, P., & Cieplak, M. (2008). Stabilizing effect of knots on proteins. *Proceedings of the National Academy of Sciences*, *105*(50), 19714-19719.

182. Huang, L., & Makarov, D. E. (2008). Translocation of a knotted polypeptide through a pore. *The Journal of chemical physics*, *129*(12), 121107.

183. Perutz, M. F., Kendrew, J. C., & Watson, H. C. (1965). Structure and function of haemoglobin: II. Some relations between polypeptide chain configuration and amino acid sequence. *Journal of Molecular Biology*, *13*(3), 669-678.

184. Gromiha, M. M. (2010). *Protein bioinformatics: from sequence to function*. Academic Press.

185. Chothia, C. (1975). Structural invariants in protein folding. *Nature*, *254*(5498), 304-308.

# 7 REFERENCES

186.  Virnau, P., Mirny, L. A., & Kardar, M. (2006). Intricate knots in proteins: Function and evolution. *PLoS Comput Biol*, *2*(9), e122.

187.  Lim, N. C., & Jackson, S. E. (2015). Molecular knots in biology and chemistry. *Journal of Physics: Condensed Matter*, *27*(35), 354101.

188.  Dill, K. A. (1990). Dominant forces in protein folding. *Biochemistry*, *29*(31), 7133-7155.

189.  England, J. L., & Haran, G. (2011). Role of solvation effects in protein denaturation: from thermodynamics to single molecules and back. *Annual review of physical chemistry*, *62*, 257.

190.  Jackson, S. E. (1998). How do small single-domain proteins fold?. *Folding and Design*, *3*(4), R81-R91.

191.  Wolynes, P. G. (2005). Recent successes of the energy landscape theory of protein folding and function. *Quarterly reviews of biophysics*, *38*(04), 405-410.

192.  Mallam, A. L., & Jackson, S. E. (2006). Probing nature's knots: the folding pathway of a knotted homodimeric protein. *Journal of molecular biology*, *359*(5), 1420-1436.

193.  Mallam, A. L. (2009). How does a knotted protein fold?. *Febs Journal*, *276*(2), 365-375.

194.  Wallin, S., Zeldovich, K. B., & Shakhnovich, E. I. (2007). The folding mechanics of a knotted protein. *Journal of molecular biology*, *368*(3), 884-893.

195.  Virnau, P., Mirny, L. A., & Kardar, M. (2006). Intricate knots in proteins: Function and evolution. *PLoS Comput Biol*, *2*(9), e122.

196.  Sułkowska, J. I., Rawdon, E. J., Millett, K. C., Onuchic, J. N., & Stasiak, A. (2012). Conservation of complex knotting and slipknotting patterns in proteins. *Proceedings of the National Academy of Sciences*, *109*(26), E1715-E1723.

197.  Soler, M. A., Nunes, A., & Faísca, P. F. (2014). Effects of knot type in the folding of topologically complex lattice proteins. *The Journal of chemical physics*, *141*(2), 025101.

198.  Nureki, O., Shirouzu, M., Hashimoto, K., Ishitani, R., Terada, T., Tamakoshi, & Shibata, T. (2002). An enzyme with a deep trefoil knot for the active-site architecture. *Acta Crystallographica Section D: Biological Crystallography*, *58*(7), 1129-1137.

199.  Sułkowska, J. I., Sułkowski, P., Szymczak, P., & Cieplak, M. (2008). Stabilizing effect of knots on proteins. *Proceedings of the National Academy of Sciences*, *105*(50), 19714-19719.

200.  Sanchez-Ruiz, J. M. (2010). Protein kinetic stability. *Biophysical chemistry*, *148*(1), 1-15.

201.  Elkins, P. A., Watts, J. M., Zalacain, M., van Thiel, A., Vitazka, P. R., Redlak, M., & Holmes, W. M. (2003). Insights into catalysis by a knotted TrmD tRNA methyltransferase. *Journal of molecular biology*, *333*(5), 931-949.

202.  Watanabe, K., Nureki, O., Fukai, S., Ishii, R., Okamoto, H., Yokoyama, & Hori, H. (2005). Roles of conserved amino acid sequence motifs in the SpoU (TrmH) RNA methyltransferase family. *Journal of Biological Chemistry*, *280*(11), 10368-10377.

203.  Sasisekhar, B., Suba, N., Sindhuja, S., Sofi, G. M. A., & Narayanan, R. B. (2005). Setaria digitata: identification and characterization of a hypodermally expressed SXP/RAL2 protein. *Experimental parasitology*, *111*(2), 121-125

204.  Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, & Ahringer, J. (2003). Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. *Nature*, *421*(6920), 231-237.

# 7 REFERENCES

205. Tytgat, T., Vercauteren, I., Vanholme, B., De Meutter, J., Vanhoutte, I., Gheysen, G., & Gheysen, G. (2005). An SXP/RAL-2 protein produced by the subventral pharyngeal glands in the plant parasitic root-knot nematode Meloidogyne incognita. *Parasitology research*, *95*(1), 50-54.

206. Fiumara, F., Fioriti, L., Kandel, E. R., & Hendrickson, W. A. (2010). Essential role of coiled coils for aggregation and activity of Q/N-rich prions and PolyQ proteins. *Cell*, *143*(7), 1121-1135.

207. Lupas, A., Van Dyke, M., & Stock, J. (1991). Predicting coiled coils from protein sequences. *Science*, *252*(5009), 1162-1164.

208. Kim, M. W., Chelliah, Y., Kim, S. W., Otwinowski, Z., & Bezprozvanny, I. (2009). Secondary structure of Huntingtin amino-terminal region. *Structure*, *17*(9), 1205-1212.

209. Mitchell, P. J., & Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science*, *245*(4916), 371-378.

210. Karlin, S., & Burge, C. (1996). Trinucleotide repeats and long homopeptides in genes and proteins associated with nervous system disease and development. *Proceedings of the National Academy of Sciences*, *93*(4), 1560-1565.

# 8 RISK AND SAFTY STATEMENTS

## 8. RISK AND SAFTY STATEMENTS:

### 8.1 Chemicals used (GHS classification):

| Compound | CAS-No. | Supplier | GHS hazard | Hazard Statements | Precautionary Statements |
|---|---|---|---|---|---|
| Acetic acid | 64-19-7 | Chem-solute | GHS02 GHS05 | H226, H314 | P280, P305+351+338, P310 |
| Acrylamide 30% | 79-06-1 | Carl Roth | GHS06 GHS08 | H301, H312, H316, H317, H319, H332, H340, H350, H361f, H372 | P201, P280, P301+310, P305+351+338, P308+313 |
| ATP | 34369-07-8 | Sigma | - | - | - |
| Agarose | 9012-36-6 | Serva | - | - | - |
| $(NH_4)_2SO_4$ | 7283-20-2 | Carl Roth | - | - | - |
| $NH_4NO_3$ | 6484-52-2 | Applichem | GHS03 | H272 | P210 |
| Ampicillin | 69-52-3 | Carl Roth | GHS08 | H334, H317 | P280, P261, P302+P352, P342+P311 |
| AMP-PCP | 7414-56-4 | Sigma | GHS06 | H301, H311, H315, H319, H331, H335 | P261, P280, P301+P310, P305+P351+P338 |
| APS | 7727-54-0 | Carl Roth | GHS03 GHS07 GHS08 | H272, H302, H315, H317, H319, H334; H335 | P280, P305+351+338, P302+352, P304+341, P342+311 |
| Bromphenol blue | 115-39-9 | Applichem | - | - | - |
| $CaCl_2$ | 10043-52-4 | Merck | GHS07 | H319 | P305+351+338 |

# 8 RISK AND SAFTY STATEMENTS

| Compound | CAS-No. | Supplier | GHS hazard | Hazard Statements | Precautionary Statements |
|---|---|---|---|---|---|
| $Ca(H_3CCOO)_2$ | 114460-21-8 | Sigma | - | H315, H319, H335 | P261, P305+P351+P338 |
| Citric acid | 77-92-9 | Sigma | GHS05 | H318 | P305+351+338, P311 |
| Coomassie Brilliant Blue R250 | 6104-59-2 | Serva | - | - | - |
| CHES | 9005-64-5 | Sigma | - | H319 | P305+P351+P338 |
| DTT | 578517 | Applichem | GHS07 | H302, H315, H319, H335 | P302+352, P305+351+338 |
| NaOAc | 127-09-3 | Applichem | - | - | - |
| NaBr | 7647-15-6 | Merck | - | - | - |
| $(CH_3)_2AsO_2Na)$ | 124-65-2 | Sigma | GHS09, GHS06 | H301, H331, H410 | P261, P273, P301+P310, P311, P501 |
| NaCl | 7647-14-5 | Carl Roth | - | - | - |
| $NaH_2PO_4$ | 10049-21-5 | Applichem | - | - | - |
| NaOH | 1310-73-2 | Merck | GHS05 | H314 | P280, P310, P305+351+338 |
| $Na_3$ citrate | 6132-04-3 | Sigma | - | - | - |
| $Ni(II)SO_4$ | 10101-97-0 | Applichem | GHS08, GHS09, GHS07 | H332, H315, H334, H317, H341, H350i, H360D, H372 H410 | P280, P273, P201, P342+P311, P308+P313, P302+P352 |
| Paraffin | 8002-74-2 | Applichem | - | - | - |
| PEG 10000 | 25322-68-3 | Merck | - | - | - |
| PEG 1500 | 25322-68-3 | Fluka | - | - | - |
| PEG 2000 MME | 25322-68-3 | Fluka | - | - | - |

| | | | | | |
|---|---|---|---|---|---|
| PEG 300 | 25322-68-3 | Applichem | - | - | - |
| PEG 3350 | 25322-68-3 | Sigma | - | - | - |
| PEG 400 | 25322-68-3 | Sigma | - | - | - |
| PEG 4000 | 25322-68-3 | Merck | - | - | - |
| PEG 6000 | 25322-68-3 | Merck | - | - | - |
| PEG 8000 | 25322-68-3 | Sigma | - | - | - |
| PMSF | 329-98-6 | Applichem | GHS06, GHS05 | H301, H314 | P280, P305+P351+P338, P310 |
| SDS | 151-21-3 | Sigma | GHS02 GHS06 | H228, H302, H311, H315, H319, H335 | P210, P261, P280, P312, P305+351+338 |
| Sodium citrate | 1545832 | Sigma | - | - | - |
| Sodium Tartrate | 868-18-8 | Applichem | - | - | - |
| EDTA | 60-00-4 | Sigma | GHS07 | H319 | P305+351+338 |
| Ethanol | 64-17-5 | Carl Roth | GHS02 | H225 | P210 |
| Ethidium bromide | 1239-45-8 | Sigma | GHS06, GHS08 | H302, H330, H341 | P260, P281, P284, P310 |
| Glycerol | 56-81-5 | Sigma | - | - | - |
| Guanidinhydro chlorid | 50-01-1 | Applichem | GHS07 | H302, H315, H319 | P305+P351+P338, P302+P352 |
| Hepes | 7365-45-9 | Sigma Aldrich | - | - | - |
| Hydrochloric acid >25 % | 7647-01-0 | Merck | GHS05, GHS07 | H314, H335 | P261, P280, P310, P305+351+338 |

# 8 RISK AND SAFTY STATEMENTS

| | | | | | |
|---|---|---|---|---|---|
| Isopropanol | 67-63-0 | Carl Roth | GHS02, GHS07 | H225, H319, H336. | P210, P233, P305+351+338 |
| Imidazole | 288-32-4 | Carl Roth | GHS05, GHS06, GHS08 | H301; H314; H361 | P260, P281, P303+P361+P3 53, P301+P330+P3 |
| KCl | 7447-40-7 | Carl Roth | - | - | - |
| LiCl | 7447-41-8 | Merck | GHS07 | H302; H315, H319, H335 | P302+352, P305+351+338 |
| Li$_2$SO$_4$ | 10102-25-7 | Merck | GHS07 | H302 | - |
| Mg(HCOO)$_2$ | 6150-82-9 | Fluka | - | - | - |
| MgCl$_2$ | 7786-30-3 | Carl Roth | - | - | - |
| MgOAc | 16674-78-5 | Merck | - | - | - |
| MgSO$_4$ | 7487-88-9 | Merck | - | - | - |
| Methanol | 67-56-1 | Carl Roth | GHS02, GHS06, GHS08 | H225, H301, H311, H331, H370 | P210, P280, P233, P302+P352, P309, P310 |
| 2-Mercaptoethanol | 60-24-2 | Fisher Scientific | GHS06, GHS09 | H302, H411, H315, H335, H311, H319 | P280, P312, P302+P350, P261, P273, P301+P312, P305+P351+P3 38 |
| MPD | 107-41-5 | Carl Roth | GHS07 | H315, H319 | - |
| Seleno-methionine | 3211-76-5 | Sigma | - | H301, H331, H400, H410 | R23/25, R33, H 50, H 53 P 261, P273 |
| NaOAc | 127-09-3 | Applichem | - | - | - |
| NaBr | 7647-15-6 | Merck | - | - | - |
| (CH$_3$)$_2$AsO$_2$Na) | 124-65-2 | Sigma | GHS09, GHS06 | H301, H331, H410 | P261, P273, P301+P310, P311, P501 |
| NaCl | 7647-14-5 | Carl Roth | - | - | - |
| NaH$_2$PO$_4$ | 10049-21-5 | Applichem | - | - | - |

# 8 RISK AND SAFTY STATEMENTS

| | | | | | |
|---|---|---|---|---|---|
| NaOH | 1310-73-2 | Merck | GHS05 | H314 | P280, P310, P305+351+338 |
| Na$_3$ citrate | 6132-04-3 | Sigma | - | - | - |
| Ni(II)SO$_4$ | 10101-97-0 | Applichem | GHS08, GHS09, GHS07 | H332, H315, H334, H317, H341, H350i, H360D, H372 H410 | P280, P273, P201, P342+P311, P308+P313, P302+P352 |
| Paraffin | 8002-74-2 | Applichem | - | - | - |
| PEG 10000 | 25322-68-3 | Merck | - | - | - |
| PEG 1500 | 25322-68-3 | Fluka | - | - | - |
| PEG 2000 MME | 25322-68-3 | Fluka | - | - | - |
| PEG 300 | 25322-68-3 | Applichem | - | - | - |
| PEG 3350 | 25322-68-3 | Sigma | - | - | - |
| PEG 400 | 25322-68-3 | Sigma | - | - | - |
| PEG 4000 | 25322-68-3 | Merck | - | - | - |
| PEG 6000 | 25322-68-3 | Merck | - | - | - |
| PEG 8000 | 25322-68-3 | Sigma | - | - | - |
| PMSF | 329-98-6 | Applichem | GHS06, GHS05 | H301, H314 | P280, P305+P351+P338, P310 |
| SDS | 151-21-3 | Sigma | GHS02 GHS06 | H228, H302, H311, H315, H319, H335 | P210, P261, P280, P312, P305+351+338 |
| Sodium citrate | 1545832 | Sigma | - | - | - |
| Sodium tartrate | 868-18-8 | Applichem | - | - | - |

# 8 RISK AND SAFTY STATEMENTS

| Compound | CAS-No. | Supplier | GHS hazard | Hazard Statements | Precautionary Statements |
|---|---|---|---|---|---|
| TEMED | 110-18-9 | Merck | GHS02 GHS05 GHS07 | H225, H302, H314, H332 | P261, P280, P305+351+338 |
| tert-Butanol | 75-65-0 | AppliChem | GHS02 GHS07 | H225, H319, H332, H335 | P210, P305+351+338, P403+233 |
| Tris | 1185-53-1 | Fluka | GHS07 | H315, H319, H335 | P261, P305+351+338 |
| Tween 20 | 9005-64-5 | Carl Roth | - | - | - |
| Yeast Extract | 8013-01-2 | Serva | - | - | - |

## 8.2. Commercial Protein Screens and Kits:

| Name | Supplier | Risk label | Risk phrases | Safety phrases |
|---|---|---|---|---|
| PCT | Hampton | T, N | R41, R42, R36/37/38 | S20, S26, S45, S53, S61, S36/37/39 - |
| Floppy Choppy | Jena Bio Science | C, Xn, Xi | R35, R41, R42, R36/37/38 | S22, S26, S45, S24/25, S36/37/39: |
| Macrosol | Molecular Dimensions | T, N | R10, R45, R46, R60, R61, R25, R36/37/38, R48/20/22, R51/53 | S20, S26, S45, S53, S61, S36/37/39 |
| Morpheus | Molecular Dimensions | T, N | R10, R45, R46, R60, R61, R63, R23/25, R36/37/38, R48/20/22, R51/53 | S20, S26, S45, S53, S61, S36/37/39 |

# 8 RISK AND SAFTY STATEMENTS

| Name | Supplier | Risk label | Risk phrases | Safety phrases |
|---|---|---|---|---|
| PACT premier | Molecular Dimensions | T | R23/25, R52/53 | S20, S36, S45, S61 |
| Stura /Footprint | Molecular Dimensions | T, N | R10, R45, R46, R60, R61, R25, R36/37/38, R48/20/22, R51/53 | S20, S26, S45, S53, S61, S36/37/39 |
| AmSO4 Suite | Qiagen | T+, N | R10, R25, R26, R45, R46, R60, R61, R48/23/25, R51/53 | S45, S53, S61, S36/37. |
| Classic Suite | Qiagen | T, N | R10, R45, R46, R60, R61, R23/25, R36/37/38, R48/20/22, R51/53 | S20, S26, S45, S53, S36/37/39. , |
| ComPAS Suite | Qiagen | T | R10, R45, R23/24/25, R36/38, R39/23/24/25, R51/53 | S13, S26, S45, S53, S61, S36/37/39. |
| Cryos Suite | Qiagen | T, N | R10, R45, R46, R60, R61, R23/25, R36/37/38, R48/20/22, R51/53 | S20, S26, S45, S53, S61, S36/37/39. |
| JCSG+ Suite | Qiagen | T, N | R10, R21, R41, R45, R23/25, R37/38, R51/53 | S13, S20, S26, S45, S53, S36/37/39 |

# 8 RISK AND SAFTY STATEMENTS

| | | | | |
|---|---|---|---|---|
| Pure Link PCR Purification Kit | Invitrogen | Xn | R22, R36/38 | S28, S24/25 |
| peqGOLD Plasmid Mini Kit | peqlab | - | - | - |

# 8 RISK AND SAFTY STATEMENTS

| 8.3 | GHS and risk symbols and information about hazard-, risk-, safety- and precaution-statements |
|-----|------------------------------------------------------------------------------------------|



Fig.37: GHS pictograms (source: http://www.unece.org/trans/danger/publi/ghs/pictograms.html).



Fig.38: Hazard symbols (according to http://www.hse.gov.uk/) for formulations and respective risk labels

| GHS Hazard Statements | |
|-----------------------|---|
| H 225 | Highly flammable liquid and vapour |
| H 226 | Flammable liquid and vapour |
| H 228 | Flammable solid |
| H 272 | May intensify fire; oxidizer |
| H 301 | Toxic if swallowed |
| H 302 | Harmful if swallowed |
| H 311 | Toxic in contact with skin |
| H 312 | Harmful in contact with skin |
| H 314 | Causes severe skin burns and eye damage |
| H 315 | Causes skin irritation |
| H 316 | Causes mild skin irritation |
| H 317 | May cause an allergic skin reaction |
| H 318 | Causes serious eye damage |
| H 319 | Causes serious eye irritation |

# 8 RISK AND SAFTY STATEMENTS

| H 330 | Fatal if inhaled |
|---|---|
| H 331 | Toxic if inhaled |
| H 332 | Harmful if inhaled |
| H 334 | May cause allergy or asthma symptoms or breathing difficulties if inhaled |
| H 335 | May cause respiratory irritation |
| H 336 | May cause drowsiness or dizziness |
| H 340 | May cause genetic defects |
| GHS Hazard Statements | |
| H 360 | May damage fertility or the unborn child |
| H 341 | Suspected of causing genetic defects |
| H 350 | May cause cancer |
| H 350i | May cause cancer by inhalation |
| H 360D | May damage the unborn child |
| H 361 | Suspected of damaging fertility or the unborn child |
| H361f | Suspected of damaging fertility |
| H 370 | Causes damage to organs |
| H 372 | Causes damage to organs through prolonged or repeated exposure |
| H 410 | Very toxic to aquatic life with long lasting effects |
| H 411 | Toxic to aquatic life with long lasting effects |
| GHS Precautionary Statements | |
| P201 | Obtain special instructions before use |
| P210 | Keep away from heat/sparks/open flames/hot surfaces – No smoking |
| P233 | Keep container tightly closed |

# 8 RISK AND SAFTY STATEMENTS

| | |
|---|---|
| P260 | Do not breathe dust/fume/gas/mist/vapors/spray |
| P261 | Avoid breathing dust/fume/gas/mist/vapors/spray |
| P264 | Wash thoroughly after handling |
| P273 | Avoid release to the environment |
| P281 | Use personal protective equipment as required |
| P280 | Wear protective gloves/protective clothing/eye protection/face protection |
| P284 | Wear respiratory protection |
| P309 | IF exposed or you feel unwell |
| P310 | Immediately call a POISON CENTER or doctor/physician |
| P311 | Call a POISON CENTER or doctor/physician |

| Risk Statements | |
|---|---|
| R8 | Contact with combustible material |
| R10 | May cause fire |
| R20 | Flammable |
| R21 | Harmful by inhalation |
| R22 | Harmful in contact with skin |
| R25 | Harmful if swallowed |
| R35 | Toxic if swallowed |
| R36 | Causes severe burns |
| R38 | Irritating to eyesIrritating to skin |
| R41 | Risk of serious damage to eyes |

# 8 RISK AND SAFTY STATEMENTS

| | |
|---|---|
| R42 | May cause sensitisation by inhalation |
| R43 | May cause sensitisation by skin contact |
| R45 | May cause cancer |
| R46 | May cause heritable genetic damage |
| R60 | May impair fertility |
| R61 | May cause harm to the unborn child |
| R39/23/24/25 | Toxic: danger of very serious irreversible effects through inhalation, in contact with skin and if swallowed |
| R36/37/38 | Irritating to eyes, respiratory system and skin |
| R23/24/25 | Toxic by inhalation, in contact with skin and if swallowed |
| R20/21/22 | Harmful by inhalation, in contact with skin and if swallowed |
| R48/20/22 | Harmful: danger of serious damage to health by prolonged exposure through inhalation and if swallowed |
| R23/25 | Toxic by inhalation and if swallowed |
| R36/38 | Irritating to eyes and skin |

| Safetey Statements | |
|---|---|
| S20 | When using do not eat or drink. Do not breathe dust |
| S22 | In case of contact with eyes, rinse immediately with plenty of water and seek medical advice |
| S26 | In case of accident or if you feel unwell seek medical advice immediately (show the label where possible). |
| S28 | After contact with skin, wash immediately with plenty of water (to be specified by the manufacturer) |
| S45 | If swallowed, seek medical advice immediately and show this container or label |

# 8 RISK AND SAFTY STATEMENTS
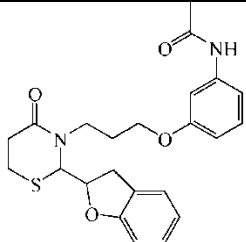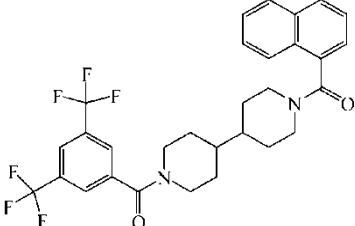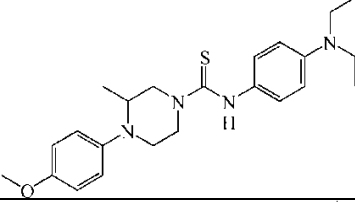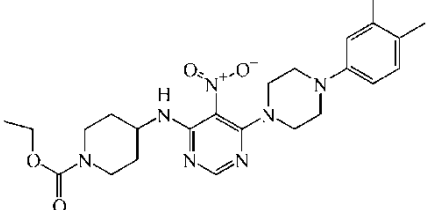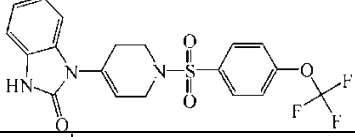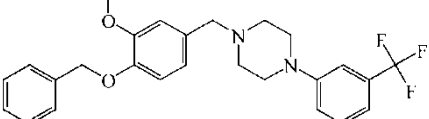
| S46 | Avoid exposure - obtain special instructions before use |
|---|---|
| S53 | Avoid release to the environment |
| S61 | Refer to special instructions/safety data sheet |
| S24/25 | Avoid contact with skin and eyes |
| S36/37 | Wear suitable protective clothing and gloves |
| S36/37/39 | Wear suitable protective clothing, gloves and eye/face protection |

# 9 APPENDIX
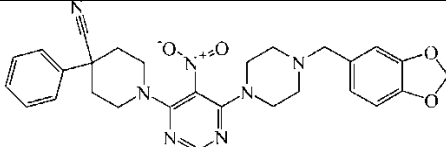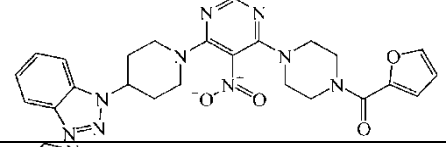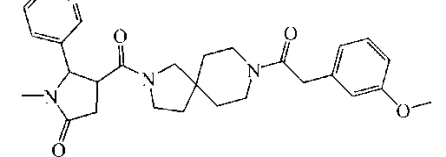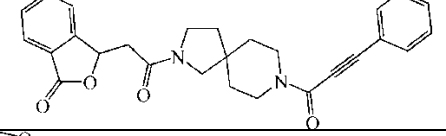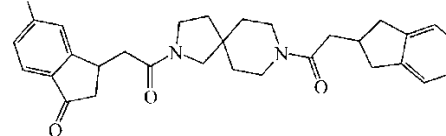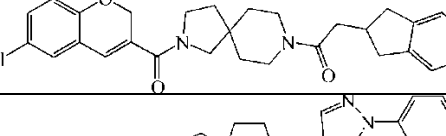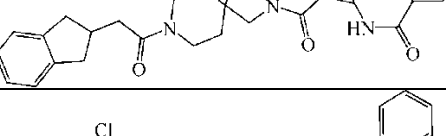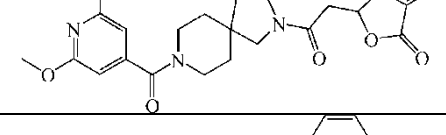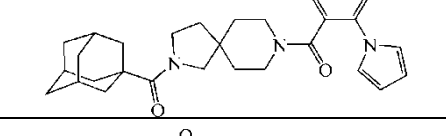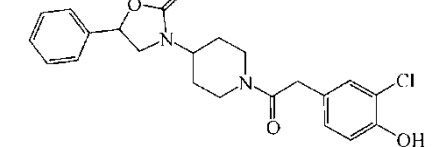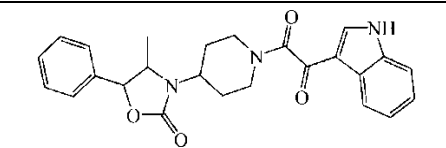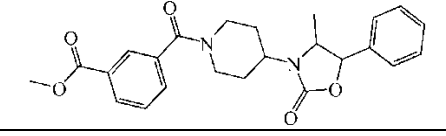
Table 12: Targets amino acid sequences.

| Target | Sequence |
|--------|----------|
| OvEC- SOD | >sp\|Q07449\|SODE_ONCVO Extracellular superoxide dismutase [Cu-Zn] OS=Onchocerca volvulus GN=sod-4 PE=1 SV=1<br>MINSFIVIFLSFLIFINYANLVCVEATHVYGRRSHSNGMHGNGARRAVAVLRGDAGV SGIIYFQQGSGGSITTISGSVSGLTPGLHGFHVHQYGDQTNGCTSAGDHYNPFGKTHG GPNDRIKHIGDLGNIVAGANGVAEVYINSYDIKLRGPLSVIGHSLVVHANTDDLGQGT GNMREESLKTGNAGSRLACGVIGIAAVS |
| Ov Ral-2 | >sp\|P36991\|OV17_ONCVO OV-17 antigen OS=Onchocerca volvulus GN=OV17 PE=2 SV=1<br>MKFVILLTIGLLVVAAIPQRRQQQQQQQQQQQQRDEREIPPFLEGAPPSVIDEFYNLLKT DENKTDQQTEADVEAFINRLGGSYKVRFTQFMEEVKKARADYERIHQQAVARFSPA AKDADARMSAIADSPHLTTRQKSQQIQAIMDSLSESVRREIINALSPPQE |
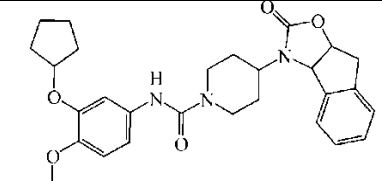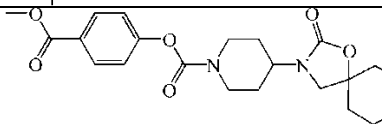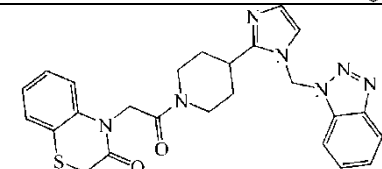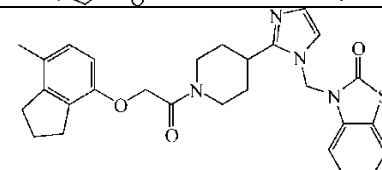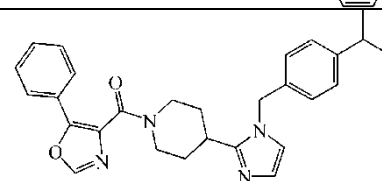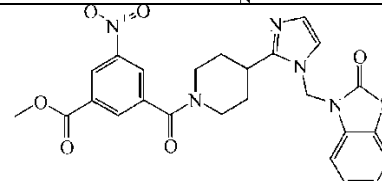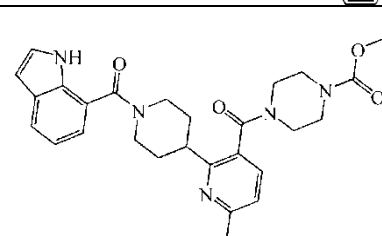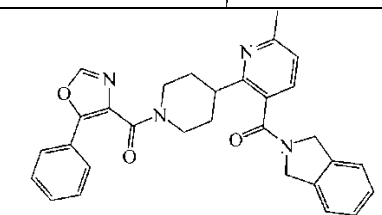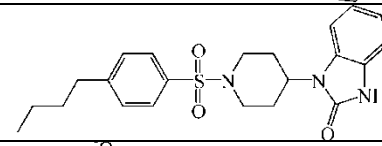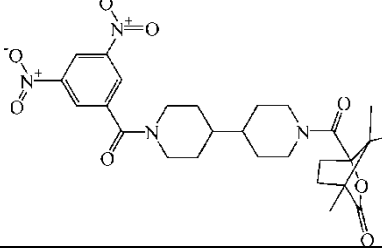
Table 13: LeadQuest® compounds liberary descriptions and 2D structures of fifty predicted potential inhibitors used in docking studies.

| Cmpd | 2D strucutre | LeadQuest ID |
|------|--------------|--------------|
| 1 |  | 1545-07806 |
| 2 |  | 1460-00055 |
| 3 |  | 1526-26128 |
| 4 |  | 1502-13317 |
| 5 |  | 1554-02903 |
| 6 |  | 1524-09920 |

| 7 |  | 1502-14336 |
|---|---|---|
| 8 |  | 1502-15362 |
| 9 |  | 1520-00421 |
| 10 |  | 1520-00623 |
| 11 |  | 1520-01059 |
| 12 |  | 1520-01078 |
| 13 |  | 1520-01545 |
| 14 |  | 1520-02418 |
| 15 |  | 1520-02489 |
| 16 |  | 1534-01314 |
| 17 |  | 1534-02951 |
| 18 |  | 1534-05756 |

| 19 |  | 1534-07719 |
|---|---|---|
| 20 |  | 1534-09941 |
| 21 |  | 1535-06553 |
| 22 |  | 1535-06614 |
| 23 |  | 1535-06726 |
| 24 |  | 1535-06896 |
| 25 |  | 1548-02640 |
| 26 |  | 1548-02825 |
| 27 |  | 1559-04569 |
| 28 |  | 1463-00056 |

| 29 |  | 1506-04337 |
|---|---|---|
| 30 |  | 1511-00170 |
| 31 |  | 1520-03165 |
| 32 |  | 1520-03293 |
| 33 |  | 1520-03820 |
| 34 |  | 1520-04175 |
| 35 |  | 1534-09974 |
| 36 |  | 1548-02420 |
| 37 |  | 1535-01458 |
| 38 |  | 1520-03716 |

| 39 |  | 1534-09962 |
| 40 |  | 1502-14336 |
| 41 |  | 1506-05646 |
| 42 |  | 1520-02828 |
| 43 |  | 1520-02675 |
| 44 |  | 1520-02847 |
| 45 |  | 1506-00980 |
| 46 |  | 1520-01548 |
| 47 |  | 1520-04427 |
| 48 |  | 1520-01219 |
| 49 |  | 1520-00342 |
| 50 |  | 1535-01457 |
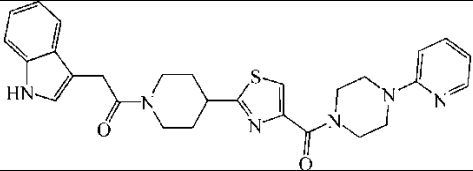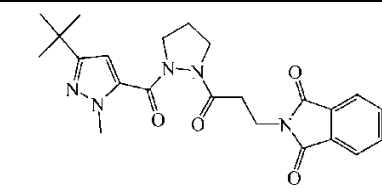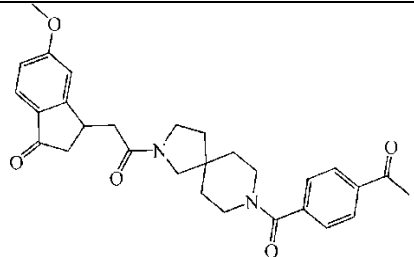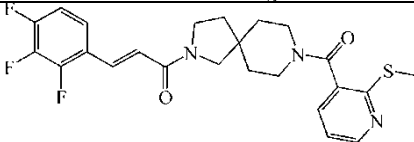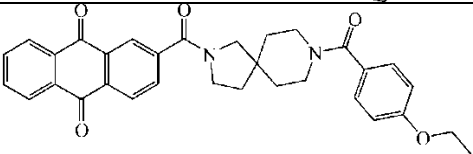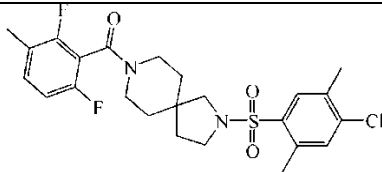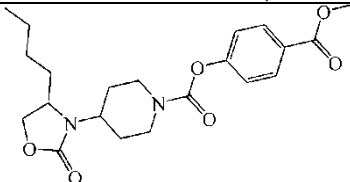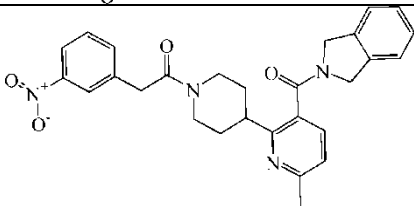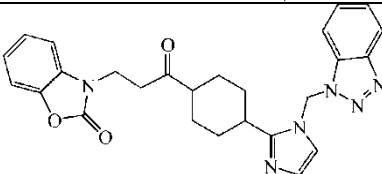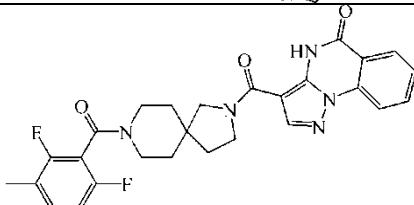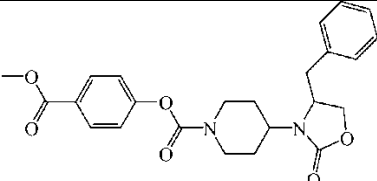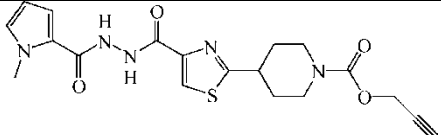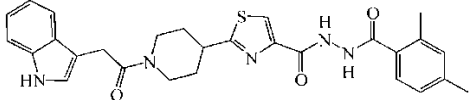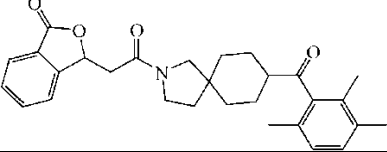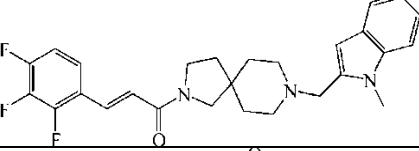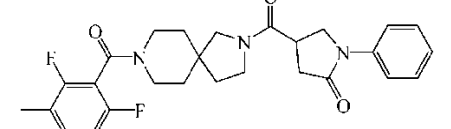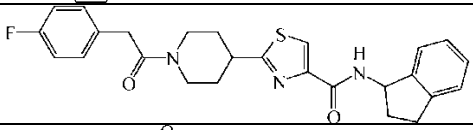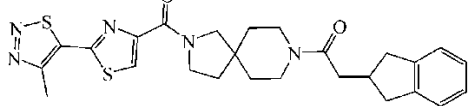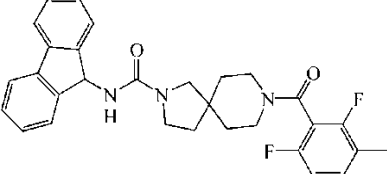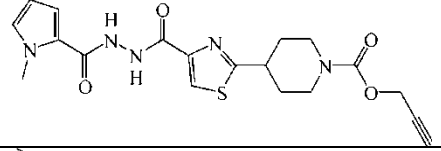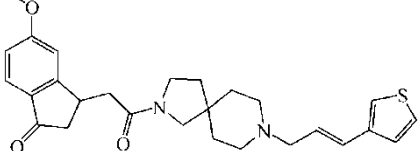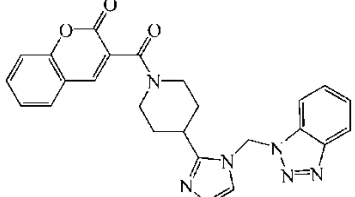
# 10 AKNOWLEDGEMENTS

# 11 CURRICULUM VITAE

## Personal data

| | |
|---|---|
| Date of Birth | 01/10/1982 |
| Place of Birth | Zagazig/Egypt |
| Marital Status | Married |

## Education

**Since 04/2012:** PhD Student, University of Hamburg / Institute for Biochemistry and Molecular biology and Laboratory for Structural Biology of Infection and Inflammation, C/o DESY, Build. 22a, Notkestr. 85 22607, Hamburg, Germany.
Thesis title: Structural and functional analyses of secretory and excretory proteins from *Onchocerca volvulus* as basis for rational drug design

**05/2005 - 03/2008:** Master of Veterinary Medical Science/Biochemistry, Department of Biochemistry and molecular biology, Faculty of Vet. Med, Zagazig University, Egypt. Thesis title: some biochemical and molecular biological studies on immune substance in Arabian Camels, Grad: Very Good with honour Degree.

**09/1999 - 05/2004:** Bachelor of Veterinary Medical Science, Faculty of Vet. Med, Zagazig University, Egypt. Grade: Very Good with honour degree.

**08/1996 - 05/1999:** High school graduation, Zagazig high school for boys, Sharkia providence, Egypt.

## Publications:

- Amr A. Moustafa (2008). Some biochemical studies on Immune substances in camels. M.V.Sc. Thesis. Fac. Vet. Med. Zagazig Univ.

- Amr A. Moustafa, Hussien I. El-belbasi, Mohamed F. Dowidar, Hamad A. El-Sadawy Purification and characterization of camel heavy chain antibodies in Arabian camels. 10th annual conference of Fac. Vet. Med., Zagazig Univ. 516-522.

- Amr A. Moustafa, Hussien I. El-belbasi, Mohamed F. Dowidar, Hamad A. El-Sadawy. Purification and evaluation of camel heavy chain antibodies against tetanus toxoid in comparative with commercial anti tetanic hyper immune sera. 14th international Congress of Immunology, Cube, Japan, oral presentation.

## Mansuscripts in preparation:

- Crystal structure of extra cellular Cu/Zn Superoxide Dismutase from *Onchocerca volvulus* at 1.5 Å; Insights into Novel Binding Site and New Inhibitors.

- The First Crystal Structure of Novel Nematode Specific Immunodominant Hypodermal Antigen (RAL2) from *Onchcerca Volvulus* at 1.7 Å; Insights into Structural Functional Relationship.