Università di Roma "La Sapienza"

DOTTORATO DI RICERCA IN BIOCHIMICA

CICLO XVII

A STRUCTURAL APPROACH TO SCHISTOSOMIASIS: CRYSTAL STRUCTURE OF TWO ANTIGENIC PROTEINS WITH IMPORTANT PHYSIOLOGICAL ROLES

Dottorando: **Francesco Angelucci**
Università di Roma "La Sapienza", Dip. Di Scienze Biochimiche "A. Rossi Fanelli", P.le Aldo Moro 5, 00185 Roma

Docente guida: **Prof. Maurizio Brunori**
Università di Roma "La Sapienza", Dip. Di Scienze Biochimiche "A. Rossi Fanelli", P.le Aldo Moro 5, 00185 Roma

Direttore della scuola: **Prof. Paolo Sarti**
Università di Roma "La Sapienza", Dip. Di Scienze Biochimiche "A. Rossi Fanelli", P.le Aldo Moro 5, 00185 Roma

Commisione giudicatrice:
Prof. Giovanni Antonini
Dip. di Biologia, Università di Roma 3
V.le G. Marconi, 446, 00146 Roma

Prof. Nazzareno Capitanio
Dip. Scienze Biomediche, Università di Foggia
Via Gramsci, 71100 Foggia

Prof. Maurizio Paci (Presidente)
Dip. di Scienze e Tecnologie Chimiche, Università di Roma "Tor Vergata"
Via della Ricerca Scientifica, 00133 Roma

CHAPTER 1

1. INTRODUCTION

1.1 Schistosomiasis: a global view

Schistosomiasis, also known as bilharziasis, is a parasitic disease that leads to chronic illness. It is the greatest health risk in the rural areas of Central China and Egypt and ranks high in several other developing countries. Schistosomiasis has been recognized since the time of the Egyptian pharaohs. The worms responsible for the disease were eventually discovered in 1851 by Theodor Bilharz, a young German pathologist, from whom the disease took its original name, Bilharziasis (Halawany, 1952). The disease is indicated either by the presence of blood in the urine or, in the case of intestinal schistosomiasis, by initially atypical symptoms which can lead to serious complications involving liver and spleen (Jordan, 1972).

The main forms of human schistosomiasis are caused by three species of the flatworm, or blood flukes, known as schistosomes (WHO, 1993-1996-2002): *Schistosoma mansoni*, *Schistosoma haematobium* and *Schistosoma japonicum*.

S. mansoni causes intestinal schistosomiasis and is prevalent in 52 countries and territories of Africa, Caribbean, the Eastern Mediterranean and South America (Figure 1.1).

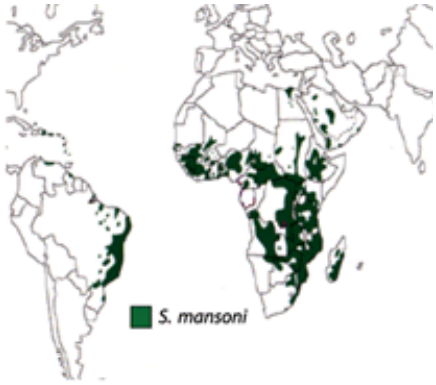


Figure 1.1 *Geographical distribution of S. mansoni*

S. japonicum causes intestinal schistosomiasis and is prevalent in 7 African countries and the Pacific region

S. haematobium causes urinary schistosomiasis and affects 54 countries in Africa and the Eastern Mediterranean (Figure 1.2).

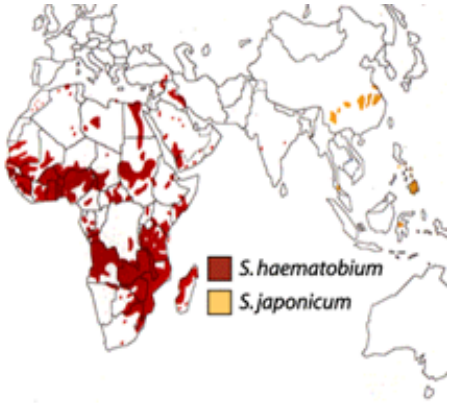


Figure 1.2. *A geographical distribution of S. mansoni and S.*

People are infected by contact with water used in normal daily activities such as personal or domestic hygiene and swimming, or by professional activities such as fishing, rice cultivation, irrigation.

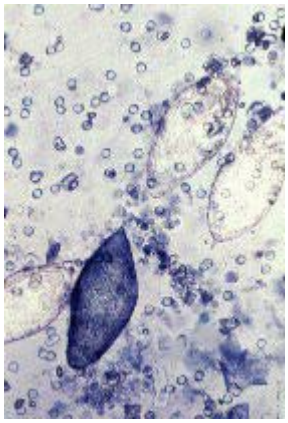


Figure 1.3. *Eggs of schistosomes in the excreta*

Due to lack of information or insufficient attention to hygiene, infected individuals may contaminate their water supply with faeces or urine. The eggs of the schistosomes (Figure 1.3) in the excreta of an infected person open on contact with water and release the larval stage of the parasite, the miracidium. To survive, this motile form must infestate a fresh water snail (Figure 1.4).

Figure1.4. *The Intermediate host: a snail*

Once it has found its host, the miracidium undergoes the intermediate phase of its life cycle, that leads to the production and release of thousands of new parasites (cercariae). The cercariae erupt from the snail into the surrounding water. They can penetrate an individual's skin within a few seconds, continuing their biological cycle (in Figure 1.5) once they have made their way to the victim's blood vessels. Within 30 to 45 days, the parasite is transformed into a worm long 1-2 mm, which is either male or female. The female lays from 200 to 2000 eggs per day over an average of 5 years, according to the species.

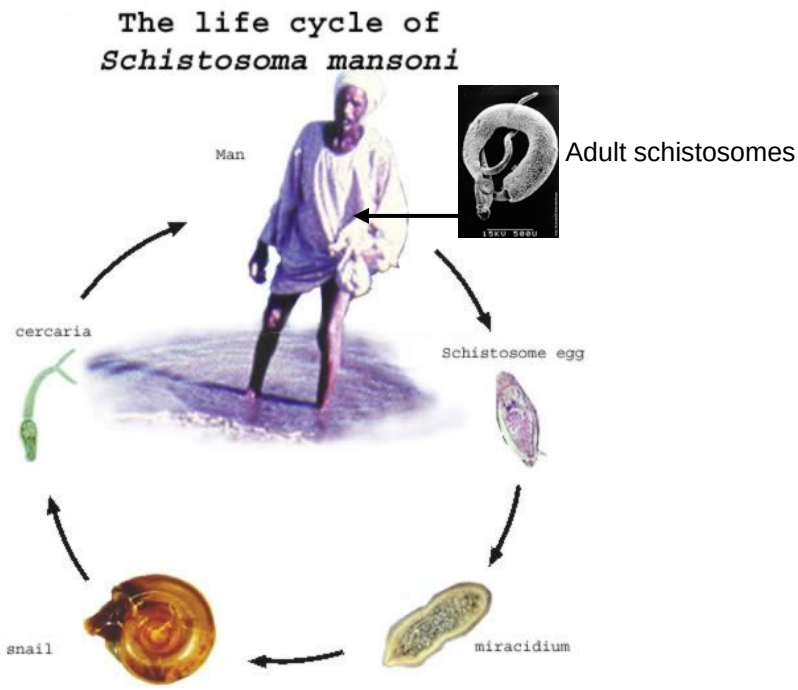


Figure 1.5. *The life cycle of Schistosoma mansoni provides an example for all species of schistosomes. After the eggs of the human-dwelling parasite are emitted with the faeces into the water, the ripe miracidium hatches out of the egg. The miracidium searches for a suitable fresh water snail, an intermediate host. Following penetration, the parasite develops via a so-called mother-sporocyst and daughter-sporocyst generation to the cercaria. The purpose of the growth in the snail is the substantial multiplication of the parasite. From a single miracidium a few thousand cercaria, each one capable of infecting men, are produced. The cercaria propel themselves in water with the aid of their bifurcated tail and actively seek for their final host. When they recognise human skin, the penetration occurs within a very short time. Following a migration through the body within the bloodstream, if they meet a partner of the opposite sex, they develop into sexually mature adults.*

1.2 Infection and Pathogenesis

1.2.1 Epidemiological considerations.

Schistosomiasis is endemic in 74 tropical developing countries. Some 600 million people are at risk of becoming infected. It is estimated that 200 million people are already infected. Extreme poverty, unawareness of the risks, inadequacy or total lack of public health facilities together with the unsanitary conditions in which millions of people lead their daily lives are all factors contributing to the risk of infection. In many areas, a high proportion of children between the ages of 10 and 14 are infected, causing grow delay (Figure 1.6).

Figure 1.6 *25 year-old man with advanced*

Urinary schistosomiasis affects 66 million children throughout 54 countries. In some villages around Lake Volta in Ghana, over 90% of the children are infected by the disease. As with other tropical diseases, population

movements and refugees in unstable regions contribute to the transmission of schistosomiasis. Rapid urbanization has been accompanied by new foci of transmission. The increase in "off-track" tourism has led to increasingly serious infections with previously unencountered sequelae, including paralysis of the legs. The large fresh water reservoirs associated with dams such as Akosombo Dam in Ghana, the Kainji Dam in Nigeria and the Kariba Dam in Zimbabwe as well as smaller reservoirs in the Sahel and irrigation systems throughout Africa are major transmission foci and thus endemic areas for schistosomiasis. Although the majority of people in endemic areas have only light infections or no symptoms, the impact of schistosomiasis on economic conditions and the general health situation should not be underestimated. In the north-east of Brazil, Egypt and Sudan, the work capacity of rural workers has been estimated to be seriously undermined. The disease also substantially affects children's growth and school performance. Moreover the close association of schistosomiasis with natural and artificial water reservoirs frustrates the efforts aimed at developing agricolture.

1.2.2 Life cycle and infection behaviour

The eggs are responsible for most of the pathology associated with the disease. The eggs pass through the walls of the mesenteries, and through the intestinal walls into the gut lumen. How they achieve this is still not well understood, however, it is likely to be a result of a number of interacting factors. Physical factors such as the mechanical action of the egg spine, helped by the host blood pressure and then by the peristaltic action of the gut helps driving the egg into and through the tissues. In addition, the miracidium within the egg has been shown to release proteolytic enzymes which may help digesting its way through the host tissue. The host inflammatory

reaction, which is a delayed type hypersensitivity reaction and forms a granuloma around the egg, also seems to be essential for successful migration of the egg to the lumen of the intestine. Experimental infections in mice given anti-inflammatory agents results in reduced granuloma formation with the eggs becoming trapped in the intestinal tissue. Not all the eggs pass out via the intestine. Many of the eggs are swept back to the liver where they are trapped and form liver granuloma. The schistosome eggs travel in the host tissues for an average of 6 days, between the time they are laid and the moment they leave the host, by which time they are fully embryonated and ready to hatch. There are three main factors which are important in the hatching of schistosoma eggs: temperature (25-30 °C), light and osmotic pressure. On entering a hypotonic environment (fresh water) the increase in the osmotic pressure as water enters the egg and the activation of the enzyme leucine amino peptidase, inhibited by NaCl, result in rupture of the egg's shell. (Note: when isolating schistosome eggs from mouse livers the eggs are maintained in isotonic saline, 0.85%, preventing them from immediate hatching) Each egg, which is mechanically ruptured along its long axis, releases a highly motile (2 mm/sec) ciliated miracidium. It is noteworthy that the sex of the adult worm which will eventually be produced from the miracidium is already determined at this stage. Therefore, if a snail is infected with a single miracidium all the resulting cercariae will produced adults worms in the definitive host which are either all male or all female. The miracidium seeks for and penetrates the snail intermediate host. It can remain infective for 8 - 12 hr. To increase the chance of the miracidia locating the host, they have a negatively geotactic and positively phototactic behavioural response which tends to place them in the general environment of the snail host, *Biomphalaria glabrata.* Chemical attractants from the snail

such as mucus, long chain fatty acids and even amino acids attract the miracidia. After the miracidium makes contact with the snail there is a period of exploratory behaviour prior to penetration. Seventy percent of the miracidia appear to penetrate through the foot of the snail, other penetration sites include the tentacles and the edge of the mantle. Penetration is a combination of mechanical motion of the apical papillae and histolytic secretions released from the penetration glands. The cilia are not lost until after penetration is complete. The location of the the sporocyst, the next developing stage within the snail, is dependent on the schistosome species: *S. mansoni* and *S. haematobium* remain at the site of penetration, usually the foot, whereas *S. japonicum* has a preference for cavity organs, viscera and heart. The sporocyst enters into mitosis and produces about 35-600 daughter sporocysts after about 3 weeks. The daughter sporocysts migrate to the digestive glands of the snails and produce the next infective stage, the cercaria. The average cercarial output from an infected snail has been estimated to be about 1500/day, for a total about 18 days. The period from penetration of the snail to release of the cercariae is about 4 weeks. The cercariae are released on a circadian (24 hr) cycle during daylight hours. It has been determined that the pattern of cercarial shedding is dependent on the focus of the infection and the behaviour of the host population. For example, one study in Guadeloupe island, where there were both a human and a sylvatic foci of infection, showed three different shedding in different regions. In the urban area, where the human population was the focus of infection, cercariae were shed early in the morning. In a remote country area, where the focus of the infection cycled through wild hosts, shedding occurred in late afternoon. However, in a rural community where there humans and rodents acted as hosts, shedding was found to be intermediate to the other

two. The cercaria are very active, non-feeding stages which rely on stored glycogen food reserves. These reserves appear to vary considerably (4.8 - 14.2 ng of glycogen /cercariae) and may depend on the nutritional state of the snail and level of infection. Glycogen reserves in the cercaria have been shown to decline exponentially with time after shedding. The rate of loss is, not surprisingly, greater in the tail than the rest of the cercarial body.

Once released from the snail, periodic bursts of activity keep the cercariae just under the surface of the water. Sudden shadows also evoke cercarial activity. Once in close contact with the next host, products released from the mammal's skin have a dramatic effect on the cercarial behaviour causing continuous swimming motion interspersed with frequent reversals which is a typical example of what is known as klinokinetic behaviour.

Penetration into the host skin has distinct phases: attachment, creeping over the skin surface (both triggered by chemical and thermal stimuli) and penetration into the epidermis (triggered by chemical stimuli, such as aliphatic hydrocarbons and free fatty acids, produced by bacterial esterases working on tri-glycerides on the skin). The actual penetration is a combined mechanical and secretory process whose initial phase may take as little as a few minutes. As penetration proceeds there are structural and physiological changes which occur and accompany the transformation from a free-living infective stage to a parasitic larval schistosomulum.

This metamorfosis takes less than one hour: (i) changes in outer membrane (The cercarial tegument is surrounded by a trilaminate plasma membrane with a thick glycocalyx, wich is lost with the tail at the time of penetration) (ii) empting the various penetration glands (iii) release of vesicles from the tegumental cell bodies beneath the muscular, into the tegumental cytoplasm to form a multi-laminate tegumental surface, which replaces the trilaminate

membrane (iv) finally, release of contents from oesophageal glands into the lumen heralding the start of feeding. This process requires two physiological triggers: elevated temperatures and iso-osmotic conditions. The glycocalyx is thought to control the surface permeability in fresh water and its loss coincides with osmotic sensitivity.

As well as physical changes there are a variety of biochemical changes which also take place during the methamorphosis. Over the first 24 h, there is a switch in the larval energy production from an aerobic glycogen based metabolism to a predominantly anaerobic one, accompanied by an increase in lactate production. At the same time, there is a remarkable turnover of surface molecules on the schistosomula surface, marked by the appearance of more surface molecules which have low immuno-reactivity. By 24-48 h the schistosomula has become completely refractory to antibody mediated immune cell cytotoxicity. In addition to providing molecules on the surface which mask antigenic epitopes, the schistosomula surface also has the ability to coat itself with host molecules. Host erythrocyte surface glycolipids are adsorbed onto the schistosomula surface helping to mask sensitive parasite epitopes from the host defenses system. The parasites enter the initial epidermal layer of the skin very rapidly (less than 30 min) and then come to rest when they reach the dermis which appears to present a temporary barrier to further penetration. They remain at this location for about 40 h. Then they require extra 10 h to locate a venule and a further 8 h to penetrate the venule wall. Once in the blood capillaries the schistosomula are carried to the first capillary bed, the lungs, where they become lodged and double in size over the next few days. This so called “lung phase” lasts from 3-8 days. Following this period the larvae make someway their way to the liver, but the exact pathway is unknown. Upon reaching the liver the schistosomula, mature to

young adults pairing 28-35 days post-infection. When the worms are mature the paired adults migrate out of the liver to the mesenteries where the female begins to lay the eggs.(Figure 1.7).

Figure 1.7. *Schistosoma pairs in the mesenteries*

The male and female worms remain in close association, the slender female lying in a ventral groove on the male surface called the gynecophoral canal. This is an intimate association with the female receiving products, such as glucose, trans-membranously from the male. The importance of the association is highlighted by the fact that if the female fails to mate it does not mature properly and remains stunted. It has been estimated that in human infections adult worms can survive in the host for 20-30 years.

In the mouse model system only about 20% of the initial cercarial inoculum makes it to the adult stage. Although there has been much controversy over the site of attrition of the larval worms, it now seems clear that the greatest loss of larvae occurs during the migration through the lungs, with relatively small losses during migration through the skin.

Although many of the eggs pass through the gut mucosa and exit the host with fecal material, as many as 50% of the eggs can be swept by the blood stream back to the liver, where they become lodged in the liver parenchyma. The lodged eggs cause the host inflammatory response leading to granuloma formation and liver fibrosis, that is the chief cause of pathology.

1.2.3 Pathology of Schistosome Infections.

The pathology associated with a schistosome infection arises primarily from the schistosome eggs (single sex infections, thus no egg production, cause little pathology), wherever they are present, both those in the intestine and those swept back by the blood to the liver, and on occasion to other tissues such as the brain and lungs. At least 50% of the laid eggs fail to reach the exterior.

The clinical signs of the disease pass through different phases:

The acute stage, often known as Katayama fever, is normally found in young children or young adults with no previous exposure to the disease, and is particularly prevalent in individuals with *S. japonicum* infections. The acute reaction is in response to the sudden high level of antigen exposure and is usually associated with the onset of egg deposition. Clinical symptoms consists of skin rashes, asthma-like episodes, daily fever, dizziness, diarrhoea, swollen lymph nodes and aching joints and a number of other non-specific symptoms. Pathologically one sees large florid granuloma with a rich infiltration of eosinophils. Frequently, heavy infections can lead to fibrotic chronic schistosomiasis or the death of the patient.

In chronic schistosomiasis the patient experiences diarrhoea and fever. In children the infection can depress their growth rate. The infection also leads to enlargement of the liver and spleen. Fibrosis of the liver can

result in portal hypertension, hepatosplenomegaly, ascites formation, oesophageal varices leading to fatal hematemesis (vomiting blood). In *S. haematobium*, fibrosis of the bladder may lead to ureteric obstruction, pyelonephritis and hydronephrosis and finally to renal failure.

Over about a 16 day period, after the deposition, occurring predominantly in the periportal area, the egg, becomes surrounded by a dense infiltrate composed mainly by lymphocytes, macrophages and a variable number of eosinophils, held together in a extracellular matrix (Figure 1.8).

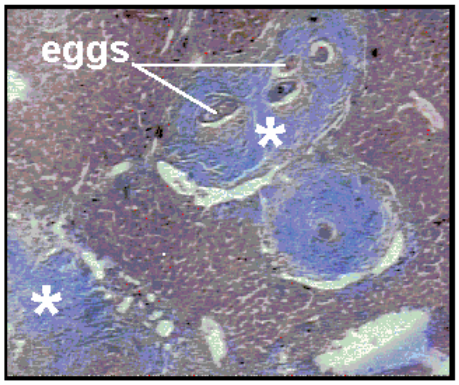


Figure 1.8 *Granuloma formation around*

In athymic mice such lesions do not occur, whereas mice with impared B-cell function still exhibits normal granuloma formation. Therefore granuloma formation is a T-cell dependent and T-cell mediated process.

Eggs cannot pass through the capillary beds as the blood flows through the liver because of their size. The size of granuloma and cell composition vary depending on the schistosome species, the host species and the intensity and duration of the infection, and even tissue location; however, the main factor, is how immunoresponsive the host is to the schistosome egg antigen. Although some of the tissue responses are due to the physical presence and

damage by the eggs, most of the outcome is due to the host response to the soluble egg antigens (SEA) released through submicroscopic egg's shell pores.

In normal hosts, reactivity to SEA, peaks early, producing large florid lesions but as the infection becomes more chronic, i.e. by 8-10 weeks post infection, granulomas tend to become relatively smaller due to a modulation of the host hypersensitivity response.

Activated T-helper cells are instrumental in the induction of IL2, which is the principle cytokine required for the formation of normal granulomas. Down regulation of the production of this cytokine is initiated by a subset of suppressor-inducer T-cells, but modulation is immunologically complex.

Finally, immunoregulation of the granulomas requires multiple effector systems working together to achieve a deceptively simple host adjustment to the persistent generation of parasite antigens. This modulation can be regarded as beneficial to the long term maintenance of the adult parasite and its life cycle; but it is obviously a delicate balance: in fact both the cases of no immuno-depression and complete immunosuppression result in high host mortality.

1.3 Control of Infection

The WHO priorities to control infection are: health education, supplying drinking water and planning adequate healthcare facilities, diagnosis and treatment, management of the environment, control of the intermediate hosts (freshwater snails)

The global distribution of the disease has changed significantly in the past 50 years, with control successes achieved in Japan and in other restricted areas

in Asia, the Americas, North Africa and the Middle East. However, despite this progress the disease remains endemic in a lot of developing countries (as already showed).

Immunological intervention in the form of a schistosome vaccine would complement the success of present efforts in controlling strategies. Nevertheless, it is doubtful whether an effective vaccine will be ready for use within national control programmes, in the next future, the main reason for pessimism being that schistosomiasis does not confer immunity and multiple reinfestations are common.

As we have already mentioned in the previous paragraph, adult schistosomes reside in the capillary net of the abdominal organs and avoid host immune attack by various mechanisms, the most important of which is the adsorption of host antigens to the worm surface. The disease is caused by a granulomatous reaction against the large share of the eggs production, most of which fails to be excreted from the host. Therefore the main clinical forms of disease are due to the eggs, either arrested in the liver, causing hepatomegaly and associated pathology, or occupying the bladder wall resulting in urinary obstruction, kidney damage and frequently bladder cancer. The severity of morbidity is related to the intensity of infection and, as the parasite does not reproduce in the human host, clinical signs will not appear until the worm burden has been boosted by recurrent infections and a sufficient number of egg-associated granulomas have been generated. This process is decelereted by the diverse degrees of immunity developed by endemic populations, but, because a balance must be struck to reduce the risk of overstimulating the granuloma reaction towards the egg, protection builds up slowly and is only partial. It is apparent that, the equilibrium between

opposing immunological mechanisms is a delicate one, and the need for it precludes the classical approach to vaccine development.

Most schistosome antigens were initially identified in *S. mansoni* as this species can be adapted to the laboratory with relative ease. Several antigens have reached an advanced phase of development. WHO/TDR (Special Programme for Research and Training in Tropical Diseases) has developed a programme to assist selection of promising antigens for further development. (Table 1.1).

Table 1.1 *Mayor Schistosome antigens selected by WHO.*

Antigen	Size (kDa)	Stage expressed	Description	Protection (%) (in inbred mice)	Place of development	Legal status
Glutathione S-transferase (P28/GST)	28	Adult/ somula/ egg	Enzyme	30–60	Institut Pasteur, Lille, France	Patented
Paramyosin (Sm97)	97	Adult/ somula	Muscle protein	30	Case Western Reserve University/ National Institutes of Health/ Cornell University, USA	Public domain
IrV-5[a]	62	Adult/ somula/ egg	Muscle protein	50–70	Johns Hopkins School of Medicine, Baltimore, USA	Patented
Triose phosphate isomerase (TPI)	28[b]	Adult/ somula/ egg	Enzyme	30–60	Harvard School of Public Health, Boston, USA	Public domain
Sm23	23[c]	Adult somula/ egg	Integrated membrane protein	40–50	Johns Hopkins School of Medicine/ Harvard School of Public Health, USA	Public domain
Sm14	14	Adult/ somula	Fatty acid-binding protein	65[d]	Instituto Oswaldo Cruz Rio de Janeiro, Brazil	Patented

[a] Abbreviations: IrV, irradiated vaccine antigen; MAP, multiple antigenic peptides.
[b] Tested in the form of a synthetic peptide (MAP-4).
[c] Tested in the form of a synthetic peptide (MAP-3).
[d] In outbred mice.

With the great majority of published work carried out with *S. mansoni*, it is a safe bet that a vaccine against this species will be developed first.

Anyway, it is unclear how proteins that are predicted to be cytosolic, such as Sm14 (Moser et al. 1991), GST (glutathione S-transferase; Balloul et al. 1987), TPI (triose-phosphate isomerase; Harn et al., 1992), could be attached to the schistosome surface, since none has a conventional transmembrane domain and there is no evidence of a lipid anchor. One possiblity is that parasites could be damaged during the course of an infection and soluble proteins are adsorbed to the surface perhaps as part of its strategy to acquire host antigens that act as an antigenic disguise (Simpson, 1992). On the other hand the adsorption of soluble proteins derived from the parasite to the surface seems to be relevant for stimulation of a strong protective immunity (as is the case of the vaccine candidates GST and TPI). Interestingly, there is a striking accumulation of actin in areas of tegument recovering from damage (Matsumoto et al., 1989). Thus, actin or other tegumental proteins are possibly capable of interacting with some soluble proteins allowing attachment to the tegumental surface. There are currently investigation on possible mechanisms of protein adsorbtion to the tegumental surface and the interaction of protein species within the tegumental membrane.

The selection of vaccine candidates favoured antigens showing high rates of protection, together with additional advantages, such as the effect of glutathione s-transferase (GST) on schistosome fecundity (Capron et al 1992) and the cross-reactivity displayed by the fatty acid binding protein antigen against fasciolasis (an economically important disease of cattle; Tendler et al. 1996). Other criteria included the ease of large scale production, the use of adjuvants accepted for human use, the quality of research reported and the number of publications.

Chemotherapy can delay morbidity but must be reiterated periodically. However, morbidity could possibly be suspended permanently, even with

continued trasmission, if chemotherapy were followed by vaccination. The impact of “Trikle infection” is , however, an open question as the short life span of mice and hamsters makes it difficult to emulate this situation experimentally.

Effective drugs, e.g. trivalent antimonials, were introduced in 1918. In the 1920s copper sulphate was shown to be lethal to the aquatic vectors of *S. mansoni* and *S. haematobium* and lime was first used to attack the amphibious vectors of *S. japonicum* (Taylor, 1994).

Currently, diagnosis followed by appropiate chemotherapy remains the milestone of control strategy. At the moment, only three effective and safe drugs are available for treating schistosomiasis: oxamniquine, metrifonate and praziquantel, all against the adult worms. At present praziquantel is the drug of choice for the treatment of all forms of schistosomiasis. It has thus been administered to millions infected individuals in endemic countries. The progressive reduction in costs over the past ten years has also made the drug more readily available. This scenario, however, has recently been tainted by reports of low cure rates in Senegal and the isolation of praziquantel-resistant schistosomes in laboratory. In the meantime, there is a need for increased vigilance in both monitoring and reporting of any emerging praziquantel tolerance/resistance, which would obviously have major implications for control strategies in the future. There is also a strong need to promote research and development of additional anti-schistosomal drugs. The detailed molecular mechanism of action of praziquantel has not yet been elucidated, but a few phenomena connected with its effects are well known (Cioli and Pica-Mattoccia, 2003). The most obvious and immediate modification that can be observed in schistosomes exposed to the drug either in vivo or in vitro is a spastic paralysis of the worm musculature and morphological alterations

in the worm tegument. These alterations are accompanied by an increased exposure of schistosome antigens on the parasite surface (Harnet and Kusel, 1986). Some of the drug exposed antigens have been identified and appear to be connected with the host immune response that is required for a complete activity of praziquantel (Brindley et al. 1989).

The other antischistosomal drug available on the market, oxamniquina, has an excellent record of effecacy and safety for the tratment of infections caused by *S. mansoni* (Foster 1987), but is not active against the other human schistosomes.

The third antischistosomal drug, metrifonate, active only against *S. haematobium*; rarely employed and is no longer available as a brand product, although it can still be found as a generic drug. Therefore a control strategy based on early diagnosis and treatments is actively schemed by research laboratories.

1.4 Actual research directions

Two main areas of research can be identified, one aiming at the production of vaccines and the other at improving on the limited a number of antischistosomal drugs. As with all infection diseases, a preventive vaccine would be the ideal solution, but it may not be easy to implement since natural infection leaves very little immunity. For the same reason, chemotherapy can be effective but inevitably requires multiple expensive treatments due to the high rate of reinfection. Interestingly, one of the effects of chemotherapy with praziquantel is to blister the surface skin of the schistosome, thereby penetrating the shield that makes it invisible to the host, and rendering it vulnerable to specific immune attack and destruction by the host effector cells. A combination of chemotherapy with a vaccine cocktail strategy could therefore well be the main hope for controlling schistosomiasis.

Compared to diseases of the developed world, very little investiment is being made on schistosomiasis research. Innovative approaches based on technological advancements developed in other areas, could be particularly beneficial in this relatively neglected field of schistosomiasis. The resources necessary for a structural biology project on Schistosome will require extensive collaboration between groups performing the genomic, proteomic and the structural and functional work.

In this respect, we started a collaboration with the Institute Pasteur of Lille (France) and the Institute of Cell Biology of Monterotondo of the Consiglio Nazionale delle Ricerche (Italy). Target proteins of adult worms were selected considering their potentialities for a possible cure. The selection criteria for the proteins were: (i) having a role in some of the specific

processes of the schistosomes; (ii) the possibilities to be drug targets; (iii) a demonstrated antigenicity.

CHAPTER 2

2. RESEARCH THESIS PLAN

2.1 The target metabolic pathway: uptake of Arachidonic acid and biosynthesis of prostaglandin D_2

The inhibition of Langerhans Cell (LC) migration plays a key role in immune defense mechanisms and in numerous immunological disorders. It was found that PGD_2, but not the major eicosanoids produced by the schistosome, specifically impedes the tumor necrosis factor-α (TNF-α) triggered migration of LCs through the adenylate cyclase-coupled PGD_2 receptor. Thus, it was proposed that the inhibition of LC migration could represent an additional stratagem for the schistosomes to escape the host immune system; and that the PGD_2 may play a key role in the control of cutaneous immune response (Mountford and Trottein, 2004).

Among the various life cycle stages, the tissue migratory stages of the parasite appeared to produce substantial amounts of PGs.

Previous studies showed that the infective stage of *S. mansoni* (cercaria) requires and is able to produce a wide range of eicosanoids and prostaglandins which are potent regulators of immune response (Fusco et al., 1986). Cercarial eicosanoid production was related to the fatty acid concentration (Salafasky et al., 1984). Various eicosanoid classes were associated with cercarial penetration and transformation. Penetration rates were correlated with increasing leukotriene and hydroxyeicosatetraenoic acid levels, while transformation rates were correlated with increasing prostaglandin levels. A Correlation between eicosanoid production and penetration and transformation rates, strengthened the hypothesis that successful cercarial penetration and metamorphosis are dependent on both skin essential fatty acid levels and cercarial eicosanoid production (Fusco et al., 1986). Interestingly *in vitro*, in the absence of polyunsaturated fatty acids,

no trasformation of cercariae in schistosomula occured and no PGs were detected in the culture supernatant (Ramaswamy et al., 2000). This means that schistosomes, have a metabolism for PG production similar to that of mammals.

In fact schistosomes lack pathways for long chain fatty acid *de novo* synthesis, and thereby derive essential lipids from host's blood (Moser et al., 1991). It has been reported that there is a 32.5% increase in the surface area, and therefore of the lipid membrane, during the first 3 days in the schistosomal development in the vertebrate host (Samuelson et al., 1980). The fatty acid status of mice 6 weeks after infection with *S. mansoni* was investigated and compared with that of uninfected mice. Some fatty acids including linoleic acid and arachidonic acid, were significantly reduced in the plasma of infected mice (Baumheuer et al., 1994); moreover, it was observed that arachidonic acid, the immediate precursor of eicosanoid hormones, is incorporated by the parasite more readily than other lipids (Rumajanek et al., 1980). A cercariometric method to monitor the presence of cercariae based on its need of fatty acids has been developed(Ahmed et al., 2002).

Summarizing: (i) the cercariae are not able to synthetize fatty acids and respond to the presence of unsaturated fatty acids (such as ACD) to start skin penetration and transformation in schistosomuli, (ii) the schistosomulum starts prostaglandin synyhesis in a quantitative way only in presence of unsaturated FAs and (iii) a product of this metabolism, PGD_2, is crucial to circumvent the host immune response and thus, to start infection.

Therefore, incorporation of arachidonic acid and its transformation into PGs has been chosen as a promising subject to investigate in this research thesis, since interfering with either fatty acid uptake or metabolism may constitute an important therapeutical approach.

Within this framework we decided to choose one candidate per each path.

2.1.1 Fatty acid transport.

Schistosomes have a strong requirement for lipids in order to synthetise fatty acid derivatives and to maintain their complex membrane system. Thereby, these parasites need a specific mechanism of transport to acquire lipids from host cells. Uptake and transport of fatty acids and other lipids in *S. mansoni* depend (probably to a large extent) on the fatty acid binding protein (Sm14) (Moser et al., 1991). Sm14 is present in all the stages of the life cycle and is localized in the external cell layer, i.e. near the interface between the parasite and the host (Brito et al., 2002).

2.1.2 Prostaglandin D synthase activity

The key parasite enzyme responsible for PGD_2 production has recently been identified as being a 28 kDa glutathione-S-transferase (termed 28GST) which has PGD_2 synthase (PGDS) activity (Figure 2.1; Herve et al., 2003). These findings suggest a crucial role for 28GST, and its product PGD_2, in the inhibition of Langerhans cell and in the regulation of the immune response during early stages of schistosome infection. 28GST also possesses important functions for the parasite because it actively participates in the detoxification of parasite and host-derived products (through its GST activity) and appears to be crucial for parasite fertility (probably through its PGDS activity). Therefore, inhibition of the enzymatic (GST and PGDS) activities of 28GST might be critical for the parasite survival. We focused on the protein expressed by *S. haematobium* and therefore will refer to it as Sh28GST.

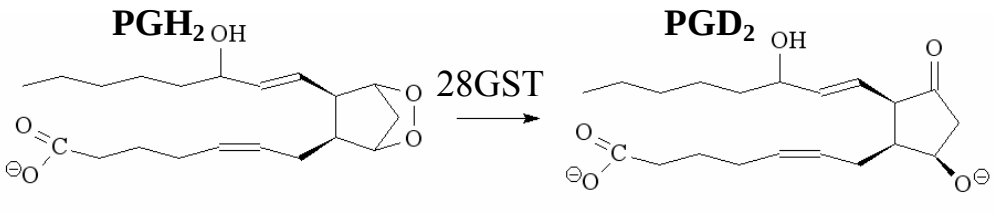


Figure 2.1 *Synthesis of PGD_2 performed by schistosomial GSTs*

Moreover, these two proteins fulfil also the third criterion (see section 1.4) of choice, being two of the six vaccine candidates against schistosomiasis selected by WHO (see Table 1.1).

2.2 Aim of the work

The structural and functional characterization of Sm14 and Sh28GST was the aim of my Ph.D work. No structural studies have been done on either these two proteins, if one excludes the molecular modeling of Sm14 appeared in 1996 (Tendler at al., 1996).

The structural characterization has been achieved through X-ray crystallography, while the functional one has taken advantage of spectroscopic and rapid kinetic techniques. The goal of the structural analyses of these proteins was to obtain a detailed atomic description of the molecular interactions and reactivities with the physiological ligands in order to (i) design new inhibitors, and (ii) explain their antigenic properties, thus leading to more efficacious drugs and vaccines and/or new diagnostic test for the infection.

A thorough functional and structural characterization requires substantial quantities of the protein of choice. Therefore heterologous expression in *E.*

coli. has been used to scale up production. The proteins so obtained were purified to homogeneity and suitable crystallization conditions were discovered. Crystal diffractions data allowed, through different experimental techniques, to obtain electron density maps, from which the protein model is finally built and refined.

In this thesis I report the crystal structures of the Sh28GST, a multifunctional enzyme involved in host-parasite interactions and presently considered as a promising vaccine candidate against schistosomiasis and of Sm14, a protein involved in lipid trafficking.

The structures of the GSH-free enzyme, as well as the partially and fully GSH-saturated enzyme, exhibit a unique feature, absent in GST structures: the invariant Tyr10 side chain occupies two positions (Johnson et al., 2003). The structural constrains responsible for the double conformation are also conserved in the structure of the Y10F mutant. The presence of the 2 conformers of Tyr10 provides a clue about clarifying the multiple catalytic function of Sh28GST and allows us to hypothesize some new enzymatic mechanisms of reaction. Understanding the mechanism of Sh28GST in the activation of (GSH) and in the isomerization of PGs is the first step towards the design of new drugs with a higher inhibitory activity.

Moreover, the proposed mechanism could be general for other GSTs belonging to similar classes; GSTs were related to various pathogenesis phenomena in human, and the study of the structure-function relationships in the interaction with its ligands, has a high pharmaceutical relevance (Armstrong, 1997), as demonstrated by the fact that the search of new inhibitors for this family of proteins is a fundamental and up-to-date task for several pharmaceutical companies.

Sm14 was crystallised with bound oleic acid (OLA) and arachidonic acid (ACD).

The 3D structures of the complexes of Sm14 with ACD and oleic acid (OLA) were solved by means of X-ray crystallography at 2.4 and 1.85Å, respectively. Moreover, we characterized the binding and release reactions of Sm14 and several important FAs using 1,8-anilino-naphtalene-sulphonic acid (ANS) as a competitor. The ligand with highest affinity, when water solubility is taken into account (Richieri et al., 2000), is ACD, a compound essential to schistosomes. ACD is incorporated by the parasite more readily than other lipids (Rumjanek et al., 1980) and is the immediate precursor of eicosanoid hormones, including the prostaglandins, critical in facilitating the skin penetration process of cercariae (Angeli et al., 2001). Kinetic experiments were performed to clarify the mechanism of competition between FAs and ANS; a ternary complex between the protein and the two ligands was postulated to account for the complex reaction mechanism. Since affinities for FAs were shown to be pH dependent, pH jump experiments were performed which proved the presence of a slow pH dependent conformational change, whose physiological relevance is also discussed.

A comparison of the functional data and the two structures obtained by X-ray crystallography explains the structural basis of ligand selectivity. Finally, the 3D-structure of Sm14 also allowed us to structurally highlight two previously described antigenic determinants of the protein.

Comparison with homologus FABPs suggests that the binding site of Sm14 is optimized to fit ACD, the precursor of prostaglandins (Angelucci et al., 2004).

CHAPTER 3

3. GLUTATHIONE S-TRANSFERASE FROM *SCHISTOSOMA HAEMATOBIUM*

3.1 GSTs: What are they good for?

The discovery in 1961 of a catalytic activity for the addition of glutathione (γ-glutaminylcysteinylglycine, GSH) to 1,2-dichloro-4-nitrobenzene in cytosolic liver extracts (Booth et al., 1961) initiated three decades of increasing interest in the genetics and enzymology of the glutathione transferases (GSTs), culminating in the elucidation of the three-dimensional structures of several cytosolic isoenzymes in the first half of this decade (Wilce and Parker, 1994; Armstrong, 1997). The structural information had a profound impact on our understanding of the catalytic mechanisms of the glutathione transferases, the evolution of their fold, and the molecular basis for their enzymatic activity. Mammalian GSTs play several physiological roles, which include: detoxification of endogenous and xenobiotic compounds, by means of the S-linked addition of GSH to electrophilic compounds to facilitate extrusion from cells; simple binding of target compounds for transport or sequestering; isomerization of steroid intermediates and other compounds; and finally eicosanoid biosynthesis (Sheehan et al., 2001).

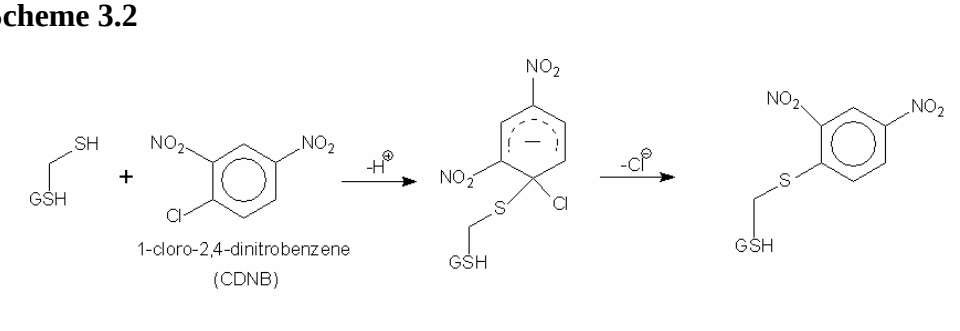
The GSTs catalyze the general reaction shown below (Scheme 3.1):

Scheme 3.1

$$GSH + R\text{-}X \rightarrow GSR + HX$$

and in Scheme 3.2 in more details, with 1-chloro-2,4-dinitrobenzene (CDNB, the model compound used to test the activity of GSTs, as R-X):

Scheme 3.2



The issues with respect to how these enzymes function can be framed by two general categories of questions, addressing on one hand how the enzyme recognizes and activates glutathione for nucleophilic attack and, on the other, how, or even if, the enzyme specifically recognizes electrophilic substrates (R-X). There are both cytosolic and microsomal GSH transferases that are involved in the metabolism of xenobiotics. The cytosolic or soluble enzymes derive from a super-family of genes that encodes at least six individual classes of enzymes, five of which are known to be present in vertebrates. The soluble enzymes defined to date are invariably either homodimeric or heterodimeric species. However, heterodimers can be formed only between subunits of the same enzyme class. Three-dimensional structures of at least one representative of five of the enzyme classes are known. The structures have revealed basic features of each enzymatic class and catalytic features of the enzymes. For example, both the subunit interface and the active site residues involved in the ionization of GSH have evolved considerably though not in tandem. Unfortunately, the nomenclature for the various enzyme classes from species other than humans remains unsettled. A nomenclature has been agreed with the human cytosolic enzymes which is, in theory, applicable to all vertebrate GSTs and extendible to prokaryotes and other

organisms (Mannervick et al., 1992). The human enzymes are named with respect to the class in which they fall (alpha, mu, pi, kappa, sigma and theta).

3.2 Schistosomial glutathione s-transferases: discovery and relevance of their antigenic properties

The 28 kDa glutathione S-transferase of *S. mansoni* (Sm28GST) could be considered today as the best characterized [(molecular cloning, full sequence of the chromosomal gene (McNair et al. 1993) and crystallisation (Trottein et al. 1992b)] and the most tested vaccine candidate antigen of the parasite. The protective properties of this molecule have stimulated work both on the GSTs of other schistosome species and on other enzymes involved in detoxification and antioxidant pathways.

Schistosome GSTs were first identified as potential targets of protective immunity by two separate approaches. The identification of a 28 kDa fraction of soluble *S. mansoni* adult worm antigens that elicits a protective antibody response and confers a high degree of protection to both rats and mice (Balloul et al. 1985) was followed by cloning of the cDNA encoding the major antigenic component of the fraction (Balloul et al. 1987a) which was identified as a GST (Taylor et al. 1988). The recombinant protein elicits protective immunity in rats, hamsters (Balloul et al. 1987a), baboons (Balloul et al. 1987b) and mice (Boulanger et al. 1991). The level of protection conferred against infection varies from an average of 50-70% in rats to 40-50% in mice and 40% in baboons. The relevance of the choice of the 28GST as vaccine candidate has been strengthened by vaccination experiments performed in Sudan against cattle schistosomiasis due to *S. bovis* (Bouchara et al. 1993).

General discussions of immune protection against schistosome infection have concerned a reduction in the numbers of parasites recovered from immunized animals compared to non-immunized controls. The experiments in baboons using recombinant Sm28GST highlighted a different protective effect of immunization, which until recently had not been demonstrated with any other antigen. This effect was a reduction in female worms fecundity reflected in a reduced excretion of eggs by immunized animals. Moreover, a reduction in tissue egg loads was reported in mice immunized with Sm28GST and challenged with *S. mansoni* (Boulanger et al. 1991) and in cattle immunized with native Sb28GST and infected with *S. bovis* (Bouchara et al. 1993). This particular protective effect could be related to the inhibition of the GST enzymatic activity by antibodies (Xu et al. 1991), a vital activity for the survival of the parasite in the definitive host, and probably to the PGDS activity as well (demonstrated for Sm28GST and Sh28GST; Angeli et al., 2001) that may modulate the host immune response to infection by inhibiting the migration of epidermal Langerhans cells to the draining lymph nodes. Thus, these findings open new and accessible concepts in vaccine research against schistosomiasis, dedicated to the reduction of the parasite fecundity, more in terms of anti-pathological vaccines than in terms of eradication of the parasitic disease in the definitive host. Indeed, a vaccine, aimed at preventing the development of the chronic forms of the disease, has become a major goal for applied research, reinforced by the recent emergence of schistosome strains naturally resistant to chemotherapy (Fallon et al. 1996). Indeed more experimental work is required to elucidate how host's antibodies reach an intracellular target such us Sh28GST.

3.3 The 28kDa Glutathione S-transferase from *S. haematobium*

The recent achievements on *S. haematobium* infection in human populations obtained from epidemiological studies, and the results acquired from the research on experimental models of urinary schistosomiasis, give the opportunity to define the parameters for a vaccination in man using Sh28GST. Moreover, the *S. haematobium* protein showed to be the most cross-protective among the 28GSTs; in that way, Phase I clinical trials of Sh28GST in conventional vaccine formulation took place in Europe in December 1997. Phase II has been already designed in three endemic countries for the following years. These trials will represent a major milestone in the achievement of the future vaccine against schistosomiasis.

On the basis of sequence comparison, schistosomal 28 kDa GSTs would belong to the so-called sigma class of GSTs. The enzyme was shown to have high transferase activity with model compound (CDNB), fatty acid hydroperoxide-GSH peroxidase activity (Taylor at al., 1988) and eicosanoid synthesis (Hervé et al., 2003). Thus, Sh28GST is a multifunctional enzyme which plays key role in the host-parasite interactions. In this work we present high-resolution structures of Sh28GST, at different level of saturation with GSH, and the structure of the Y10F mutant. These structural data allowed us to understand the particular chemical properties of this enzyme, to hypothesise new enzymatic mechanism for both its activities and to locate new portions of the polypeptide chain as possible targets for drugs-design.

3.4 Materials and Methods

3.4.1 Expression, Purification, and Preparation of Sh28GST for Crystallization.

The full-length cDNA of Sh28GST was subcloned into pET-24d(+) (Novagen), and *Escherichia coli* BL21(DE3) cells were transformed with this expression vector (M. Hervé et al., 2003). Purification of Sh28GST was carried out as previously described (Trottein et al., 1992). Initially, the recombinant protein was purified by affinity chromatography on a GSH-conjugated Sepharose column, and eluted at pH 7 using GSH. However, to obtain an enzyme preparation without bound GSH, the recombinant Sh28GST was affinity-purified twice on GSH-conjugated Sepharose columns using 50 mM glycine (pH 10) as eluant (Lien et al., 2001). After elution, the protein was dialyzed extensively in phosphate-buffered saline (PBS, pH 7.4) containing 1 mM dithiothreitol, passed through polymyxin columns, concentrated to 66 mg/mL by ultrafiltration (Amicon PM10), and sterile-filtered.

The expression, purification and preparation of the Y10F mutant followed the same procedures developed for the wild type, except that the eluition from GSH-conjugated Sepharose column was performed ony with GSH.

3.4.2 Crystallization.

Crystals were obtained under conditions similar to those reported for Sm28GST (Trottein et al., 1992). Sh28GST was crystallized by the hanging drop vapor diffusion method using a well solution of 2.1 M ammonium sulfate, 100 mM Tris (pH 7.2), and 5 mM β-mercaptoethanol. The hanging drops consisted of a 1:1 mixture of the well solution with the solution of

protein at 66 mg/mL in PBS. Large crystals grew within a few days. Crystals of the GSH-saturated enzyme were obtained from protein equilibrated with a 500 M excess of GSH prior to crystallization trials. Crystals of the GSH-free and the saturated enzymes were obtained following essentially the same procedure, starting with the enzymes prepared in the absence of GSH and incubated, prior to crystallization, with a buffered (KPi 20mM) GSH solutions, respectively.

The Y10F mutant crystallised in the same conditions.

3.4.3 Data Collection and Processing.

Data for the structures were collected at ELETTRA (Trieste, Italy) and at DESY (Hamburg, Germany) at ~100 K. Data were processed with the HKL suite (Otwinowski and Minor, 1997). Autoindexing produced a list containing two likely choices for the space group and unit cell with equivalent scores, either $C222_1$ or $P2_1$. The data from a home source and detector [Rigaku R-axis IV++ detector, coupled to an RU-200 rotating anode X-ray (Cu K) generator equipped with focusing mirrors and operating at 50 kV and 100 mA] to 2.1 Å were used to determine the structure of the partially saturated enzyme. The data were originally processed in space group $C222_1$ with one Sh28GST monomer in the asymmetric unit and the following cell dimensions: a = 72.6 Å, b = 77.4 Å, and c = 77.7 Å. During refinement of the first structure, it appeared likely that the second possibility (space group P21, a = 53.1 Å, b = 77.7 Å, c = 53.2 Å, and = 93.4 which has a dimer of Sh28GST in the asymmetric unit) was afterall the correct choice (see below). Thus, the data sets used for refinement were processed using the $P2_1$ unit cell.

3.4.4 Molecular Replacement, Model Building, and Refinement.

The structure of the partially saturated enzyme was determined by molecular replacement in the $C222_1$ unit cell. The ninth search model which was used, e.g. the structure of Sj26GST (PDB entry 1GTB), resulted in a clear rotation and translation solution using the program AMORE (Navaza, 1994) from the CCP4 suite (CCP4). The structure was refined using REFMAC (Murshudov et al., 1997), and fit to electron density maps using XtalView (McRee, 1999). However, even after the addition of many observed waters, the R-factor converged to 26% which $C222_1$ clearly indicated a problem with the structure. The most likely problem was due to the choice of space group especially given the almost identical length of the b and c axes. Thus, the structure was transferred to the $P2_1$ unit cell and refined easily, leading to a final R factor of less than 23%. This special case of twinning arises from the combination of very similar monomers and the almost identical length of the b and c axes. Thus, the noncrystallographic 2-fold axis of the dimer found in the P21 space group can be "twinned" to a crystallographic 2-fold axis in space group C2221. The two monomers in the P21 asymmetric unit are very similar but do exhibit fine differences especially in the active site.

All the structures with different percentage of GSH and the structure of the Y10F mutant reported in this work belong to the same space group, with very similar unit cell and are solved by difference Fourier method, using as a model the partially saturated Sh28GST structure.

3.4.5 Equilibrium Fluorescence Titrations.

The affinity of Sh28GST and its mutant (Y10F) for reduced GSH was determined by taking advantage of the tryptophan fluorescence quenching occurring upon binding (Dietze et al. 1996). The experiments were carried

out using a Spex Fluoromax spectrofluorimeter in 0.1 M potassium phosphate (pH 7). Sample solutions of Sh28GST were prepared by diluting the stock to concentrations ranging between 0.1 and 5 M and adding reduced GSH (Sigma, St. Louis, MO); usually the dilution due to the addition of titrant was negligible. The excitation wavelength was 280 nm, and emission spectra were collected over the range of 300-400 nm; emission peaked at 337 nm, and readings at this wavelength were used for the quantitative analysis. Measurements were carried out using both preparations of protein: the one purified from the GSH affinity column by elution with glycine at pH 10 and the other with GSH at pH 7. Addition of GSH up to ~1 μM induced a 30% Trp fluorescence quenching. However, in some experiments, further quenching (10%) was observed upon addition of very large amounts of GSH (up to 0.1 M), suggesting the possible existence of low-affinity GSH binding sites, as already reported for similar proteins (Jakobson et al. 1979).

3.4.6 Stopped-Flow Fluorescence Measurements.

Stopped-flow experiments were carried out using an Applied Photophysics (Leatheread, U.K.) MV18 apparatus equipped for fluorescence signal detection. The excitation wavelength was 285 nm, and emission was collected using a filter with cutoff at 320 nm. Solutions of Sh28GST (10 μM) in 0.1 M phosphate (pH 7) were mixed with buffered solutions of GSH at concentrations ranging between 15 and 80 μM. The transmittance of the samples at 285 nm was higher than 75% (with a 1 cm path length). The time courses were fitted to a pseudo-first-order equation to obtain the apparent combination rate constants k_{app}. Linear regression of k_{app} versus the concentration of GSH yielded the intrinsic rate constants for combination (k_{on}) and dissociation (k_{off}) of GSH.

3.4.7 Steady State Determinations.

The catalytic activity of Sh28GST and of the Y10F mutant was measured spectroscopically under steady state conditions, using the chromogenic substrate 4-chloro-1,2-dinitrobenzene (CDNB, Sigma). Two typical assays were performed, either at a fixed GSH concentration or at a fixed CDNB concentration. In the former case, the sample was prepared as follows: a 0.33 μM solution of Sh28GST (0.5 mL) in 0.1 M potassium phosphate buffer (pH 7) was pipetted into a 2 mm spectrophotometric cuvette; then GSH was added to a final concentration of 2 mM. The reaction was started by addition of the appropriate volume of a 0.25 M solution of CDNB in 95% ethanol; the final volume ranged between 0.510 and 0.525 mL (Mannervick and Guthenberg, 1981). The reaction was followed by the absorbance increase at 340 nm in a Hewlett-Packard diode array spectrophotometer. The alternative experiment was carried out in a similar way, at a constant concentration of 2.5 mM CDNB.

3.5 Results

3.5.1 Quality of the Structure.

The structures of Sh28GST free enzyme (**A**), of Sh28GST partially GSH saturated (**B**, the first one to be obtained), of Sh28GST cocrystallized with 3.3 mM GSH (**C**), of Sh28GST cocrystallized in presence of 15 mM GSH (**D**) and of mutant Y10F (**E**) were refined using high-resolution data, 1.9, 1.85, 1.65, 1.5 and 1.9 Å, respectively, with each model achieving final R-factors of 22.8, 22.7, 22.3, 22.5 and 22.9%, respectively (see Table 3.1).

Table 3.1. *Data collection and refinement summary*

	A	B	C	D	E
Resolution	20÷1.9	20÷1.8	20÷1.65	20÷1.5	20÷1.9
R_{merge}	0.07(0.58)[a]	0.07(0.22)[b]	0.04(0.47)[c]	0.04(0.7)[d]	0.07(0.47)[e]
I/σ	39.2(4.2)[a]	21.5(3.4)[b]	40.9(2.7)[c]	15.5(3.9)[d]	11.4(5.8)[e]
B_{wilson}	29.3	25.2	24.9	22.5	33.8
R	0.228	0.227	0.223	0.225	0.229
R_{free}	0.298	0.296	0.286	0.263	0.301
N° H_2O	236	226	317	485	220
N° residue	1(GSH)	1(GSH)	1(GSH)	1(GSH)	1(GSH)

[a]Last shell (1.9-1.95 Å); [b]Last shell (1.81-1.86 Å); [c]Last shell (1.65-1.71 Å); [d]Last shell (1.5-1.56 Å); [e]Last shell (1.9-1.96 Å)

Except where stated differently, the description of the overall structure is referred to the structure at 1.65 Å, the one cocrystallized with 3.3 mM GSH (approximately 60-80% GSH saturated). This is possible because the structures reported in this work are practically identical (r.m.s. deviations of the main chain atoms of the three structures with respect to the one taken as reference are less than or equal of 0.35 Å), with similar statistics and showing few differences in the active sites, that, anyway, will be reported and discussed later in the text.

The distribution of main chain angles in both models shows 93.7% falling in the preferred region of a Ramachandran plot with the remaining 6.3% in the allowed region, and none in the less favourable or generously allowed regions. Only Glu70, found in the tight turn between β-strand 4 and helix 3 (see below), exhibits unusual main chain angles, as already reported in all

other GSTs for the Gln, Glu, or Asp residues found at this position. These unusual main chain angles are thought to arise from the dual function of this residue which is involved in contributing to the dimer interface and also in the formation of the GSH binding site (Widersten et al. 1992). The geometry of all models is good, given that the geometric restraints which were used during refinement of the models led to root-mean-square deviations that were canonical for average bond distances and angles of structures at these resolutions. Analysis of the many geometrical factors of the two final models gives an overall resulting "goodness" or g-factor which is also average for structures at similar resolution (Laskowski et al. 1993). At the moment only the structures at 1.85 and 1.65 Å (the partially and almost fully saturated with GSH) have been deposited in the Protein Data Bank with entries 1OE7 and 1OE8, respectively (Johnson et al., 2003).

3.5.2 Overall Structure.

Sh28GST is a dimer, each monomer having the typical N- and C-terminal domains (Figure 3.1) making up the canonical GST fold. The model contains residues 4-207 in each monomer. The N-terminal domain (residues 4-87), which consists of a four-stranded β-sheet (order 2134) flanked by three helices, forms a thioredoxin-like fold. The C-terminal domain (residues 88-207) consists of six α-helices, or seven in case helix 5a and helix 5b are considered as separate. The N-terminal domain forms the binding site for GSH, called the G-site, whereas the C-terminal domain provides the residues of the H-site where the electrophilic substrate is bound. Small but significant differences are observed between the two monomers, mostly in the GSH binding sites; hence the two monomers are named A and B (Figure 3.1).

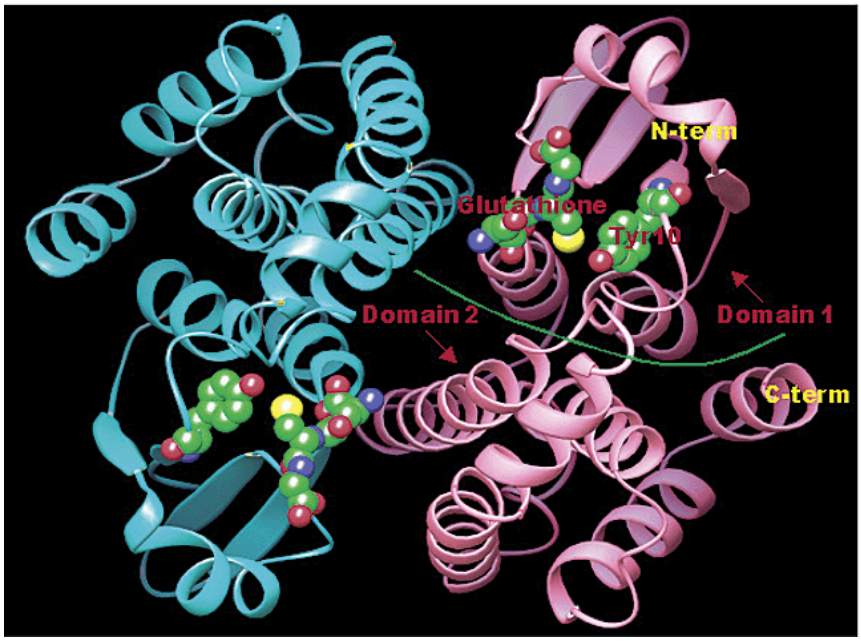


Figure 3.1 *Ribbon-style view of the GSH-saturated Sh28GST dimer, showing the bound GSH and the activating conformer of active site residue Tyr10. Each monomer is composed of two domains. The N-terminal domain (called as domain 1) has a thioredoxin-like fold consisting of a four-stranded β-sheet (order 2134) surrounded by three α-helices, and the C-terminal domain (called domain 2) is composed entirely of α-helices. The N-terminal domain functions mainly to bind GSH, whereas the C-terminal domain forms the xenobiotic binding site. The two monomers are indeed very similar; fine differences in the active sites are discussed in the text.*

The typical overall fold is maintained despite the quite low degree of sequence similarity with other GSTs. This is illustrated by a structure-based sequence alignment of Sh28GST with (i) the *Schistosoma japonicum* 26 kDa GST (Sj26GST) which was used to determine the structure by molecular

replacement and (ii) the structure in the Protein Data Bank whose sequence is most identical to that of Sh28GST (32%), i.e., that of rat hematopoietic PGDS [H-PGDS, PDB entry 1PD2 (Kanaoka et al., 1997)] (see Figure 3.2). Sh28GST, like H-PGDS and the sigma class GSTs, lacks a long C-terminal extension which occludes the active site in Sj28GST and other alpha class GSTs. One distinguishing feature of the Sh28GST fold as compared to other GSTs is the increased length of the β-sheet, which is also evident as a gap in the sequence alignment between strand 3 and strand 4 (Figure 3.2).

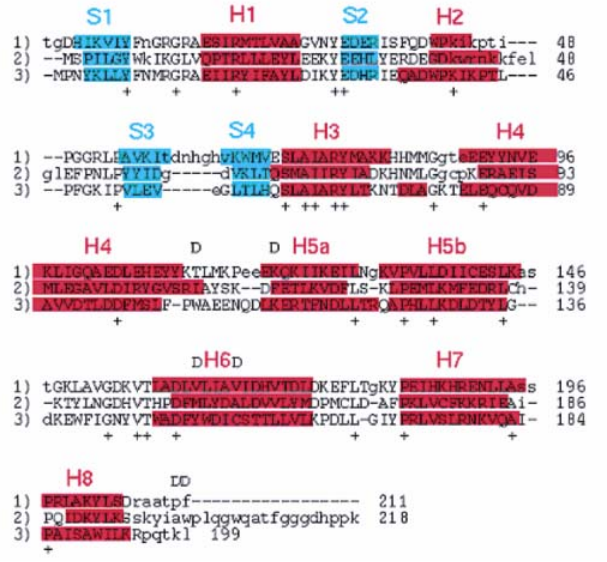


Figure 3.2 *Structure-based sequence alignment using the program SEQUOIA of (1) Sh28GST and two structures discussed in the text, i.e., (2) S. japonicum 26 kDa GST with bound praziquantel [Sj26GST, PDB entry 1gtb, 2.2 Å root-mean-square deviation (rmsd) for superposition of 176 equivalences with 46 identities] and (3) the hematopoietic prostaglandin D synthase (HPDS, PDB entry 1PD2, 1.7 Å rmsd for superposition of 192 equivalences with 61 identities). Equivalent residues (capital letters) are those which after superposition are within 3 Å of the corresponding residues*

in Sh28GST. Helices are denoted in red and strands in cyan. A D appears above residues in HPDS which contribute to the putative prostaglandin binding cleft. Residues Tyr10, Arg16, and Arg21 are discussed in the text.

The crystallographic structure of Sh28GST is compatible with the expectation that the protein forms a dimer in solution, as most GSTs do (Wilce and Parker, 1994, Armstrong et al., 1997). The dimer interaction surface area in the crystal covers 1142 $Å^2$ and 137 shorter contacts than 4 Å, is dominated by nonpolar interactions, but has 21 polar interactions, including four salt bridges. The relatively high ratio of polar to total contacts (21/137) identifies Sh28GST as having a dimer interface similar to those of the sigma class GSTs (Sheehan et al., 2001; Ji et al., 1995). The overall buried surface area, however, is at the lower end of the range typical for GSTs, with monomer-monomer surface areas ranging from ~1000 to ~2000 $Å^2$ (from the Protein Quaternary Structure Server at EMBL-EBI). Like other sigma class GSTs, Sh28GST lacks the hydrophobic "lock-and-key" motif found in the dimer interface of the alpha, pi, and mu classes of GSTs (Sheehan et al., 2001). The level of sequence identity of Sh28GST with nonschistosomal GSTs is at most 32%, which would tend to place the schistosomal GSTs in a separate class given that the cutoff of ~30% sequence identity has been suggested to distinguish GST classes (Sheehan et al., 2001). Schistosomal GSTs are highly homologous, with the lowest level of pairwise sequence identity being 77% (Trottein et al., 1992b).

3.5.3 GSH Binding Site.

Our first structure proved to be partially saturated with GSH despite extensive dialysis against phosphate-buffered saline prior to crystallization;

the active site in each monomer appears to be somewhat less than 50% saturated with GSH in the crystal. The occupancy of GSH was estimated by comparing the B-factors of the C atoms from GSH with those of the polypeptide chain that interact with it (e.g., Glu70). A second structure was obtained by crystallizing Sh28GST exposed to 3.3 mM GSH prior to crystallization; this structure was very little changed but the fractional GSH saturation was higher (Figure 3.3, top panels). Crystallographic data obtained at a yet higher GSH concentration (15 mM) suggested an even higher occupancy. These data are consistent with the hypothesis that the binding of GSH to Sh28GST may display negative cooperativity (as suggested for other GSTs, Lien et al., 2001), a point to be further investigated.

Sh28GST is characterized by a unique feature which, to the best of our knowledge, has not been reported for other GSTs, since the crucial and invariant active site Tyr10 side chain occupies two distinct positions in each monomer (Figure 3.3).

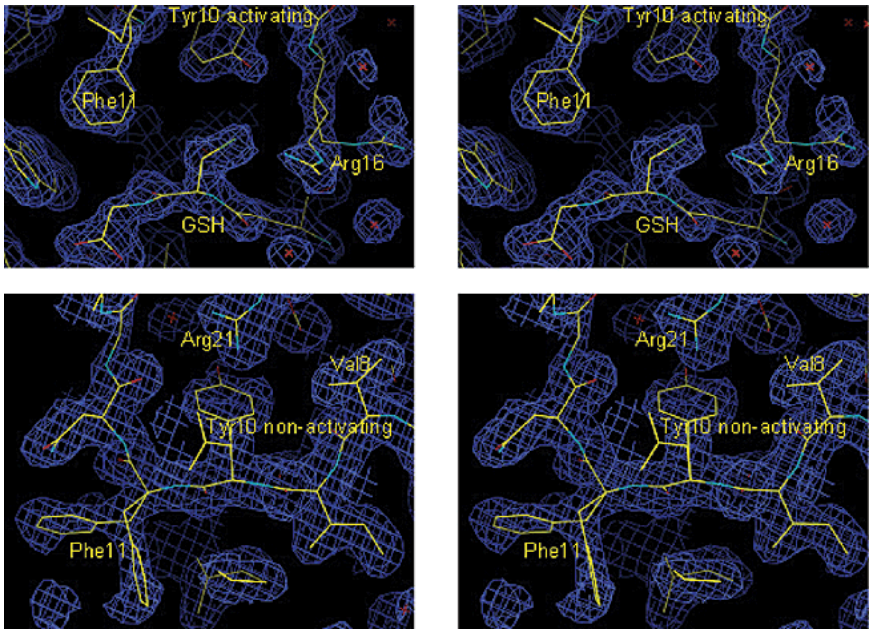


Figure 3.3 *Stereoview of the electron density (2Fo - Fc at 1; blue contours)of the structure of the almost fully saturated Sh28GST at 1.65Å, showing evidence for two conformers of Tyr10 in the active site of monomer B, from the structure which is almost fully saturated with GSH. The stick model uses yellow for carbon, red for oxygen, blue for nitrogen, and green for sulfur. The top panels illustrate the level of disorder in three active site residues, namely, Phe11, Tyr10, and Arg16. The side chain of Tyr10, shown in the activating conformation, forms H-bonds with the S atom of GSH and the main chain N atom of Arg16. The guanidinium group of Arg16 forms an H-bond with the S atom of GSH and a salt link with the carboxylate group of the -Gln of GSH. For Phe11, only one conformer is shown. The bottom panels show the nonactivating conformer of Tyr10 (center) and its polar interaction with the side chain of Arg21 (at the top). The activating conformer of Tyr10 is only partly visible (facing into the page). The two conformers of Phe11 are clearly seen.*

Tyr10 has been proposed (Wilce and Parker, 1994; Armstrong, 1997) to interact with and help forming the active thiolate of GSH by lowering the pKa of the sulfhydryl group from 9 to 6-7. In the free-GSH and in the partially (40-60%) and almost fully (60-80%) GSH-saturated crystal structures the position of the Tyr10 is poorly dependent by the presence of GSH, considering its relative percentage in the two positions as judged from the B-factors and the electron density map that outlines the two conformers.
Anyway, it seems that there is a relationship between the percentage of bound GSH and the double position of the Tyr10 side chain, in fact crystallographic data obtained at a yet higher GSH concentration (15 mM) indicate that the relative amount of the two alternative conformers of Tyr10

may be altered, even though no experimental conditions were found to induce Tyr10 to adopt only the activating position conformer.

GSH is bound to both monomers in the extended conformation, as observed in all other GST-GSH complexes (Koelher et al., 1997), and forms salt bridges with three residues from the N-terminal domain (Arg16, Lys45, and Glu70) of the same monomer and a fourth one with Asp104 from the C-terminal domain of the other monomer. Since the side chain of Arg16 is also found in two conformations, the number of salt bridges might be counted as fewer than four. In addition, GSH forms three H-bonds with protein main chain atoms and seven H-bonds with protein side chain atoms, all from the N-terminal domain of the same monomer. There are also four H-bonds to water molecules, including one to the sulfur thiolate of GSH in the A monomer. It is likely that this water plays a role in stabilizing the ionized state of bound GSH. In comparison, a survey of GSH-protein contacts in high-resolution structures in the Protein Data Bank [alpha class 1F3A (Gu et al. 2000), pi class 8GSS (Oakley et al. 1997), mu class 6GST (Xiao et al., 1999), and Sj26GST 1GNE (Lim et al., 1994)] reveals three salt links for the alpha, pi, and mu class GSTs and two for the Sj26GST. In addition, the mu class GSTs form two H-bonds with protein main chain atoms, whereas the alpha, pi class, and Sj26GST proteins form three H-bonds with protein main chain atoms.

As stated above, two side chains in the G-site have double conformations: namely Tyr10 and Arg16 (Figure 3.3, top and bottom panels). One conformer, the hydroxyl oxygen of Tyr10 makes a H-bond with the sulfur atom of GSH which stabilizes the activated nucleophilic thiolate. In all other GST structures with an active site tyrosine, in the alpha, pi, mu, and sigma classes (Wilce and Parker, 1994; Armstrong, 1997), this activating conformer

is the only one observed. On the other hand, in the GSH-free Sh28GST as well as the partially GSH-bound structures, the Tyr10 side chain is found in the canonical activating conformer at approximately 40-60% occupancy and in the fully saturated enzyme this conformer is occupied at about 80%. Both conformers of Tyr10 are stabilized by polar interactions. The activating conformer is stabilized by (i) a H-bond between the hydroxyl oxygen of Tyr10 and the main chain N atom of Arg16 (as observed in all the other structures of GST-GSH complexes) and (ii) a H-bond with the sulfur atom of GSH. The nonactivating conformer of Tyr10 is stabilized by an electrostatic interaction with the guanidinium group of the side chain of Arg21 in both monomers. Arg 21 is kept in place by the interactions between the oxygens of Asp33 and Glu18 and the NE and NH2 of Arg21, respectively, as shown in figure 3.4.

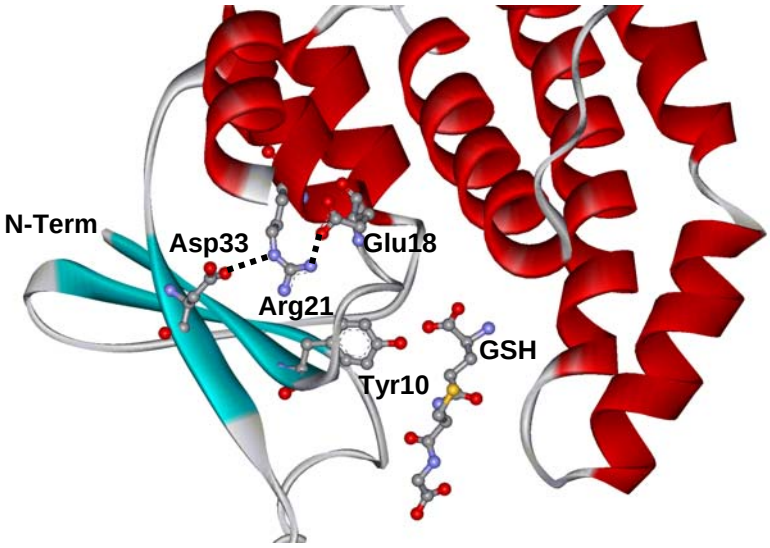


Figure 3.4. *Arg21 is stabilized in Sh28GST by interactions with Asp33 and Glu18. In the picture are shown the relative positions of the catalytic conformer of Tyr10 and GSH with respect to Arg21.*

The side chain of Arg16 also displays a double conformation, with one conformer forming a salt bridge with the carboxylate moiety of the γ-glutamine of bound GSH and the other forming a H-bond with the sulfur atom of GSH, as in alpha class GSTs (Armstrong, 1997). A closer look at inspection of the maps (Figure 3.3, top panels) reveals more electron density surrounding the activating conformer in the almost fully GSH-saturated enzyme. Thus, the conformer of Arg16 which interacts with the GSH sulfur atom appears to have a higher occupancy when the active site is 60-80% GSH saturated.

A third residue in the active site, Phe11, appears to have more than one conformation (Figure 3.3, bottom panels). The major conformer, observed most clearly in the highest-resolution structure of the GSH-saturated complex, the side chain in an unstable position which contacts the side chain of Phe38 in a symmetry-related molecule belonging to another dimer in the crystal unit cell. It is likely that the side chains of Phe11 and Phe38 lie in multiple conformers which make stable crystal contacts but which differ throughout the crystal.

Structural data collected on the Y10F mutant partially saturated with GSH show a conformation of the active site identical to that of the wild type enzyme, and the double conformation of the residue in position 10 (Figure 3.5).

The electrostatic interaction between Arg21 and the aromatic ring of Phe10 is still present. This proves that the two conformers of the catalytic residue, observed in the w.t., is not due to crystallographic artifacts. It is interesting to note that the electron density map looks different for Phe10 (as already observed for Tyr 10 in the w.t.). The electron density for Phe10, when it is in the G-site, is not complete and only the density for the last part of the ring is

visible, while for the conformer stabilized by Arg21 the density showns no interruptions. This could be explained considering that Arg21 interacts with all the electrons of the carbon atoms of the aromatic rings and so the the aromatic ring is constrained to be firmly perpendicular to Arg21, while in the catalytic site the aromatic ring maintains some rotational freedom; moreover the H-bonds formed by the phenolic oxygen of the Tyr10 in the w.t. do not prevent this rotation on the long axis of the ring and thus the density map looks the same in the w.t. structures.

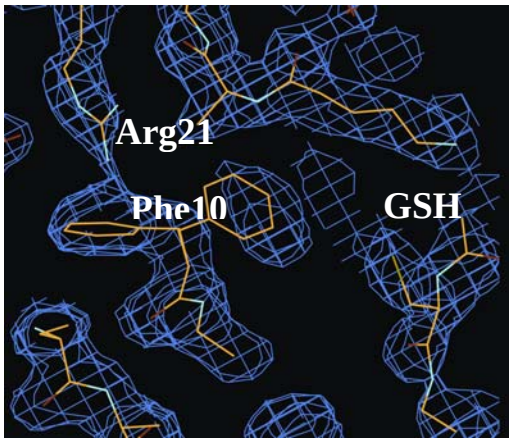


Figure 3.5. *Double conformation of the Phe10 in the structure of the Y10F mutant (1.9Å)*

3.5.4 Comparison of the G-Site in A and B Monomers.

In the almost fully GSH saturated (1.65Å), asymmetry between the A and B monomers in the homodimer is extremely weak, and is limited to the details in the positions of some amino acid residues in the G-site. In any case, the weak asymmetry between the two monomers does not significantly correlate with the alternative conformations of Tyr10, Arg16, and Phe11. The difference in the two monomers is amplified by the position of GSH, since

the orientation of the S atom is different in the A and B monomers. Comparison reveals some differences in geometry around the crucial sulfur atom of bound GSH. In the A monomer, the Tyr10 hydroxyl of the activating conformer is 4.5 Å from the S atom of GSH; close to this atom are also a water molecule (at a distance of 2.95 Å) and the NH1 atom of Arg16 (at 2.86 Å distance). In the B monomer, the Tyr10 hydroxyl in the activating conformation is at a distance of 3.4 Å and the NH1 atom of Arg16 is 2.64 Å from the S atom of GSH, with no water within H-bonding distance. The calculated distances for the H-bonds between a negatively charged S atom and a hydroxyl group, or between a negatively charged or neutral S atom and a variously charged water molecule, have been reported (Liu et al., 1993 and Orozco et al., 1997, respectively). These calculated distances suggest that in the A monomer, the sulfur atom of GSH is negatively charged, therefore a thiolate, and makes one H-bond with a positively charged water molecule. On the other hand, the Tyr10 hydroxyl oxygen at 4.5 Å is too far from the S atom to form a H-bond. In contrast, in the B monomer, the Tyr10 hydroxyl oxygen (3.5 Å) is likely to be neutral and protonated and to form a stable H-bond with the negatively charged S atom. In both the A and B monomers, the NH1 atom of Arg16 is probably also positively charged and within H-bonding distance, though calculations were not reported for this type of atom. The asymmetry between the A and B monomers in the homodimer is possibly relevant with respect to an observed feature of GSTs (Lien et al., 2001), namely negative cooperativity, that would be hard to explain in a perfectly symmetric dimer.

The structures of all other GSTs show that invariant Tyr10 is present only in the activating conformation, which is most relevant for catalysis. The structures of Sh28GST GSH-free, partially and almost fully GSH-saturated,

reveal that, at least in the crystal, Tyr10 occupies a non-activating conformation (see Figure 3.6). Our data do not allow us to precisely estimate if the fraction of Tyr10 in the two alternative conformations is the same in both the A or B monomer; however, the relevant electron densities and B-factors are comparable in the two monomers, and thus, we conclude that the two conformers of Tyr10 are approximately equally populated in both monomers of the active Sh28GST dimer.

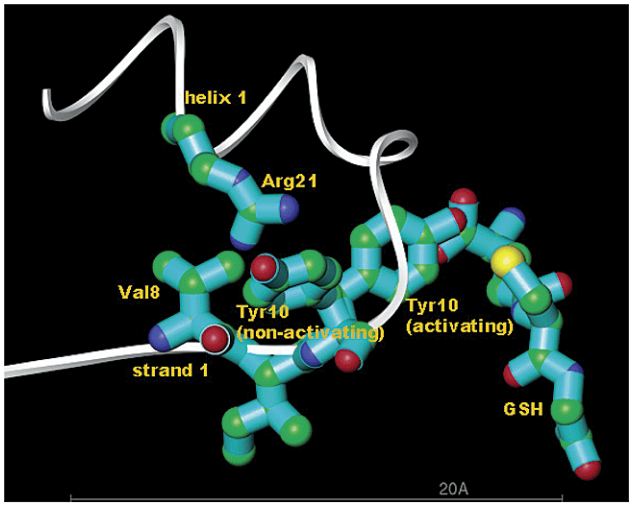


Figure 3.6 *Ball-and-cylinder representation of the two Tyr10 conformers in the B monomer of the almost fully saturated enzyme, with a ribbon-style (white) representation of -strand 1 and -helix 1. Carbon atoms are in green, oxygens in red, and nitrogens in blue, and sulfur is shown in yellow. dThe activating conformer of Tyr10 forms an H-bond with the S atom of GSH; the nonactivating conformer forms a polar interaction with the guanidinium group of Arg21. Interconversion between the two conformers of Tyr10 is blocked in the crystal structure by the side chain of Val8 and the loop between -strand 1 and -helix 1.*

Interestingly, in the fully GSH-saturated structure (1.5Å) the monomers are identical and the oxygen of the tyrosines in the activating conformation of both monomers are found at the same distance to the sulphur atom of GSH (3.4Å). Thus, an excess of GSH force the tyrosine to be predominantly in the G-site by H-bond interaction, and overcomes the tendency of the two monomers to adopt asymmetric conformations.

3.5.5 Comparison of PGDS and Sh28GST H-Site.

GSTs in general catalyze the nonspecific conversion of the prostanoid precursor prostaglandina H_2 (PGH_2) to three products: PGD_2, PGE_2, and PGF_2 (Christ-Hazelhof et al., 1976). A class of GSTs which catalyzes the specific conversion of PGH_2 to PGD_2 has been described (Urade et al., 1987), and for rat H-PGDS (sequence in Figure 3.2), the crystal structure has been published (1PD2, Kanaoka et al. 1996; Pinzar et al., 2000). Sh28GST, whose sequence is 32% identical to that of rat H-PGDS, has been identified as a specific PGDS (Hervé et al., 2003). In the H-PGDS structure, there are two clefts near the G-site which have been proposed to bind the J- and K-arms of the prostanoid precursor, PGH_2. In Sh28GST, the side chains of Glu106, Tyr110, and especially His169 would interfere with binding of PGH_2 in the same mode proposed for H-PGDS. Specifically, the side chain of His169 is sterically hindered from moving and forms two H-bonds, the first between the NE2 atom and the main chain O atom of Arg14 and the second between the ND1 atom and the side chain OE1 atom of Glu106. However, the H-site near the bound GSH in Sh28GST is quite very open with space for both the approach of the PGH_2 headgroup near the GSH thiolate and the positioning of both arms of the substrate in contact with the protein surface.

3.5.6 Helix 5A Major Antigenic Peptide.

Natural antibodies to schistosomal GSTs are known to inhibit the transferase activity (Xu et al., 1991); however, this effect is not due to antibodies against the most potent antigenic and cross-reactive region of the protein molecule, identified from investigations of various antigenic peptides of schistosome 28GSTs as the Pro115-Lys131 fragment. This region may be important in the cross-protection against different schistosome species, including *S. mansoni*, *S. haematobium*, and *Schistosoma bovis* (Trottein et al., 1992b). In the crystal structure, this peptide contains the last part of the loop between helix 4 and helix 5a up until the first residue of helix 5b. The Pro115-Lys131 peptides in the A and B monomers are located topologically close to each other and near the G- and H-sites, as shown in Figure 3.7.

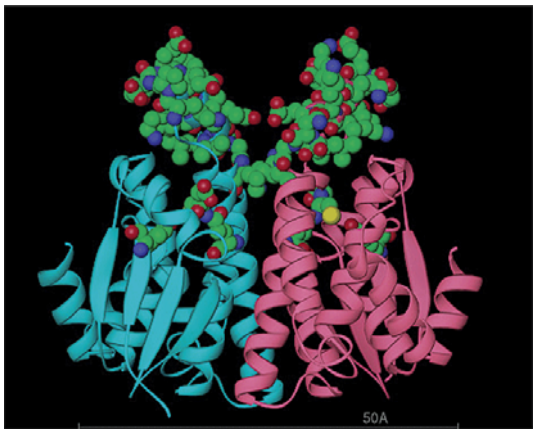


Figure 3.7 *Ribbon-style view of the Sh28GST dimer showing the location of residues 115-131 (top, in VDW balls) which make up the cross-protective antigen and its typology and distance from the active site. Also shown are the bound GSH (center) and the active site Tyr10 (center left and center right) in the activating conformer. These antigenic peptides, located near each other on the protein surface, contain many charged residues but do not contribute*

to the stability of the monomer-monomer interface (since they form no stabilizing contacts in the Sh28GST dimer). Carbon atoms are shown in green, oxygens in red, and nitrogens in blue, and the sulfur of GSH is shown in yellow.

3.5.7 Kinetic Characterization of Sh28GST.

The finding that Sh28GST was partially saturated with GSH in the first crystal structure, despite extensive dialysis prior to crystallization, indicated that GSH is bound tightly. To measure the affinity of Sh28GST for GSH, we exploited the decrease in fluorescence of the protein upon GSH binding, mostly attributed to a change in the environment of Trp 41 which forms H-bond between its side chain NE1 atom and the carbonyl oxygen of the glycine moiety of bound GSH (Dietze et al. 1996). Titrations of Sh28GST at very low concentrations (0.1 μM) with GSH under equilibrium conditions were carried out for the two preparations of the protein. As shown in Figure 3.8 A, we obtained similar results with a fitted K_d of 3 μM. In addition, we employed stopped-flow fluorescence measurements to follow the time course of GSH binding which resulted in all cases to a simple exponential. The estimated affinity for GSH from these experiments is model-dependent and assumes second-order, reversible binding to independent monomers; this expectation is consistent with the experimental observations (Figure 3.8 B). The stopped-flow experiments, also carried out on both batches of the protein, yielded an estimated K_d of 9 μM. Both estimates of K_d (3 μM by equilibrium and 9 μM by kinetics) are similar to the smallest previously reported values (e.g., 7 μM for an alpha class rat liver GST) (Jakobson et al., 1979), and are significantly lower than the usual range reported of 300 μM.

Thus, Sh28GST has an affinity for GSH which ranks among the highest, up to 100 times higher than values usually reported for GSTs.

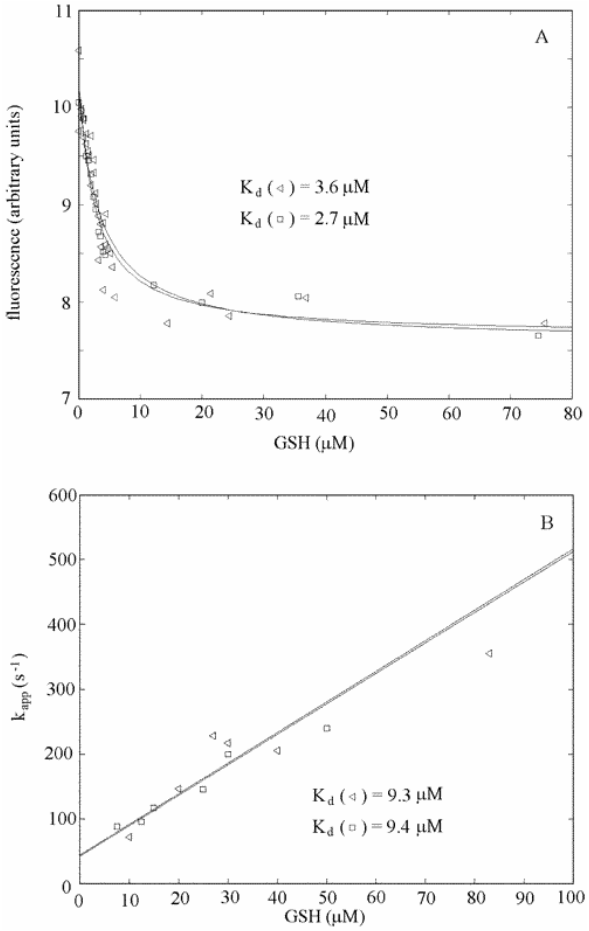


Figure 3.8 *Measurements of affinity of Sh28GST for GSH by equilibrium fluorescence titration (A) and stopped-flow kinetics (B). Experiments were carried out with two enzyme preparations purified using either GSH (squares) or glycine (triangles) to elute the protein from the GSH-Sepharose affinity column. The two enzyme preparations gave similar Kd values for GSH binding, by both methods.*

For the Y10F mutant, eluted from the GSH-conjugated column with GSH, the affinity for GSH is about 30μM, suggesting that the H-bond between the Tyr10 and GSH is relevant for ligand binding.

The transferase activity with CDNB, measured as detailed in Materials and Methods, corresponds to a specific activity of 12 mol of product formed per minute per milligram of protein, similar to those found for other GSTs, including H-PGDS (Thomson et al., 1998; Jowsey et al., 2001). On the contrary, the Y10F mutant appears to be practically inactive, for detoxification and PGDS activities (Trottein et al., unpublished experiments).

3.6 Discussion

In comparison to other reported GST structures, the residues in the active site of Sh28GST exhibit two side chain conformers (Tyr10, Phe11, and Arg16). The invariant active site Tyr10 has been extensively studied because of its essential role in the transferase mechanism (Armstrong, 1997). As far as we know, the non-activating conformer of Tyr10 clearly seen in the Sh28GST structure has never been reported before for any of the several GSTs containing Tyr at this position. In Sh28GST, the two conformers of Tyr10 illustrated above have similar occupancies in all the independent derivatives.

To investigate the conditions which might alter the distribution of the two conformers of Tyr10, four more structures were collected at the synchrotron. The structure of the GSH-free Sh28GST at pH 7.0 was determined starting from the enzyme preparation purified using glycine instead of GSH (see Materials and Methods); starting from the enzyme that was partially GSH saturated, we collected a data set for a crystal soaked at pH 8.0 to rule out the possibility that the bound GSH sulfhydryl had a pK_a of >6.2-6.5 (data not

shown), as reported for other GSTs (Armstrong, 1997); in addition, was solved a structure where Sh28GST was cocrystallized in the presence of 15mM GSH.

The structures in which we increase the presence of GSH were obtained for proteins which have been pre-equilibrated with a different concentration of GSH prior to crystallization. Because the enzymes were crystallized starting from a protein equilibrated with GSH in solution prior to crystallization, it is likely that the distribution of conformers found in the crystal would be present also in solution.

In summary, the three structures at pH 7.0 (GSH-free, partially saturated, and almost fully saturated) exhibit similar relative occupancies of the two Tyr10 conformers (Figure 3.6).

Can the tyrosine swing in the crystal structure? The two conformers of Tyr10 are both stabilized by specific polar interactions (see Results) as well as favorable steric contacts. As shown in Figure 3.6, rotation of the Tyr10 side chain about the C-C bond is blocked in one direction by the main chain of the loop between β-strand 1 and helix 1, and in the other by the rather rigid side chain of Val8, which is on the same β-strand as Tyr10. Thus, it is very unlikely that the two conformers can interconvert in the crystal.

The fourth structure, solved in presence of 15mM of GSH, presents a higher fraction of the Tyr10 in the G-site. This could demonstrate (i) that the tyrosine can swing in solution and (ii) that the H-bond between GSH and the oxygen of the tyrosine could stabilize this conformer, when the protein is fully saturated with GSH.

Indeed, in solution, large conformational changes between apo and GST-bound forms of several GSTs have been described (Ricci et al., 1996; Stella et al., 1998; Hitchens et al., 2001). In the human P1-1 pi class GST, a

disulfide bond can be formed between two cysteines which are 18 Å apart in the crystal structure (Ricci et al., 1991; LoBello et al., 1998). In the same GST, the short 3_{10}-helix 2 is disordered in the apo form but ordered in the GSH-bound enzyme (Hitchens et al., 2001). In Sh28GST, the short helix 2 is both preceded (Phe-Gln-Asp) and followed by a three-residue (Lys-Pro-Thr) 3_{10}-helix which suggests a similar flexibility. In addition to conformational fluctuations in tertiary structure, a specific structural rearrangement of the G-site upon binding of GSH has been suggested for several GSTs (Ricci et al., 1991).

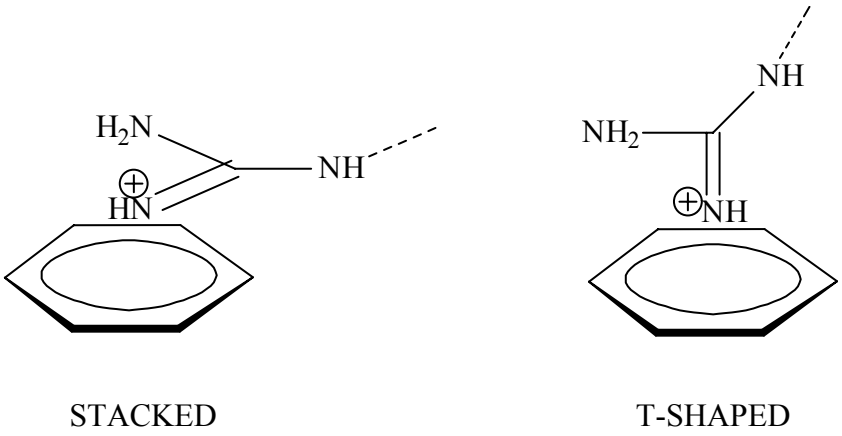
To give an explanation for the different fractions of Tyr10 conformers relatively to the concentration of GSH, we should admit that the two populations of the enzymes (the first with the Tyr10 out of the catalytic site and the second with this amino acidic residue in the G-site) have a different affinity for GSH. Our hypothesis is that there is monomer heterogeneity, and thus a possible negative coopertivity has to be taken into account. We are still carrying out studies to elucidate this point.

3.7 Structural data suggest new mechanisms for the transferase activity and for PGD_2 synthesis

3.7.1 Description of the interaction between Arg21 and the Tyr10.

Several studies have been carried out on this kind of particular interaction that is rarely found in the catalytic site of enzymes (Liu et al., 1993). Electrostatic interactions between cations and π-systems may provide appreciable interactions energies in proteins (Dougherty, 1996). Theoretical and experimental studies, both in gas media and in aqueous media, have shown that π-cation interactions could be quite strong (up to 11 kcal/mol),

although the geometry involved could be quite different and consequently the energies implicated may vary widely (Figure 3.9).



STACKED T-SHAPED

Figure 3.9. *The stacked and T-shaped conformation between an aromatic ring and an Arg.*

The stacked conformation is stabilized through the π-cation as well as additional nonelectrostatic effects such as Van der Waals and π–π cation interactions. The T-shaped geometry points one of the NH groups at the centroid of the ring. T-shaped conformation has been proposed to be stronger than the parallel geometry even if it is more rarely found in proteins and the on-face hydrogen bond is possible only in this T shaped geometry.

Weak interactions between aromatic rings and neutral hydrogen bond donors have also been suggested to contribute to protein structure and stability. However the extent to which these "unconventional hydrogen bonds" contribute to protein function remains unclear. An interaction of this kind is associated to a binding energy of 3-5 Kcal/mol and distances between 3.0-4.5 Å.

It is also clear that if this type of interaction happened between the positive nitrogen and the coniugate base of the tyrosine, the energy of binding would be improved because the electrostatic component of the interactions is raised. Moreover, this could also lower the pK_a of the tyrosine (Figure 3.10).

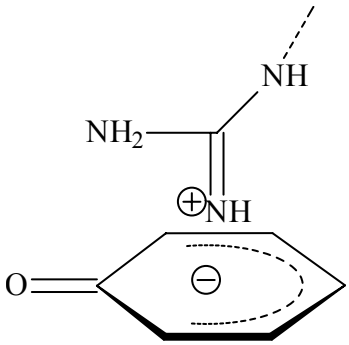


Figure 3.10. *Electrostatic stabilization of the conjugated base of Tyrosine in interaction with Arg.*

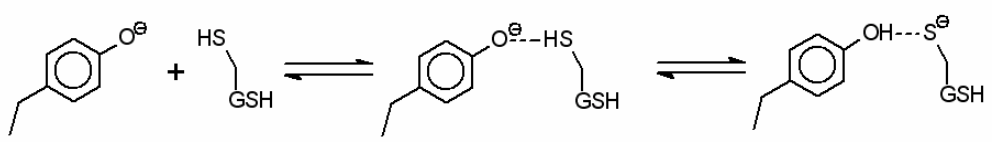
In GST structure, the contact between Arg21 and Tyr10, and that between Arg21 and Phe10 in the Y10F mutant is found at distances that ranging from 3.0 to 3.5 Å and could be described by two contributions: (i) π-cation interaction between the positively charged NH1 of Arg and electron of the benzene ring; (ii) on-face hydrogen bond. Clearly, each contribution strengthens the other.

3.7.2 Transferase (detoxifying) activity

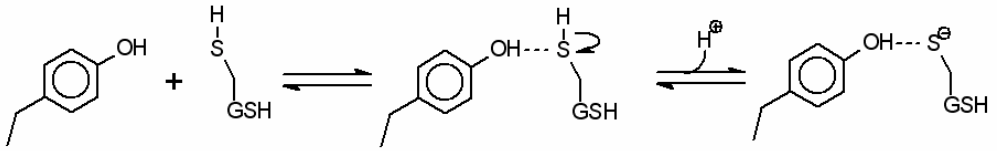
In the literature there are two different mechanisms which account for the activation of GSH, the step preceeding the nucleophilic attack on the substrate (see Scheme 3.2) in other classes of GSTs. The activation reaction is a simple acidic-base exchange but the scenario presented in the literature is

more complex. For example, the two mechanisms proposed for GST A1-1 (a class similar to that of Sh28GST) are: (a) the base-catalyzed mechanism (Scheme 3.3; Dietze et al., 1996) and (b) the lowering of the pK_a of GSH (Scheme 3.4; Bjornestedt et al., 1995).

Scheme 3.3



Scheme 3.4



In the first mechanism (Scheme 3.3) the ionized tyrosine acts as a base extracting the proton from GSH. Several results are in agreement with the base-catalyzed mechanism. First of all, the tyrosine has an unusually low pK_a (8.1-8.3) compared to that of a tyrosine in solution (~10); moreover both the unprotonated and protoned states of the tyrosine were shown to be present and finally the protonation of the tyrosine seems to be dependent on the presence of GSH.

In the second mechanism (Scheme 3.4) the tyrosine is protonated and works as a hydrogen bond donor towards the sulphur atom of GSH, lowering its pK_a. The experimental data supporting this mechanism are: the pK_a of the bound GSH is about 6.8 (while in solution the pK_a is about 9.8); and the

catalytic activity (K_{cat}/K_m) decreases whenever the pH is raised and thus the active form cannot be the unprotonated species.

Anyway, both mechanisms cannot clarify the fate of the GSH proton nor the amino acid that accepts the released proton.

We now provide a functional significance to the presence of Tyr10 in double conformation in Sh28GST and try to fit our proposed mechanism with the experimental data reported in literature for GST A1-1. We remind that in the conformer with the tyrosine out of the binding site, the phenolic oxygen is in contact with the solvent. Considering the nature of the interaction between Arg21 and Tyr10, a unique mechanism for the release of the GSH proton comes to mind whereby the swing of the Tyr10 is coupled to H^+ dissociation.

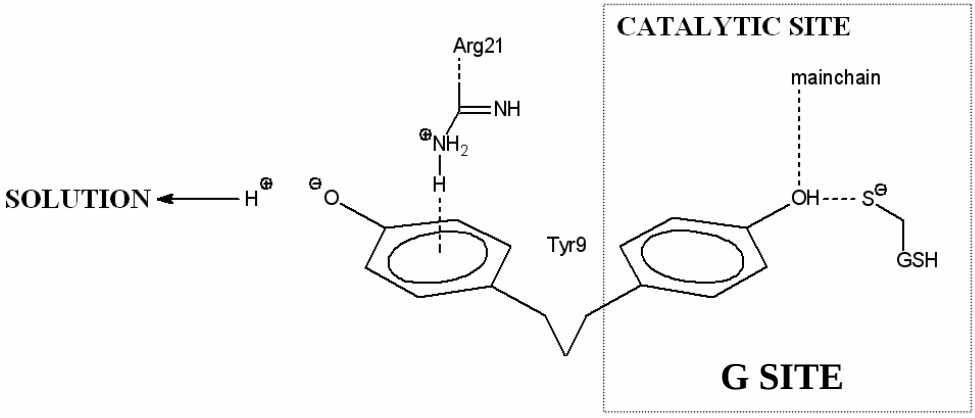


Figure 3.11 *Mechanism for the release in solution of the GSH proton for Sh28GST*

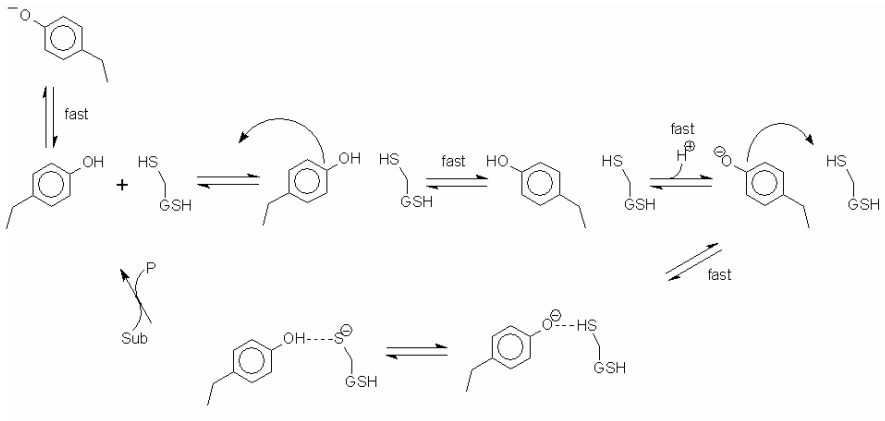
As shown in Figure 3.11 the tyrosinate in the G-site can extract the proton from the bound GSH; transition of Tyr10 out of the catalytic site and interaction with Arg21, lowers its pK_a leading to H^+ dissociation.

This mechanism can account also for experimental data on GST A1-1, i.e.: (i) the low pK_a for Tyr which may be proper for the non-catalytic conformer; (ii) both tyrosine states are present (protonated and unprotonated); (iii) the fraction of protonated tyrosine is directly related to the presence of GSH in the active site; (iv) the decrease in catalytic activity with increasing pH, which stabilizes the tyrosine conformer out of the binding site and hinders its return to the G-site.

Interestingly, such a mechanism whereby a tyrosine can work as a proton shuttle, changing its acidic properties through interaction with Arg, has not been reported

On the assumption of a rapid equilibrium between the two conformers, the following hypothetical mechanism (Scheme 3.5) can be drawn:

Scheme 3.5



3.7.3 Comparison with other GSTs.

Interestingly, Arg21 is totally conserved in all classes of GST which share a tyrosine as a catalytic active site residue (Figure 3.12); whereas it is not conserved in those classes that have a serine or a cysteine as the catalytic residue. Thereby the question arises whether this proposed mechanism could be general for this class of GSTs.

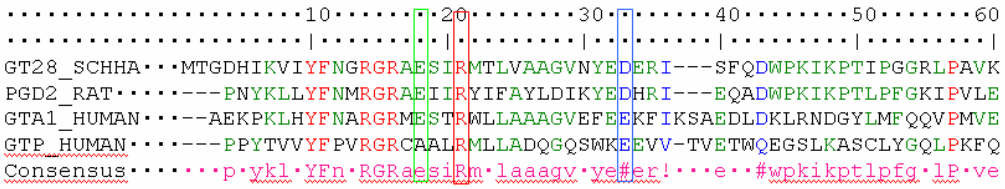


Figure 3.12. *Sequence alignment between the N-term of Sh28GST (GT28_SCHHA), rat Hematopoietic PGDS (PGD2_RAT), human GST A1-1 (GTA1_HUMAN) and human GST P1-1 (GTP_HUMAN). The coloured boxes mark the residues affecting the stability of the non-activating conformer of the enzymes. The red box shows the total conservation of the residue (Arg21), the blue the conservation of the chemical properties (Asp/Glu in position 33), the green the partial conservation of the residue (Glu in position 18, which is not conserved only in human GSTP1-1).*

Since Arg21 helps stabilizing the non-activating conformer of Tyr10, we suggest that this conformer may not be unique to Sh28GST. In fact, inspection of other GST structures (McTigue et al. 1995; Ji et al., 1995; Oakley et al., 1997) reveals that the side chain of the Arg topologically homologous to Arg21 displays a conformation similar to that seen in Sh28GST (Figure 3.13).

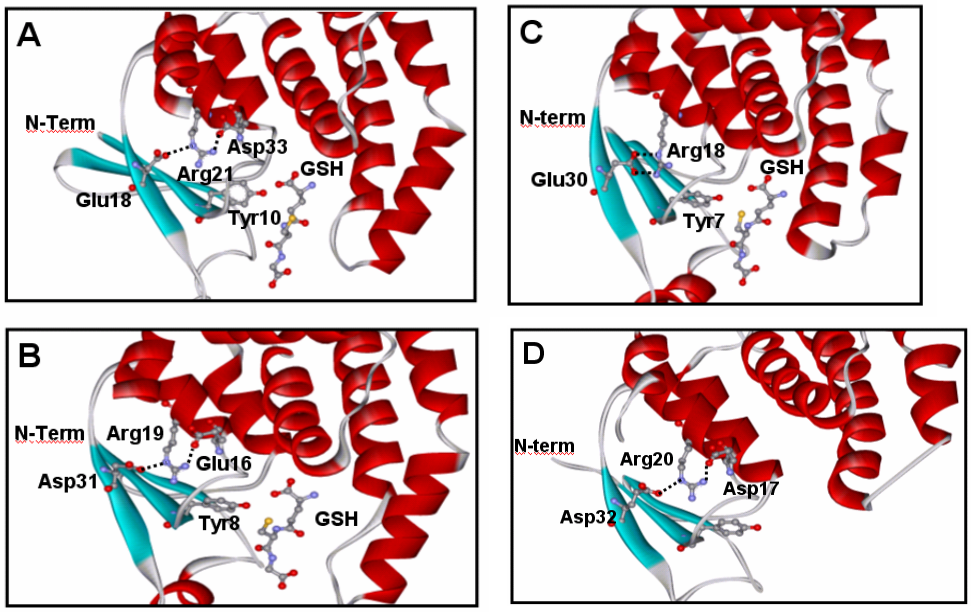


Figure 3.13 *Structural comparison of the N-terminal portion of Sh28GST (Panel A), H-PGDS (Panel B), human GSTP1-1 (Panel C) and human GSTA1-1 (Panel D).*

In all GSTs containing Tyr as the catalytic residue, the conserved Arg21 is kept in place by interactions with an Asp and a Glu, as observed in Sh28GST, with the exception of the pi class GST where only one of the two residues that bind Arg through a double H-bond is conserved (see Figure 3.12 and 3.13). In other GSTs, water molecules fill the space near the guanidinium group expected for the nonactivating Tyr10 conformer (if present), requiring a slight rearrangement of structure to make room for the phenoxyl group.

Since the experimental data for the alpha class GSTs are in agreement with our proposed mechanism and given the structural observations reported above, we propose that the mechanism which seems valid for Sh28GST, may be more generally applicable. This is not a minor point if we consider that the

research of GST inhibitors is a fundamental and up-to-date research task. Moreover, the GSTs are detoxifying enzymes, characterized by wide substrate specificity, which makes it intrinsically difficult to design inhibitors (in this case substrate analogues) active towards enzymes from different classes. From a theoretical stand point, if the swing off of the tyrosine is involved in the detoxifying activity, as we propose, it may be profitable to orient the drug design towards the pocket in which the tyrosine is held by the interaction with the conserved Arg (Figure 3.14). On the other hand, if the mechanism is unique to Sh28GST, it may be a specific lead to inhibit the schistosomial enzyme with clearly interesting therapeutical implications.

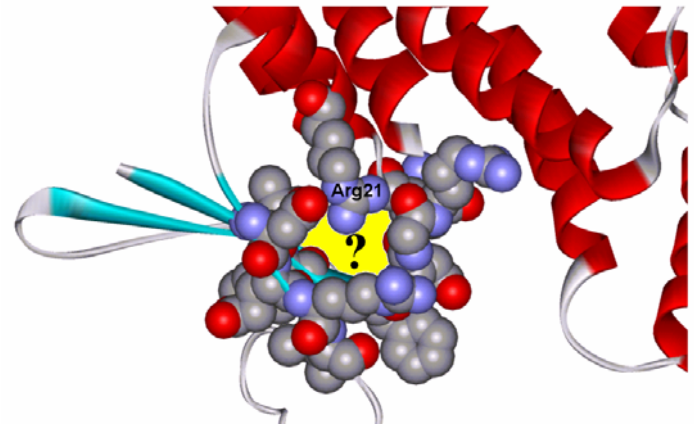


Figure 3.14 *Possible target pocket for drug-design*

3.7.4 Prostaglandin D_2 synthase activity

Since mammalian sigma class GSTs, as well as several GSTs from invertebrates, possess a specific and well-characterized prostaglandin synthase activity (Jowsey et al., 2001), it is tempting to speculate that the non-activating conformer of Tyr10 may play a role in the prostaglandin D_2

synthase activity (PGDS) of Sh28GST (Hervé et al., 2003). The proposed mechanism for PGDS involves two steps (Kanaoka et al., 1997, Figure 3.15).

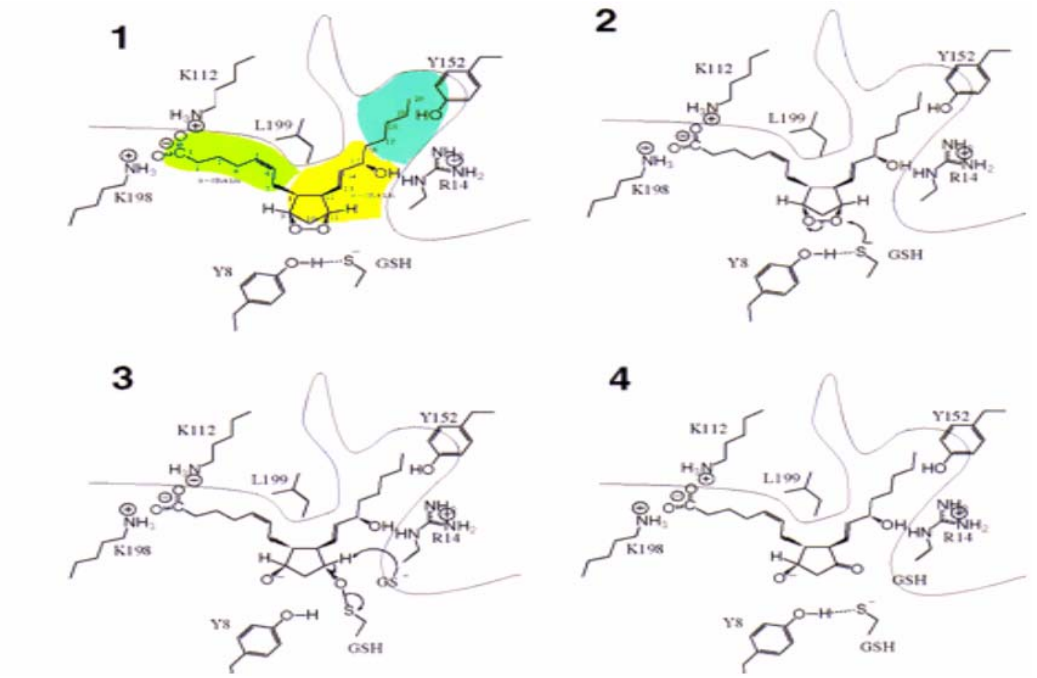


Figure 3.15 *Proposed mechanism for HPGDS (Kanaoka et al., 1997)*

In the first step, GSH attacks one oxygen atom in the peroxide ring of the prostaglandin, thereby breaking the O-O bond and forming a GSH-addition intermediate. In the second step, an unidentified general base (the authors hypothesise that could be a deprotonated GSH, coming from the solution) extracts a proton from the carbon bonded to the first oxygen atom, consequently releasing the product PGD_2 and leaving the GSH bound to HPGDS.

In the case of Sh28GST, the PGDS activity may also be facilitated by the relatively high and unusual affinity of this enzyme for GSH, which would

tend to ensure that the intermediate addition product remains bound to the protein.

In this context, apart from specific aspects of the interaction between bound GSH and tyrosine (any of the tree mechanisms reported above) and assuming that Tyr would loose the proton by swinging out of the binding site, we may propose a different and plausible mechanism as presented in Figure 3.16.

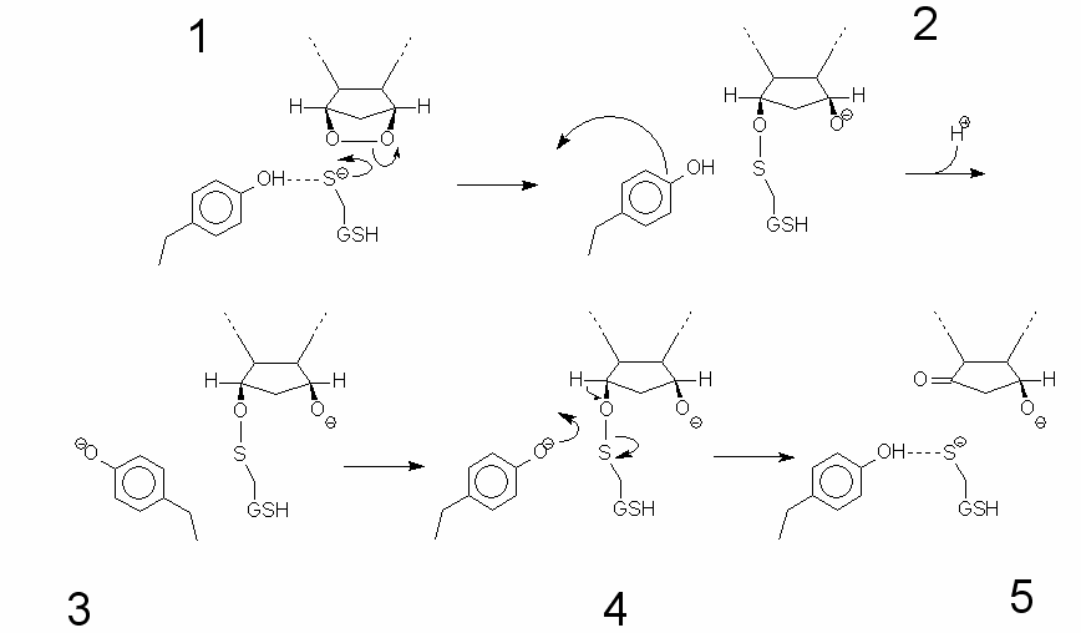


Figure 3.16 *Proposed mechanism for the isomerization of PGH_2 to PGD_2 for Sh28GST*

The hypothesis is based on the assumption that for Sh28GST as well as for HPGDS, the two conformers of the catalytic tyrosine are active and interchanging (see Figure 3.13); thus we propose that the tyrosine comes back into the G-site to extract the proton from the PG-GSH adduct (Figure

3.16). This makes it unnecessary to summon a second molecule of GSH in the active site to outdraw the proton from the PG-GSH adduct,

If this is the case, the design of new inhibitors of the detoxifying activity of Sh28GST, modelled on the Arg21-pocket (Figure 3.14), could be extended to the PGDS activity.

CHAPTER 4

4. THE FATTY ACID BINDING PROTEIN FROM *SCHISTOSOMA MANSONI*

4.1 Fatty acid binding proteins

Mammalian fatty acid binding proteins (FABPs) were discovered in the early 1970s as abundant cytoplasmatic proteins which bind long-chain fatty acids. The FABPs constitute a multigenic family of low molecular mass proteins (14-15 kDa). Analysis of their 3-D structures revealed basic structural similarities among the members of this family despite low sequence identities (Storch and Thumser, 2000). The precise physiological role of FABPs is still unclear and different hypotheses have been proposed, including protection of cell membranes and enzymes from the effect of high concentrations of free fatty acids (FAs) and their acyl-CoA derivatives; storage of FAs; lipids trafficking; regulation of cell growth and differentiation (Veerkamp et al., 1991; Weisiger, 2002). Ambiguities come mainly from the observation that these proteins bind with high affinity several hydrophobic substrates and reversibly associate with phospholipid bilayers (Kaikaus et al., 1990). Among these hypotheses, the involvement of FABPs in the intracellular trafficking of lipids to and from cell membranes is the most accredited; an important piece of evidence is that the heart fatty acid uptake is decreased in heart-fatty acid binding protein gene-ablated mice (Murphy et al., 2004).

4.2 *Schistosoma mansoni* fatty acid binding protein: Antigenic properties.

Sm14 is the second more studied vaccine candidate after GSTs. This is also evident when considering the number of published papers about the *Schistosoma mansoni* fatty acid binding protein (Sm14).

In 1996 (Tendler et al. 1996) molecular cloning of components of protective antigenic preparations has suggested that related parasite fatty acid-binding proteins could form the basis of the protective immune cross-reactivity between the parasitic trematode worms *F. hepatica* and *S. mansoni*. Molecular models of the two parasite proteins showed that both molecules adopt the same basic three-dimensional structure, consisting of a barrel-shaped molecule formed by 10 anti-parallel β-pleated strands joined by short loops, and revealed the likely presence of cross-reactive, discontinuous epitopes principally derived from amino acids in the C-terminal portions of the molecules. A recombinant form of the *S. mansoni* antigen, rSm14, protected outbred Swiss mice by up to 67% against challenge with *S. mansoni* cercariae in the absence of adjuvant and without provoking any observable autoimmune response. The same antigen also provided complete protection against challenge with *F. hepatica* metacercariae in the same animal model. The results suggest that it may be possible to produce a single vaccine that would be effective against at least two parasites, *F. hepatica* and *S. mansoni*, of veterinary and human importance, respectively.

With a view to producing peptides capable of inducing a protective immune response against *S. mansoni* and *F. hepatica*, the sequence and structure of the protective antigens Sm14 and Fh15 were analyzed (Vilar 2003). Their C-termini showed a high level of sequence conservation which, together with models for their three-dimensional structures, aided in peptide selection. Vaccination trials in Swiss mice challenged with *S. mansoni* cercaria or *F. hepatica* metacercaria showed that peptides which included the sequences VTVGDVTA or EKNSESKLTQ were capable of inducing levels of protection equivalent to the recombinant form of Sm14. These peptides may

represent an alternative to r-Sm14 for the development of a bivalent anti-helminth vaccine.

Recently, the T helper cell 1-mediated immune response elicited by Sm14 has been associated with the resistance to schistosomiasis in individuals from endemic regions of Brazil. For this reason, the study of the *S. mansoni* gene structure and polymorphism of Sm14 is of pivotal importance to identify the isoform better suited to devise an efficient vaccine. The data presented in the work of Ramos et al. (2003) indicate that the sequence of the various Sm14 proteins is relatively conserved among the American strains of *S. mansoni* and provide the first experimental evidence for the existence of a reduced polymorphism, especially with respect to *S. japonicum* FABP (SjFABP). In particular, the single mutation M20T is the principal example in terms of recurrence. Interestingly, this change induces important structural and functional modifications in the protein that can have direct consequence in the development of a schistosomiasis vaccine based on this FABP. In fact, the more structurally stable Sm14-M20 isoform appears to be the better vaccine candidate.

4.3 Materials and Methods

4.3.1 Cloning of full-length Sm14.

S. mansoni adult worms were freshly obtained by perfusion of mice, infected at least 7 weeks before, using HEPES-buffered RPMI-1640 medium and containing 100 U/ml of heparin. Parasites, suspended in a minimal amount of medium, were frozen in liquid nitrogen and ground to a fine powder. Total RNA was extracted in Trizol reagent (Invitrogen) of which 5 μg were treated with SuperScript II reverse transcriptase (Invitrogen) to synthesize cDNA,

following the protocol recommended by the manufacturer. To obtain the Sm14 full length coding region, a PCR reaction was carried out on the cDNA template using the forward primer 5' –GTTGAAACATATGTCTAGTTTCTTGGGA– 3' and the reverse primer 5' –TTGCTCGAGTTAGGATAGTCGTTTATAATT– 3'. The PCR reaction was carried out using 1µl of template, 50pmol of each primer and 2.5U of Pfu DNA Polymerase (Stratagene). The PCR product obtained was subcloned into pPCRscript Amp SK(+) (Stratagene) after which the plasmid was sequenced with an ABI PRISM 310 Genetic Analyzer using the BigDye Terminator Cycle Sequencing Kit (ABI). The same PCR product was then cloned into the Gateway His-tagged *E. coli* expression vector, pDEST17 (Invitrogen), following the protocol recommended by the manufacturer. Expression of recombinant protein was carried out in *E. coli* BL21(DE3)pLysS cells.

4.3.2 Purification of Sm14.

Wet cells of a 3l bacterial growth, were resuspended in 40 ml of buffer A (0.5M NaCl, 50mM Tris/HCl pH 8.0, 2mM β-mercaptoethanol, 1mM EDTA, 2% Triton X-100 and 1mM of protease inhibitor phenylmethylsulphonylfluoride) (Sigma-Aldrich, all reagents were of analytical grade). Cell lysis was achieved after 5 min sonication at 4 °C. The protein was found to be expressed in inclusion bodies; hence, after centrifugation at 10000×g and removal of the supernatant, the remaining pellets containing inclusion bodies were cyclically treated three times with the protocol described. Water washing of the residual pellet was employed to remove the excess detergent. The pellets were then dissolved in 50 ml of buffer B consisting of 6 M urea, 20mM Tris/HCl pH 7.8, 500mM NaCl,

5mM imidazole, 20mM β-mercaptoethanol and gently stirred overnight at 4 °C. The resulting solution was filtered with a 0.45μm syringe filter and loaded onto a Ni^{2+}-column (chelating sepharose, Amersham Biosciences). The bound protein was washed with 5-7 column volumes of buffer B without β-mercaptoethanol. A rapid change of solution with buffer C (20mM Na-phosphate pH7.8, 500mM NaCl) resulted in a properly folded His-tagged Sm14, which was eluted at about 250mM imidazole by a linear gradient obtained by mixing buffer C with the same containing 500mM imidazole. SDS-PAGE (Laemmli, 1970) showed the presence of the pure His-tagged Sm14 characterized by a molecular mass of 18.5 kDa. Then, 50mg of pure His-tagged Sm14 were dialyzed against 100mM Tris/HCl pH 8.0, 250mM NaCl, at 4°C, concentrated to 3.5-5 mg/ml followed by addition of 0.1% octyl-β-glucopyranoside. The protein solution was then diluted 1:1 with buffer D (2M urea, 400mM NaCl, 100mM Tris/HCl pH 8.0, 10mM $CaCl_2$, 2mM β-mercaptoethanol) and 10U/ml of thrombin from bovine plasma (Sigma-Aldrich) were added. The resulting solution was stirred overnight at room temperature, then dialysed against 0.1M MES pH 5.5 and loaded onto a cation-exchange column (Source S-15, Amersham Biosciences). The cleaved Sm14 was eluted with a NaCl gradient at about 100mM salt and verified by SDS-PAGE.

4.3.3 Crystallization of the complex with FAs.

Prior to crystallization the protein was equilibrated in 10mM MES pH 5.5, 50mM NaCl, 1mM β-mercaptoethanol. The complexes with the two FAs were formed by adding a 2-fold molar excess of each ligand to Sm14. FAs were added in ethanol to a 1mg/ml solution of Sm14, so that the ethanol concentration did not exceed the 1% v/v. The mixture was gently stirred

overnight at room temperature after which the excess of ligands was removed by washing the sample in 3 kDa-cutoff ultrafiltration devices (Amicon) and the protein was subsequently concentrated to 10mg/ml. Crystallization was achieved using the hanging drop vapour diffusion method, employing drops consisting of 1μl of a well solution of 0.1M MES pH 6.0, 0.2M NaAcetate, 29-31% PEG 8K (wt/v), 5mM β-mercaptoethanol, and 1μl of the protein sample. Rod-like crystals of dimensions 0.4×0.2×0.2 mm grew within few days at 21°C. Crystals were cryoprotected by using a solution containing 10% glycerol (v/v), 30% PEG 8K, 0.1M MES, 0.2M NaAcetate, 5mM β-mercaptoethanol, pH 6.0 and flash–frozen in liquid nitrogen.

4.3.4 Data Collection and Processing.

Data for the two structures were collected at ELETTRA (Trieste, ITALY) at 100K and then processed using the HKL suite (Otwinowsky and Minor, 1996). Both Sm14-OLA and Sm14-ACD complexes crystallized in the same $P2_12_12$ space group with the unit cell dimensions given in Table 4.1. The datasets from 20.0 Å to 1.85 Å for Sm14-OLA and from 20.0 Å to 2.4 Å for Sm14-ACD are complete, well measured, and uniformly distributed in reciprocal space. Solvent volume was calculated, with the method of Matthews (Matthews, 1968), to be around 45% of total crystal volume, and in both cases there was only one complex molecule per asymmetric unit.

4.3.5 Structure solution and refinement of Sm14-OLA complex.

The structure was solved by molecular replacement techniques with the program AMoRe (Navaza, 2001) of the CCP4 suite (CCP4), using the structure of the Muscle-FABP (M-FABP) complex (PDB code 2HMB; Zanotti et al., 1992) as search model after removal of FA and water

molecules. Data between 15 Å and 3.5 Å were used and gave a clear solution with a correlation coefficient of 0.42. Model building and electron density map inspection were performed with the program XtalView (McRee, 1999). For model building both 2|Fo|-|Fc| and |Fo|-|Fc| electron density maps were calculated and contoured at 1σ and 3σ levels, respectively. Most of the loops and the two α-helices were rebuilt. Subsequent rounds of rebuilding and refinement using Refmac (Murshudov et al., 1997), for 13179 reflections between 20.0 Å and 1.85 Å, resulted in a complete model of Sm14-OLA complex, with a R-value of 0.198 and a R_{free} of 0.246 (see Table 4.1). OLA, Met 20 and the side chains of His 14, Glu 110, Asp 124 and Lys 132 were refined with two alternative conformations. Two additional amino acids (Gly and Ser) resulting from the thrombin cleavage of the polyHis tag were found at the N-terminus.

4.3.6 Structure solution and refinement of Sm14-ACD complex.

Since this complex was isomorphic to Sm14-OLA, the difference Fourier method was used to solve the structure, using Sm14-OLA, without OLA and waters as a model. The refinement of Sm14-ACD proceeded in a similar fashion to that described above. The difference density located in the binding cavity, was modelled as C20:4 fatty acid. Final refinement led to an R-factor of 0.213 and an R_{free} of 0.283 for 5746 reflections between 20.0 Å and 2.4 Å. The final structure contained 135 aminoacid residues, 77 water molecules and 1 molecule of ACD (see Table 4.1) and displayed only the side chains of Met 20 and His 14 in double conformations.

4.3.7 Structural analysis.

The models were monitored for geometrical quality by using PROCHECK (Laskowsky et al., 1993). The B-factors for several structural elements were calculated using the subroutine BAVERAGE of the CCP4 program (CCP4). The contacts between FAs, residues and waters were considered up to 4.6Å and measured with the program CONTACT (CCP4). The r.m.s. deviations between structures were calculated with the program LSQKAB (CCP4). Volume and surface of the protein cavity were calculated with Cast-P (Liang et al., 1998). Sm14-OLA and Sm14-ACD structures have been deposited in the PDB with entry codes 1VYF and 1VYG respectively.

4.3.8 Spectrofluorimetry and fluorescence-based FAs binding.

The affinity of Sm14 for FAs was determined by taking advantage of the enhancement of the fluorescence quantum yield of ANS (Sigma-Aldrich) upon binding to Sm14 (Kirk et al., 1996). The experiments were carried out using a Spex Fluoromax spectrofluorimeter at 20°C in 2 ml of 0.1M MES pH 5.5. Sample solutions of Sm14 were prepared by diluting a stock to concentrations of 0.1μM for K_d determinations and to 5μM for stoichiometric experiments. The stock solution (7mM) of the fluorescent probe was prepared in ethanol. Aliquots of appropiately diluted solutions of ANS in 0.1M MES were added to the protein and mixed in a cuvette with a magnetic stirrer for 2 min before collecting spectra. ANS and protein concentrations were verified spectrophotometrically (ε_{280}=12750 $M^{-1}cm^{-1}$ for Sm14, ε_{280}=4990 $M^{-1}cm^{-1}$ for ANS in water). The excitation wavelength was 380 nm and emission spectra were collected over the range of 400-500 nm; emission peaked at 480 nm and readings between 470 and 490 nm were used for quantitative analyses. A 1:1 stoichiometry was achieved and the K_d

determined to be 0.5 μM, without significant dilution of the Sm14, assuming simple equilibrium binding between aqueous ANS and Sm14.

Competition experiments with FA were performed as described for the ANS binding measurements, starting with a protein solution containing 4.5μM ANS as competitor. OLA, ACD, palmitic (PA), myristic (MA), linoleic (LA) and decanoic acid (DA) were purchased from Sigma. All the FAs were dissolved in ethanol to a final concentration of 10mM, then diluted in the experimental buffer prior to use and used within a few hours. The concentrations of PA, LA, OLA and ACD at the end of the titration were much below their CMC determined at physiological pH (Richieri et al., 1992); presumably this applies also to DA and MA, whose CMCs have not been reported. Addition of FA aliquots resulted in quenching of the emission peak of the bound ANS. Titration curves were fitted to a simple replacement equation; the relative errors of the resulting K_d values (Table 4.2) lie within 0.2%.

4.3.9 Kinetic measurements, pH jump and double mixing experiments.
Stopped-flow experiments were carried out using an Applied Photophysics MV18 (Leatherhead, U.K.) apparatus equipped for fluorescence signal detection. The excitation wavelength was 380 nm, and emission was collected using a filter with cut-off ≤ 455 nm. Solutions of Sm14 (0.4μM) in 0.1M MES pH 5.5 were mixed with buffered solutions of ANS at concentrations ranging between 3 and 10μM at 20°C. The time courses were fitted to a double exponential.

The competition measurements were performed in two ways: Sm14 0.4μM was first incubated with ANS 50μM and mixed with different amounts of OLA (1.25-5μM); alternatively the protein was incubated with

OLA (1.5μM) and mixed with different solutions of ANS (100 to 400μM). These experiments were carried out under sub-optimal conditions, given that the transmittance of the samples ranged from 56% to 10%. As a consequence the amplitude of the recorded signal cannot be trusted, whereas the rate constants are reliable.

In pH jump experiments, a weakly buffered Sm14 solution (0.4μM) containing ANS at concentration ranging from 2 to 3.3μM at one pH (10mM MES pH 5.5 or 10mM HEPES pH 7.4) is mixed with an equal volume of a concentrated buffer solution at the reciprocal pH (0.1M HEPES pH 7.4 or 0.1M MES pH 5.5). The final pH of the solutions after the pH jump was assessed by mixing equal quantities of the buffers.

Double mixing experiments were carried out in order to test the time course of the pH dependent conformational change. Sm14 (0.4μM) dissolved in 10mM HEPES pH 7.4 was mixed with 0.1M MES pH 5.5, and the solution was "aged" for different times (ranging from 15ms to 150s). After the pre-set delay, the resulted solution was mixed with an equal volume of a buffer containing 0.1M MES pH 5.5 and 4μM ANS.

4.4 Results

Recombinant Sm14, when expressed in E.coli with a polyHis tag, is insoluble. After solubilisation of inclusion bodies with urea, refolding was achieved by a rapid change of buffer with the protein still bound to the Ni^{2+}-column, to avoid aggregation. A non-ionic detergent (1-octyl-α-glucopyranoside) was added to refolded His-tagged Sm14 to keep the protein in solution. Thrombin digestion yielded a quantitative cleavage of the His-tag, the major element of instability; an additional purification on a cation

exchange-column at pH 5.5 was necessary to obtain homogeneous Sm14. Purified Sm14-apo form did not crystallize despite extensive efforts; however its complexes with ACD and OLA did crystallize under the conditions described in the methods section. Thus the description of the protein structure will refer exclusively to Sm14-FA complexes.

4.4.1 Overall structure

Sm14-ACD and Sm14-OLA were solved at 2.4 and 1.85 Å respectively, with the two models achieving final R-factors of 21,3%and 19.8%.As assessed by the program PROCHECK (Laskowsky,1993; see Table 4.1 for a summary of crystallographic data).

Table 4.1 *Summary of Data collection and Refinement*

	Sm14-OLA	Sm14-ACD
Space group	$P2_12_12$	$P2_12_12$
no. unique reflections	13179	5746
$I/\sigma(I)$	22.3 (4.8)[a]	19.6 (8.8)[b]
completeness (%)	99.3 (92.7)[a]	95.5 (86.9)[b]
redundancy	4.5	5.5
R_{merge}	0.046 (0.38)[a]	0.067(0.17)[b]
R	0.19	0.21
R_{free}	0.24	0.28

[a] Last shell (1.88-1.85Å)

[b] Last shell (2.49-2.40Å)

The main chain and all the atoms of the two complexes of Sm14 are superimposable, with an r.m.s.d of 0.32 and 0.75Å respectively. At the N-

terminus both structures have two partially disordered amino acids, belonging to the linker region after the His-tag (see materials and methods). Numbering of the amino acids starts from the first native residue (Met 1), the two non-native residues being Gly - 1 and Ser 0.

The two complexes have a single molecule of either ACD or OLA, bound in the large internal cavity. The volume of Sm14 cavity (calculated after removing FA and the waters from the models) is ~300Å^3 (see Table 4.2), i.e. comparable to the cavity volumes of other FABPs belonging to the same family (Zanotti et al., 1992). Similarly to the heart group of FABPs (H-FABPs), Sm14 folds as a β-barrel formed by 10 antiparallel β-strands, named from A to J, following the annotation of Sacchettini et al.(1989) (Figure 4.1).

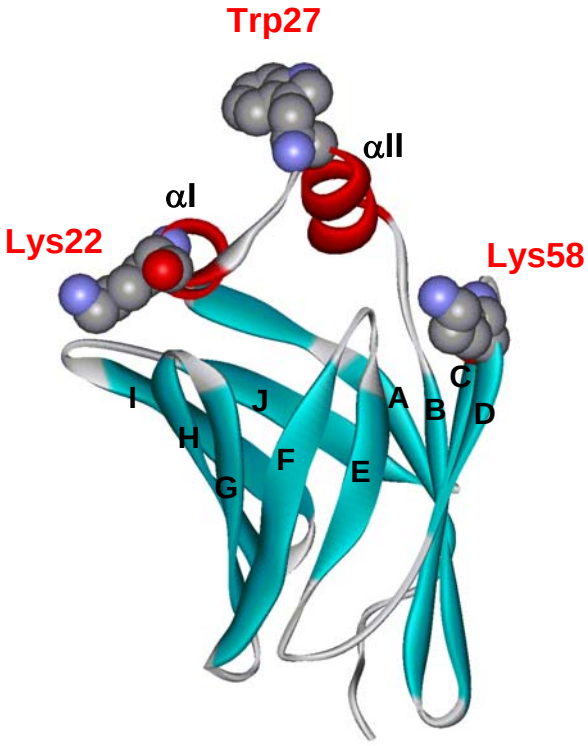


Figure 4.1 *Sm14 overall fold. Secondary structure elements (labelled in black) and the corresponding residues are: Gly 6-His 14 β-strand A, Phe 16-*

Leu 23 α-Helix I, Ala 28-Thr 35 α-Helix II, Thr 39-Asp 45 strand B, Lys 48-Ser 55 strand C, Lys 58-Cys 62 strand D, Phe 70-Lys 73 strand E, Asn 79-Lys 86 strand F, Lys 91-Asp 98 strand G, Asn 101-Asp 110 strand H, Thr 113-Val 120 strand I, Ile 126-Arg 131 strand J.
Putative amino acidic residues that may interact with membranes are displayed in CPK view and labelled in red.

The main chain of each β-strand is H-bonded to the adjacent one, with an interruption of this H-bond array between strands D and E, which are connected via their side chains. Therefore, the structure of Sm14 may be depicted as formed by two separated β-sheets, the first one formed by five β-strands from A to E and the second by those from G to J. The resulting β-barrel is twisted, closed on one side by interaction between the side chains and capped on the other side by a helix(I)- turn-helix(II) motif (Figure 4.1).
The gateway for entry of the substrate in the internal cavity is constituted by the helix(I)-turn- helix(II) motif (residues 16-35) and by two hairpin loops between strands C-D and E-F (residues 56-57 and 74-78, respectively) (Figure 4.1). In both complexes of Sm14, the gateway appears closed, but binding and dissociation of FAs must depend on the opening of this doorway. Phe 57 on the C-D loop, which has been proposed to be a key residue in the binding process (Eads et al., 1993), closes the entry of the lid and is oriented towards the interior of the cavity in both structures. The contact distances between the carbons of Phe 57 and the aliphatic chain of OLA range from 3.5 to 4.5Å; in the complex with ACD the contact distances between C14-C20 and Phe 57 are in the range 3.9-4.5Å. In Sm14-OLA complex, Phe 57 has a higher B-factor (41.1 $Å^2$), compared to the overall B factor of the protein

(36.1 $Å^2$), index of a certain mobility. In Sm14-ACD, Phe 57 has an average B factor similar to that of the whole protein (53.6 $Å^2$).

On the rim of the portal region, Lys 22 at the end of helix (I) and Lys 58 on the loop between strands C-D show the same position and orientation as that of their homologues in H-FABP (Lys 22, Lys 59), where they have been demonstrated to interact electrostatically with phospholipids head groups and supposed to drag the carboxylate of the fatty acid towards the internal cavity (Herr et al., 1996; Figure 4.1). On the external surface of Sm14 (on the back of the two alpha-helices), the hydrophobic side chains of Trp 27 and Ile 32 are exposed to the solvent. This unusual orientation of Trp 27, already predicted for SjFABP (Kennedy et al., 2000) and confirmed by our structures, may help binding to the membrane, inserting among the phospholipids tails (Figure 4.1). Moreover, the work by Storch and Thumser (2000) has clearly identified the alpha-helical regions of muscle and adipose FABPs as being involved in the interaction with membranes. Since both our structures share the same orientation of the same amino acids, we may assume that Trp 27 and Ile 32 in Sm14 have also similar functions.

4.4.2 Structure of bound fatty acids

The internal surface of the cavity is constituted by polar and hydrophobic amino acids, whose side chains are mostly oriented towards the interior of the protein (Figure 4.2 B, D)

Bound OLA is completely surrounded by protein atoms and by structural waters (Figure 4.2 B). The carboxylate is at H-bonding distance with the N(ε) of Arg 127,the oxygen atom of Tyr 129 and two structural water molecules,W12 and W78. W12 is kept in that position by specific H-bonds with N(η)of Arg 107 and the OH of Thr 116,while W78 makes H-bond with

Thr 53. Although the relative orientation of the carboxylate, within these polar residues, is well maintained in most of the FABPs, the geometry of H-bonds in the binding site is often different.

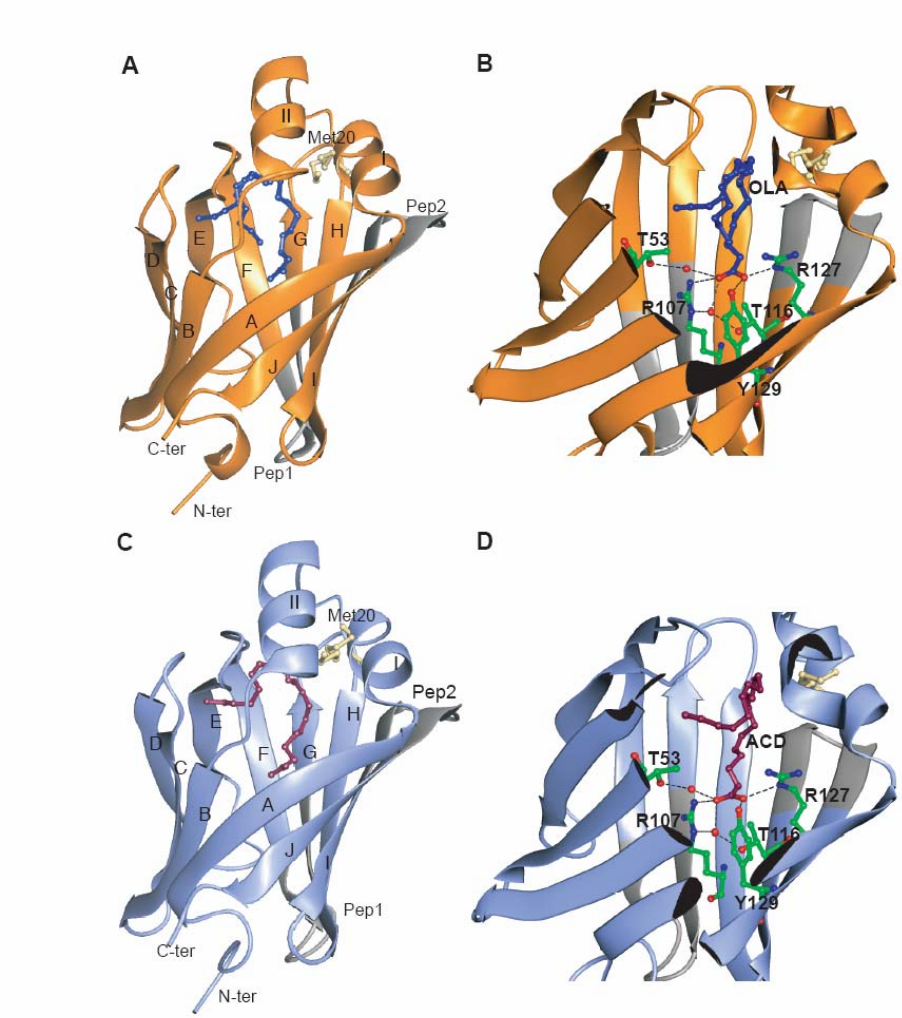


Figure 4.2. *3D structures of Sm14-OLA and Sm14-ACD complexes*

Panel A: Overall structure of Sm14-OLA with secondary structure elements highlighted. Bound OLA (in blue ball and stick) is shown in its double

conformation. Residues belonging to the antigenic peptides are coloured in light grey and Met 20 is shown in magnolia ball and stick.
Panel B: Zoom of the binding cavity with the network of H-bonds fixing the carboxyl head group of OLA highlighted.
Panel C: Overall structure of the Sm14-ACD complex. Strands and helices are labelled as in A. ACD is shown in purple ball and stick.
Panel D: Same as B, for Sm14-ACD complex.

The complex with OLA is found in two conformations, which differ in the orientation of the last five carbon atoms (Figure 4.3). Conformer 1 is found to be in the same U-shaped conformation, previously observed for other FABPs (Zanotti et al., 1992; Eads et al., 1993) and occupies the same portion of the cavity. Conformer 2 displays a chain bending similar to that found for the ACD in the Sm14-ACD complex (see below) in which the last part of the aliphatic chain is out of the plane traced by the rest of the molecule. The difference maps clearly show electron density for both conformations of OLA. The two conformers are compatible with only slight differences in the positions of the first thirteen carbon atoms (Figure 4.3). The relative occupancy of the two is 50%,and their average B factors (43$Å^2$)are above the average B factor of the side chains of the cavity in which they are buried (33$Å^2$). Moreover, in conformer 1 OLA is disordered at its terminus, as indicated by the interruption of the electron density map between C15 and C16. Hairpin conformations are induced by the cis double bond between C9 and C10, by gauche bonds between C5 and C6, C13 and C14, by the large number of Van der Waals interactions with side chains and with 7 ordered water molecules in the conformer 1 and with 6 in conformer 2. The double

bond in C9-C10 interacts in both conformers with the conserved Phe 16, Met 20 and with Val 25, Ser 75, Asp 76, Arg 78.

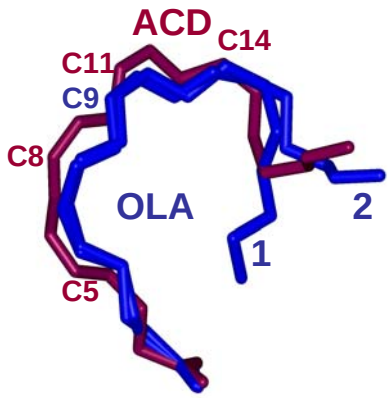


Figure 4.3 *Representation of OLA and ACD bound to Sm14. The partial superimposition of the C9-C10 double bond of OLA (in dark grey) and the C11-C12 double bond of ACD (in light grey) demonstrates a similar configuration of the two FAs in the binding pocket. OLA displays a double conformation (shown as 1 and 2). In conformer 2 the aliphatic chain adopts a fold similar to the ACD in which the last carbon atoms of the tail are out of the plane defined by the rest of the molecule; in conformer 1 the last carbon atom of OLA points towards the carboxylic moiety, yielding a V-shaped conformation.*

Of the 21 residues that contact OLA in conformer 1, 10 are hydrophobic and 11 are polar. In detail: 8 residues belong to 4 of the 10 β-strands (Gln 96 on G β-strand, Thr 103, Ile 105, Arg 107 on H, Thr 116, Val 118 on I and Arg 127, Tyr 129 on J), 3 of these amino acids (Arg 107, Arg 127 and Tyr 129) constitute the binding site of FA carboxylate, 2 residues (Val 36, Pro 38) come from the loop that connects helix(II) with B β-sheet, 4 lie on the mobile

loops that form the rim of the β- barrel (Phe57 in C-D loop, Ser75, Asp76 and Arg78 in E-F loop)and the remaining 7 belong to the helix(I)-turn-helix(II) motif (Phe 16, Met 20, Leu 23, Val 25, Thr 29, Ile 32, Gly 33).

ACD, like OLA, is entirely buried within the cavity with a 100%occupancy.The electron density map is complete without interruption. Interestingly ACD is characterized by an average B factor (37$Å^2$) lower than the surrounding polypeptide chain (48$Å^2$), indicating a strong interaction. The carboxylic moiety makes the same kind of contacts as OLA (Figure 4.2 D). ACD assumes a conformation similar to OLA conformer 2, making the same contacts, plus some additional ones. The contacts between ACD and the side chains of Thr 53, Ser 55, Lys 58 and Leu 60, characteristic of OLA conformer 2, are maintained, whereas the contact between Ile 32 and C13 is present only in OLA conformer 1.Moreover, the long chain of ACD and the presence of four cis double bond rigidify its structure increasing the convex shape and consequently the contact surface between the protein and the lipid, as compared to OLA (Figure 4.3). Several specific sets of contacts with amino acid side chains and water molecules stabilize the four double bonds (i.e. Phe 16,Met 20,Leu 23, Val 25,Thr 29,Gly 33,Phe 57,Ser 75,Asp 76,Arg 78,Gln 96,Thr 103,Ile 105,W26,W39). It is remarkable that in both ACD and OLA complexes, the residues Met 20 and His 14 adopt two alternative side chain conformations; in one of those Met 20 is closer to the double bond of the bound FA (C9-C10 for OLA and C11-C12 for ACD).

The double bonds at C9-C10 in OLA and at C11-C12 in ACD are found in the same position (Figure 4.3), interacting with the same amino acids but in closer contact in the ACD adduct. This is indicative of how different FAs arrange their aliphatic chain for a better fitting into the same protein pocket.

Relative to the overall plan defined by the lipid, the first two double bonds (C5-C6, C8-C9)of ACD are out of the plane, while the last two lie in it. Examination of the protein-ligand interactions shows that specific polar contacts stabilize the first two cis-double bonds in that out-of-plane orientation (Figure 4.4).

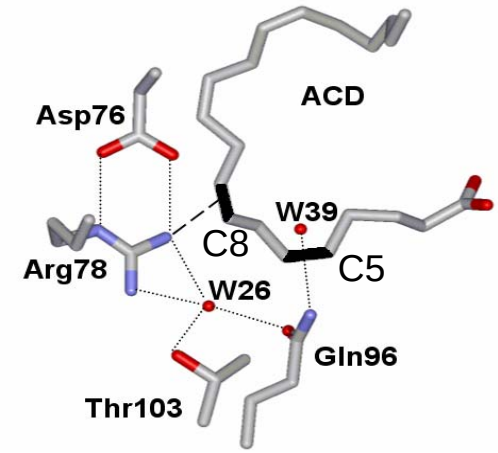


Figure 4.4 *Network of contacts in the binding site of the Sm14-ACD complex. The indicated residues interact with the double bonds of ACD in C5 and C8. The same overall architecture is conserved in Sm14-OLA complex.*

It is interesting to note the continuous contact surface generated by the side chains and the two waters interacting with the first two double bonds of ACD. This architecture is maintained in Sm14-OLA structure but it is not suitable to allow specific interactions with a C18:1 FA. Previous studies showed that the couple Arg 78-Asp 76 in the E-F loop is one of the structural elements that help the binding process in H-FABP family (Hodson and Cistola, 1997). Another residue, Thr 74, at the beginning of the same loop, takes part to this interaction making a H-bond with Asp 76, without directly contacting the FA. This network of H-bonds could help bending the E-F loop

towards the interior of the protein thus stabilising FA through a strong directional π-cation interaction between Arg 78 and C8-C9 double bond (in Figure 4.4). A consistent observation is given by the B-factors of the E-F loop in the two Sm14 complexes. In the case of Sm14-OLA, in which the interaction between FA and Arg 78 is looser, the B-factors are greater by about 10$Å^2$ with respect to the mean-value of the whole protein, whereas in the Sm14-ACD structure, they lie in the same range as the overall structure.
In conclusion, the binding pocket is relatively rigid and does not change upon binding different ligands; OLA and ACD accommodate in the pocket with respectively lower or higher shape complementarity, due to their geometry.

4.4.3 Ligand binding properties of Sm14

We studied equilibrium and kinetics of complex formation between Sm14 and several FAs using the environment-sensitive fluorescent probe ANS. The crystallographic structure of Adipocyte FABP (A-FABP)-ANS complex shows that this probe is located in the FAs binding site with the sulfonate moiety placed between the two arginines involved in the FA's carboxylate binding site (Ory et al., 1999). The functional properties of Sm14 were characterized in 0.1M MES pH 5.5 for two reasons : (i) crystals of Sm14 were obtained at a similar pH (see Matherials and Methods) hence granting a better comparison between functional and structural data and (ii) at this pH the protein is stable for a long time. Under these conditions Sm14 binds ANS with a 1:1 stoichiometry and high affinity (K_d =0.5 μM).For comparison, the Kd values reported for A-FABP and Intestinal-FABP (I-FABP9 ranged from 1 to 50 μM, in several conditions (Ory et al.,1999;Kirk et al.,1996).
Addition of FAs to the Sm14-ANS complex quenches emission at 475nm, indicating that FAs displace ANS. The best fit analysis of the titration data is

achieved assuming a single FA molecule binding. The values of K_d calculated for 6 different length FAs are reported in Table 4.2.

Table 4.2 *Equilibrium constants for the dissociation of six fatty acids from Sm14, and cavity volumes of Sm14-OLA and Sm14-ACD.*

Fatty acid		CMC[a] (nM)	K_d (nM)	Cavity Volume ($Å^3$)
Decanoic acid	(C10:0)		6100	
Myristic acid	(C14:0)		88	
Palmitic acid	(C16:0)	4,000	33	
Oleic acid	(C18:1)	6,000	9	318.9
Linoleic acid	(C18:2)	13,000	24	
Arachidonic acid	(C20:4)	20,000	10	300.1

[a]Critical Micellar Concentration (Richieri et al., 1992)

The results show that formation of the complex between Sm14 and OLA or ACD have essentially the same dissociation constant: i.e. 9 and 10 nM respectively. The values for decanoic, myristic, palmitic and linoleic acids are respectively 610, 9, 3 and 2.5 fold greater than OLA and ACD. The maximum ANS quenching found in our experiments occurs at concentrations much below the FA critical micellar concentrations (CMC) as determined at pH 7.4 under physiological conditions (Richieri et al., 1992).This is very important since above the CMC the equilibrium becomes heterogeneous. The high affinity of FABPs for their ligands is explained because: (i) FAs are poorly soluble in water, hence there is a favourable free energy change

associated with their transfer to the less polar protein interior;(ii) FAs establish extensive, though weak, contacts with the residues lining the internal cavity of FABPs, accounting for the enthalpic contribution to binding; and (iii) given pK_a= 4.7 for the long chain FAs, all molecules are predominantly in the anionic form, both at pH 5.5 and at pH 7.4, thus favouring the interactions with the positively charged guanidine group of Arg 127. We attempted ANS replacement experiments by addiction of prostaglandins, but failed to detect any fluorescence change, proving that Sm14 does not bind these derivatives of ACD, even at concentrations higher than those prevailing under physiological conditions in human body. The differences between the K_ds reported in Table 4.2 at pH 5.5 and older data (Moser et al., 1991) could be rationalised by the difference in pH. Indeed when the same ligand titrations were carried out at pH 7.4, we observed an increase in K_d, by a factor of 10 or more. This reduction in affinity as the pH is increased applies to both ANS and FAs and may be of physiological relevance, helping releasing bound FAs (see below).

The kinetic rate constants for binding and release of ANS and oleic acid were determined by fluorescence stopped flow; indeed when dealing with a carrier protein displaying such a high affinity for its ligands, the question arises whether release occurs to an extent and at a rate compatible with physiology. The combination of Sm14 with ANS (at pH 5.5) is fast, and its second order rate constant is $\geq 2 \times 10^8$ M^{-1} s^{-1}, thus approaching the limit set by diffusion. Given the affinity of ANS (K_d =0.5 μM) and the sensitivity of the instrument, it is impossible to dilute the ligand sufficiently to obtain a more precise estimate. The kinetics of dissociation of ANS bound to Sm14 can be followed by displacement, mixing with an excess of OLA. The time courses recorded have half times of ~1 s and are best described by two exponentials,

irrespective of the concentration of OLA. Although it is obvious that the slow fluorescence decrease observed is associated with the release of ANS, it is unclear whether the reaction involves a ternary complex with both ANS and OLA transiently bound to the protein cavity. Furthermore, the ratio between the overall k_{on} and k_{off} does not equal the equilibrium constant, suggesting a complex reaction mechanism.

Release of bound OLA was achieved by rapidly mixing the complex with 70 to 1000 fold excess of ANS (Figure 4.5).

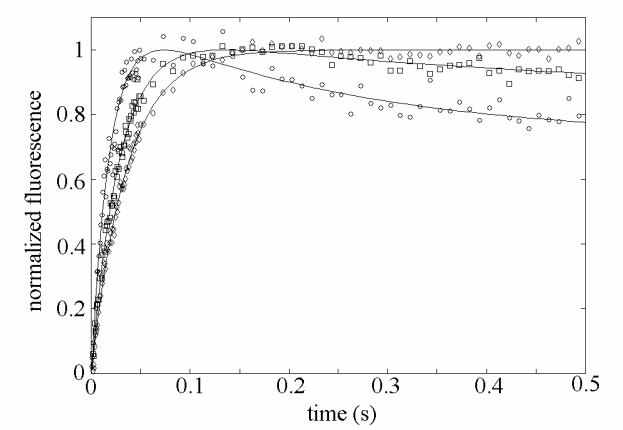


Figure 4.5 *Time course of oleic acid replacement by ANS. Sm14 (0.4μM) incubated with OLA (1.5μM) in Mes pH 5.5 is rapidly mixed with ANS buffered solutions at concentrations of 100 (◊), 200 (□) and 400 μM (○). Experimental points are fitted with a model including a ternary complex between Sm14-OLA-ANS. The slow decrease in fluorescence signal (always present but best seen at the highest ANS concentration) is the dissociation of OLA from the ternary intermediate.*

Even at its lowest concentration, ANS is able to displace over 50%of the bound OLA. The time course is complex and involves, as an unstable

intermediate, a ternary complex of Sm14 with both ANS and OLA in the pocket, from which OLA would then dissociate. Accordingly, the main fluorescence increase could be assigned to the second order formation of the ternary complex (k_{on} =3×10 ^{6} M^{-1} s^{-1}), whereas the subsequent fluorescence decrease (k_{off}=3 s^{-1}) would be assigned to the dissociation of OLA from the ternary complex and the coupled entrance of water in the pocket. This hypothesis is strengthened by data from *Echinococcus granulosus* FABP (EgFABP), a protein with high sequence identity with Sm14, which binds more than one ligand at pH 7.4 (Alvite et al., 2001).

The rate of FA release from Sm14 is independent of the ionic strength of the medium. Consistently, Sj-FABP discharges the ligand to acceptor membranes by a collision mechanism via a modulation of the membrane charge without being influenced by the salt concentration (McDermott et al., 2002).

Since the affinity of Sm14 for its ligands depends on pH and decreases by over an order of magnitude as the pH is raised from 5.5 to 7.4, it is possible to induce ANS binding or release by rapidly changing the pH of the solution. Indeed if Sm14-ANS complex in diluted buffer at pH 7.4 is rapidly mixed with concentrated buffer at pH 5.5, binding occurs; if the opposite pH jump is realized, the dissociation of ANS can be followed (Figure 4.6).

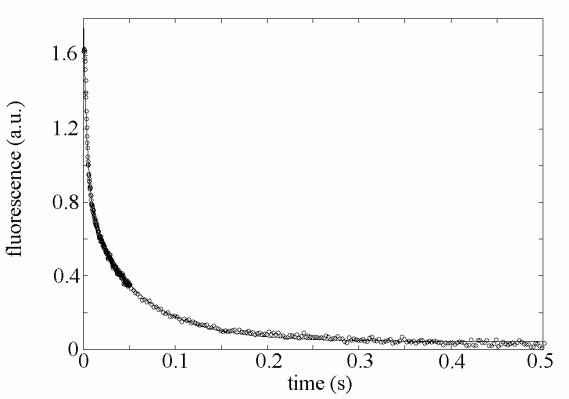


Figure 4.6 *Dissociation of ANS, following the pH jump 5.5→7.4, monitored by a decrease in fluorescence. A solution of 0.4μM Sm14 containing 2μM ANS in 0.01M MES buffer pH 5.5 was mixed in the stopped flow apparatus with an equal volume of 0.1M HEPES buffer pH 7.4.*

Unexpectedly, after a pH-jump the change in ligand affinity occurred at a relatively slow rate. Indeed when Sm14 equilibrated in dilute HEPES buffer at pH 7.4 was rapidly mixed with a concentrated buffer at pH 5.5 containing ANS (Figure 4.7, Trace 1), the binding time course was slower than that observed with Sm14 equilibrated at pH 5.5 and mixed with ANS at the same pH (Figure 4.7, Trace 2).

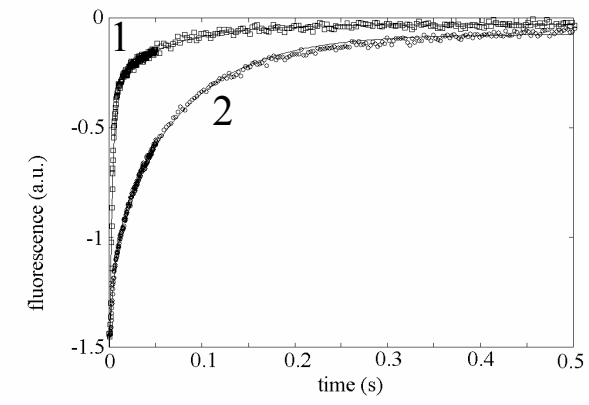


Figure 4.7 *Trace 1: Binding of ANS at pH 5.5, monitored by an increase in fluorescence, is fast: a solution of 0.4μM Sm14 in 0.1M Mes buffer pH 5.5 was mixed with an equal volume of a solution of 3.3μM ANS in the same identical buffer. Trace 2: The binding of ANS at pH 5.5 after an instantaneous pH jump 7.4→5.5 is quite different from trace 1, because the protein is still in the slowly reacting pH 7.4 conformational state; this was obtained by mixing a solution of 0.4μM Sm14 in 0.01M Hepes buffer pH 7.4 with an equal volume of a solution of 3.3μM ANS in 0.1M Mes buffer pH 5.5.*

This shows that the low-pH conformation, corresponding to the fast reacting state, was not reached during this simple pH jump experiment. Thus to determine the rate constant of the conformational change we carried out a double mixing experiment, in which Sm14 equilibrated in diluted buffer at pH 7.4 was mixed with concentrated buffer at pH 5.5 and then, after a variable delay, with ANS dissolved in buffer at pH 5.5.All time courses could be described by two exponentials whose rate constants are independent of the final pH, the relative amplitude of the processes corresponding to the two kinetic phases, being the only parameter that varies with pH.

The results of this analysis are showed in Figure 4.8. The rate of the conformational change is 0.1 s^{-1}.

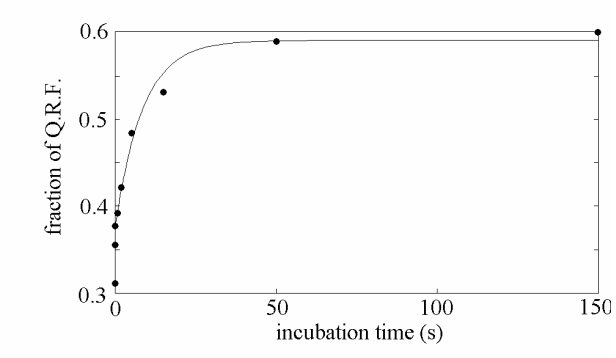


Figure 4.8 *Fraction of the quickly reacting form (Q.R.F.) of Sm14 as a function of the incubation time at pH 5.5 after a pH jump 7.4→5.5. A solution of 0.4μM Sm14 in 0.01M Hepes buffer pH 7.4 was first mixed with an equal volume of 0.1M Mes buffer pH 5.5, and than, after variable incubation time, with a double volume of a solution of 4μM ANS in 0.1M Mes buffer pH 5.5. The time courses of fluorescence increase (indicating complex formation) were fitted to double exponentials under the assumption that the*

two kinetic rate constants were independent of the incubation time, while the relative amplitude of the fast and slow binding events changed. Each experimental point represents the fraction of fluorescence increase assigned to the faster kinetic process. The solid curve drawn through the experimental points is the best fit for a simple exponential relaxation to the equilibrium condition, with $\tau^{-1}=0.1sec^{-1}$.

4.5 Discussion

4.5.1 Structure of Sm14

Sm14 is the only vaccine candidate shown to achieve significant immune protection against schistosomiasis as well as against the helminth infection of cattle caused by *F.hepatica* (Tendler et al., 1996). The protein is present in all the life cycle stages spent by the parasite inside the definitive host, i.e. schistosomulum, adult worm and eggs (Brito et al., 2002).The basal lamella of the tegument and the gut epithelium of the worm are strongly labelled by immunofluorescent probes against Sm14 (data not shown). These tissues have a high flow of lipids, supporting the putative role of Sm14 as intracellular transporter of FAs from host cells (Moser et al., 1991). Therefore a detailed investigation of its structure and function is of crucial importance.

The overall structure of Sm14 is a U-shaped β-barrel closed on top by a lid of helices and on the bottom by side chains, similarly to other FABPs. The agreement with the homology model presented by Tendler et al. (1996) is fairly satisfactory. We shall focus our discussion on some structural features relevant either to the antigenicity or to the physiological function of the protein. Vaccination trials in Swiss mice showed that two peptides derived

from Sm14 are capable of inducing a level of protection comparable to that induced by the full length protein (Tendler et al., 1996; Vilar et al., 2003). The two peptides, namely Pep1 (from 85 to 94) and Pep2 (from 118 to 125), are topologically far from each other in Sm14 crystal structure (Figure 4.2 A,C). Pep1, close to Sm14 bottom, encompasses F and G β-strands including the loop between them; Pep2 is also located in a loop, more precisely a β-turn between strands I and J, close to α I-helix. Most of the side chains of these regions are exposed to the solvent except Leu 92 and Gln 94 of Pep1 and Val 118, Val 120 and Ala 125 of Pep2, all of which point to the interior of the protein. Our structures could explain the antigenicity: both peptides possess a unique secondary feature, which is possibly retained even when removed from the protein; and both appear to be sufficiently different from the homologous human isoforms to be recognized as "non-self". Met 20 (Figure 4.2 A,C) at the beginning of the α I-helix is relevant to the immunogenic properties of Sm14 but it is far from the two antigenic epitopes; thus its effect may be indirect. In both our structures, Met 20 presents a double conformation, with 50% occupancy. In one conformation the sulphur atom lies in proximity of the C9-C10 double bond in Sm14-OLA (3.8Å) and of the C11- C12 double bond in Sm14-ACD (3.7Å). In the other conformation the sulphur atom is found in between the two helices (hence out of the binding site). The latter conformer may play a critical role in stabilizing the closed conformation of the pocket lid when a ligand is bound. When Met 20 was mutated into Ala in two other FABPs (A-FABP and I-FABP), the proteins were characterized by a reduced stability and a decreased affinity for FAs by about 6-8 kcal/mol (Richieri et al., 1998).

Position 20 of the polypeptide chain also presents a natural allelic variant, T20. Ramos and co-workers (2003) studied the effect of mutations of Met 20

in Ala and Thr with respect to antigenicity, and observed a different protective response against *S.mansoni* cercariae in mice. Sm14-M20 displays greater antigenicity than physiological Sm14-T20 and artificial Sm14-A20 (67%, 44%, 20%reduction of worm burden, respectively). Since the three isoforms have different thermodynamic stability, it is possible that this parameter could be amongst the structural determinants of the immune response.

4.5.2 Fatty acid binding:affinity,structure and physiological implications

The favourable free energy of FAs binding may be due to a combination of opposite effects: specific interactions between the protein and FA; water solubility of FA; loss of degrees of freedom in the binding process, with entropic loss. FAs aqueous solubility (CMC in Table 4.2) increases with the level of unsaturation for a given chain length and therefore the affinities might be expected to decrease with the degree of saturation, as it happens for most FABPs, where high K_d values correlate with high water solubility of polyunsaturated long-chain FAs (Figure 4.9) (Richieri et al., 2000).

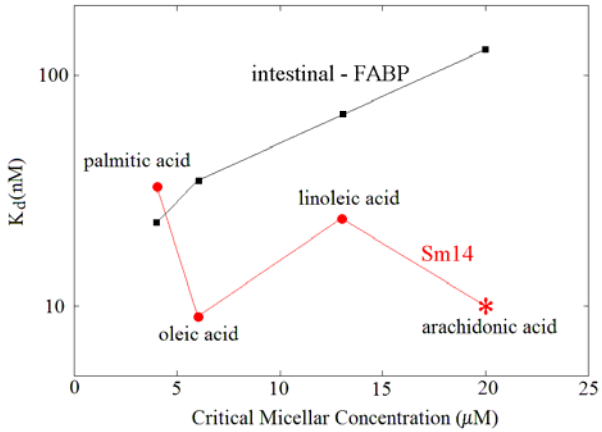


Figure 4.9 *Behavior of the equilibrium dissociation constants (K_d) of intestinal FABP (in black) and of Sm14 (in red) versus critical micellar concentration (CMC).*

This is not the case for Sm14, whose preferred ligands are long-chain unsaturated FAs. In fact the physiologically important long-chain FAs have K_d values in the range 10-90 nM, while binding of shorter FA such as decanoic acid is much weaker. Interestingly, ACD is the tightest ligand for Sm14 despite its relatively high solubility in water (Figure 4.9).

Previous calorimetric experiments on FABPs belonging to the same family as Sm14 have shown that FA binding is mostly enthalpy driven (Balendiran et al., 2000).Therefore the reason for a preference of ACD may lie on the specific contact surface and interactions between the protein and the substrate, the volume of the cavity being in all cases practically unchanged (Table 4.2).

The number of recognition sites between Sm14 and the bound OLA or ACD is similar; the aliphatic chain of OLA has 18 contacts in conformer 1 and 21 in conformer 2, and the aliphatic chain of ACD has 21. Moreover each FA carboxylic group makes 3 further contacts.

The network of interactions between C5-C9 of FAs and Sm14 involves four side chains and two water molecules (Figure 4.4). This network is substantially the same for OLA and ACD, the binding cavity being unchanged in both structures, but in the latter case the contacts are shorter. The C9-C10 and the C11-C12 double bonds of OLA and ACD, respectively, are superimposable (Figure 4.3), making the same set of contacts with Met 20,Val 25,Ser 75,Asp 76. The major difference in binding lies in the conformation of the tails of the two FAs, due to a different way of fitting into the cavity. Compared to ACD, OLA displays a reduced rigidity (showed by

the double conformation and by the higher B factors) and less specific contact sites. The structural analysis strongly suggests that the contributions of enthalpy and entropy in favouring the binding of the two FAs are different. The binding between Sm14 and ACD could result in a larger enthalpic contribution, due to specific contacts and a better shape complementarity (Figure 4.4). Indeed, our data suggest that Sm14 is specifically designed to accommodate ACD. This result is intriguing given that ACD, besides being a nutrient, is the substrate for prostaglandin synthesis, which has been demonstrated to represent an additional way to escape the host immune response through the inhibition of Langerhans cells migration (Angeli et al., 2001).

The affinity of Sm14 for FAs is higher at acidic pH than at neutral or basic pH. The pH dependence of the ligand affinity observed in mammalian FABPs is opposite, e.g. A-FABP shows a higher affinity at alkaline than at acidic pH (Wootan et al.,1993).This finding suggests that the physiological role of Sm14 may be FAs transfer in the reverse direction with respect to that observed in mammalian cells. Moreover, like A-FABP, Sm14 is a basic protein (pI=7.8) and belongs to the family that transfers ligands to/from acceptor membranes with a contribution by collision. Previous work showed that A-FABP discharges ligands rapidly when acceptor vesicles contain acidic phospholipids, and only slowly when membranes are positively charged. Thus in that case a reduction of affinity at acidic pH cooperates with an enhanced collision probability between the protein and acceptor membranes in order to release FA. The authors did not discover the bases for changes in affinity with pH, but did not rule out conformational changes (Wootan et al., 1993). In Sm14, we have kinetic evidence for a slow conformational change induced by a sudden pH change (Figure 4.8). This

effect may be due to a widening of the protein cavity at pH 7.4, causing a reduction of affinity, an observation crucial to understand the role of Sm14. In particular, negatively charged phospholipid head groups (located on the cytoplasmic side of the membranes) lower the local pH at the membrane surface (Eastman et al., 1989), thus triggering the conformational change of Sm14 to the high-affinity state. Hence we deduce that pH change next to anionic membrane phospholipids may have two distinct effects: (i) to increase the affinities of Sm14 for FAs and (ii) to facilitate collision between protein and membranes.

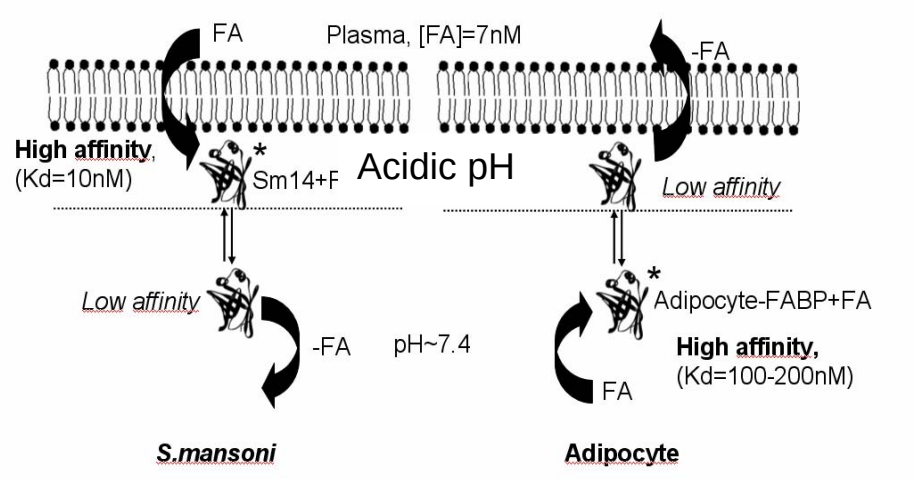


Figure 4.10 *The scheme depicts the hypothetical principal direction of FAs transport in S. mansoni tegument (left) and in mammalian adipocyte (right). The main functional difference between Sm14 and A-FABP (apart from their ligand specificity), is the opposite effect of pH (see text) that is suggested to determine the preferential direction of the FAs trafficking in the two cases.*

Since Sm14 is localized in tissues near the contact with the host, it is reasonable that its main function is to extract FAs from the host's serum,

where the free FA concentration is about 7nM, and to release them in the intracellular medium. This is consistent with the more acidic pH near the cell membrane and hence the higher ligand affinity of Sm14. On the contrary, A-FABP would mainly work in the opposite direction, loading up FAs in the cytoplasm and releasing them near cell membranes, since its pH effect on ligand affinity is reversed with respect to Sm14 (Figure 4.10). Hence, we believe that not only different affinities among different FABP families, but also different pH dependences of affinity are relevant to physiology.

FABPs markedly differ in their isoelectric points (Veerkamp et al., 1991). pI may lead FABPs to preferentially interact with one type of cell membranes, thus contributing to the physiological requirements of the tissue or organism in which they are expressed. Moreover, it has been shown that the uptake of FAs can also occur in close proximity to specific receptors (Weiseger et al.,2002) and in good agreement with our hypothesis, it was recently reported that schistosomes express receptors for human lipoproteins that display acidic properties (Fan et al.,2003).

4.5.3 Comparison with other FABPs

FABPs are grouped according to sequence comparison. In this classification, Sm14 belongs to the same family group as H-FABP [mammalian heart-, muscle-, brain-(B-), adipose tissue-and myelin-FABPs; parasite Sj-,and Eg-FABP, Esteves et al., 1997]. It seems that all these proteins have higher affinity for polyunsaturated FAs with respect to FABPs of other groups. In fact the affinity of H-,B-,M-FABP is in the range 4-7nM for OLA and 18-27nM for ACD. This is in line with our values, although Sm14 affinity for ACD remains the highest of the group. Interestingly, Eg-FABP, coming from another platyhelminth unable to de novo synthesize FAs, shows a similar

trend in affinities versus OLA, LA and ACD (Alvite et al., 2001). One main difference between Sm14 and the H-FABP group is that Sm14 affinity for FAs is maximal at slightly acidic pH, whereas that reported for mammalians is maximal at physiological pH. Several structures of FABPs are deposited in the PDB, most of them bound to FAs. The trend that FABPs (including Sm14) are easier to crystallize in the liganded state, implies that binding of FAs stabilizes the protein. The only two structures deposited in the PDB of FABP complexes with polyunsaturated FAs are human B-FABP (Balendiran et al.,2000)with docosahexanoic acid (DHA, 22:6)and murine A-FABP with ACD (PDB code:1ADL; LaLonde et al., 1994). The equilibrium constant measured for ACD binding to this latter protein was 4.4 µM (by means of calorimetry; LaLonde et al., 1994) and 196 nM (by acrylodated intestinal fatty acid binding protein ADIFAB method, Richieri et al., 2000).

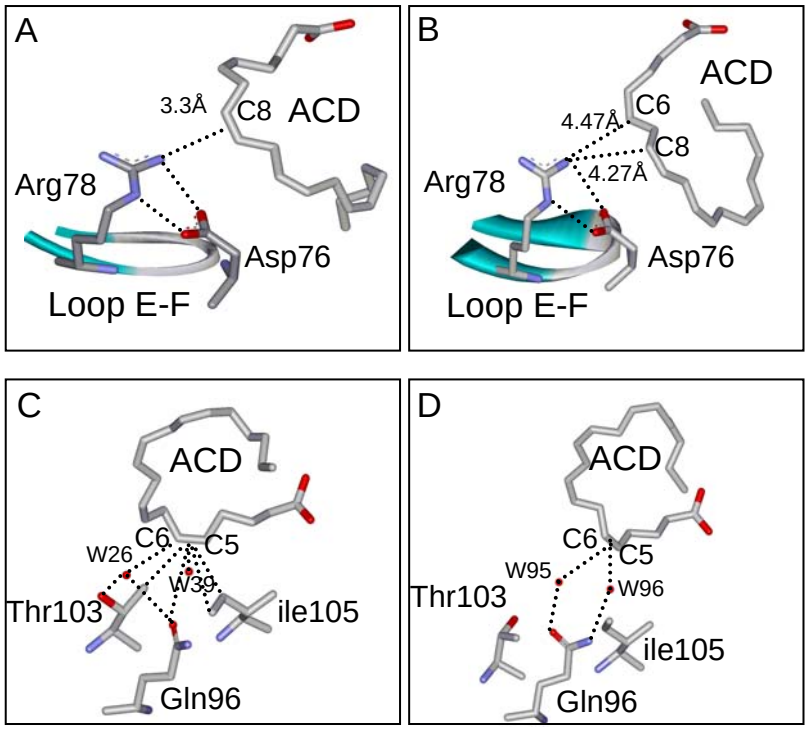


Figure 4.11 *Details of the interactions of Sm14 and A-FABP versus ACD*

A comparison between the complexes Sm14-ACD and murine A- FABP-ACD shows that our complex is characterized by more specific contacts. In our structure, the double bond in C5-C6 of ACD is stabilized by a series of polar contacts involving 3 invariant residues and two water molecules (Figure 4.11 B). Thr, Ile and Gln are conserved in murine A-FABP but only the contacts between two water molecules and C5-C6 double bond are present (Figure 4.11 B).

A further important feature is the π-cation interaction between Arg 78 and C8-C9 double bond, which is assisted by Asp 76 in correctly orienting NH2 of Arg 78 (Figure 4.11 A). Although these two residues are conserved in H-FABPs family, in A-FABP their side chains do not interact in such a precise way with ACD as they do in Sm14 (Figure 4.11 B). Moreover, these two residues are not conserved in non homologous proteins, an indication of how much this sequence architecture is important for a preference of binding polyunsaturated FAs by the H-FABP group. Therefore we can explain the tight binding of Sm14 to ACD with these two specific networks of interactions: those between Gln 96, Thr 103, Ile 105, W26, W39 and C5-C6 double bond; and those between Asp 76-Arg 78 and C8-C9 double bond.

Moreover, the double bonds C5-C6, C8-C9, C11-C12 and C13-C14 are arranged in such a way around Phe 18 to permit π -stacking interactions. From this scenario, we conclude that the lipid-specificity of FABPs, and hence the details of their interactions with FAs, may vary according to the needs of the tissue or organism in which they are expressed. A preferential region for specific interactions in the complexes of Sm14 is that of the

residues in contact with the two first double bonds of ACD (C5 and C8), where the network of interactions is optimized (Figure 4.11).

CHAPTER 5

5. CONCLUSIONS AND PERSPECTIVES

In general, this work contributes to our understanding of aspects fundamental to the pathogenesis of the schistosome and, at the same time, offers new possibilities for the search of a therapy. We focused our attention on the parasite's metabolism of Arachidonic acid (ACD). In fact, a product of this metabolism, i.e. prostaglandin D_2, inhibits the host immune response during the first stages of infections Therefore, we selected two target proteins: the Glutathione S-Transferase from *S. haematobium* (Sh28GST), involved in the isomerization of prostaglandin H_2 to prostaglandin D_2 and the Fatty Acid Binding Protein from *S. mansoni* (Sm14) implicated in the transport of fatty acids (FAs) from the host blood into the parasite interior.

In the first part of my Ph.D thesis, I presented the high resolution structures of Sh28GST at different GSH concentrations and the structure of the Y10F mutant, together with the characterization of their GSH affinities and their transferase activity.

The crystal structure of Sh28GST, a member of the very well studied GST protein family, reveals new features concerning the role of the active site residues, the structure of the dimer interface, and the affinity for GSH. In particular, an active site residue (Tyr10) with previously well defined roles in structure and function was found to have additional, or perhaps previously unrecognized, stereochemical features.

The non-activating conformer of the highly conserved Tyr10 at the active site, which was not previously identified, is especially intriguing. From inspection of the high resolution 3D structure and from crystallization experiments ranging from free, through partially saturated and than fully GSH saturated enzyme, the two conformers appear to change in relative population. Moreover, evidence from solution studies carried out on other GSTs, demonstrated a high conformational flexibility, especially of the

apoprotein, and suggested that in solution interconversion of the two Tyr10 conformers might be possible.

The presence of this new conformer allowed us to hypothesise that Sh28GST is involved not only in the detoxifying/transferase activity but also in prostaglandin D_2 synthase activity. Interestingly, both mechanisms are possibly common to other GSTs, at least those having a Tyr in the active site. The proposed mechanisms, thought still hypothetical, allow us to consider that niche occupied by Tyr10 in the non-activating conformer could be exploited in the design of inhibitors which would fill this cavity, essential to accommodate the phenolic ring of the catalytic Tyr10.

In the second part of my Ph.D thesis, the 3D structure of the two complexes of the Sm14 with ACD and oleic acid (OLA) was solved by means of X-ray crystallography at 2.4 and 1.85Å of resolution, respectively. Moreover, we characterized the binding and release reactions of Sm14 and several important FAs.

The functional and structural data of Sm14 presented in Chapter 4 enable us to infer its physiological role in the parasite life cycle; in the definitive host, Sm14 acts as a shuttle of FAs, particularly polyunsaturated ones, from the host serum to the parasite interior (directly or possibly through a receptor). The proposed mechanism implies that local pH-changes induce a modification of Sm14 ligand affinity through a conformational change, accounting for its function as a shuttle for FAs towards the cell interior. This information may also be exploited in the design of specific drugs, considering that the function of this protein is crucial for the parasite. Finally, our structures provide a rational explanation for the antigenicity of two peptides selected for vaccine trials and could guide the search for other peptides capable of eliciting highly protective antibodies.

CHAPTER 6

6. REFERENCE LIST

- Ahmed, A.A., Babiker, A., Eltash, L.A., Shiff, C. (2002) Development of a modified baited trap for detection of schistosome cercariae using natural oils rich in polyunsaturated fatty acids in Sudan. *Acta Trop.*, **82**, 363-8

- Alvite, G., Di Pietro, S.M., Santomè, J.A., Ehrlich, R. and Esteves, A. (2001) Binding properties of *Echinococcus granulosus* fatty acid binding protein, *Biochem. Biophys. Acta* **1533**, 293-302

- Angeli, V., Faveeuw, C., Roye, O., Fontaine, J., Teissier, E., Capron, A., Wolowczuk, I., Capron, M., and Trottein, F. (2001) Role of the parasite-derived prostaglandin D2 in the inhibition of epidermal Langerhans cell migration during schistosomiasis infections, *J. Exp. Med.* **193**, 1135-1147.

- Angelucci, F., Johnson, K.A., Baiocco, P., Miele, A.E., Brunori, M., Valle, C., Vigorosi, F., Troiani, A.R., Liberti, P., Cioli, D., Klinkert, M.Q., Bellelli, A. (2004) *Schistosoma mansoni* fatty acid binding protein: specificity and functional control as revealed by crystallographic structure. *Biochemistry,* **43**, 13000-11

- Armstrong, R. N. (1997) Structure, catalytic mechanism, and evolution of the glutathione transferases, *Chem. Res. Toxicol*. **10**, 2-8.

- Balendiran, K.G., Schnutgen, F., Scapin, G., Borchers, T., Xhong, N., Lim, K., Godbout, R., Spener, F. and Sacchettini, J. (2000) Crystal structure and thermodynamic analysis of human brain fatty acid binding protein, *J. Biol. Chem*. **275**, 27045-27054

- Balloul, J. M., Grzych, J. M., Pierce, R. J., and Capron, A. (1987) A purified 28,000 dalton protein from Schistosoma mansoni adult worms protects rats and mice against experimental schistosomiasis, *J. Immunol.* **138**, 3448-3453.

- Balloul, J. M., Sondermeyer, P., Dreyer, D., Capron, M., Grzych, J. M., Pierce, R. J., Carvallo, D., Lecocq, J. P., and Capron, A. (1987) Molecular cloning of a protective antigen of schistosomes, *Nature* **326**, 149-153.

- Baumheuer, M.E., Zerfass, M., Ruppel, A., Leichsenring, M. (1994) The fatty acid composition of plasma and erythrocytes in Schistosoma mansoni-infected mice. *Trop. Med. Parisitol*., **45**, 5-8

- Bergquist, N.R., and Colley D.G. (1998) Schistosomiasis Vaccines:

Research to development, *Parasitol. Today* **14**, 99-104

- Bjornestedt, R., Stenberg, G., Widersten, M., Board, P.G., Sinning, I., Jones, T.A., Mannervik, B. (1995) Functional significance of arginine 15 in the active site of human class alpha glutathione transferase A1-1. *J Mol Biol.*, **247**, 765-73.

- Board, P. G., Coggan, M., Chelvanayagam, G., Easteal, S., Jermiin, L. S., Schulte, G. K., Danley, D. E., Hoth, L. R., Griffor, M. C., Kamath, A. V., Rosner, M. H., Chrunyk, B. A., Perregaux, D. E., Gabel, C. A., Geoghegan, K. F., and Pandit, J. (2000) Identification, characterization, and crystal structure of the omega class glutathione transferases, *J. Biol. Chem.* **275**, 24798-24806.

- Booth, J., Boyland, E., Sims, P. (1961). An enzyme from rat liver catalyzing conjugation with glutathione. *Biochem. J.*, **79**, 516-524

- Boulanger, D., Reid, G.D., Sturrock, R.F., Wolowczuk, I., Balloul, J.M., Grezel, D., Pierce, R.J., Otieno, M.F., Guerret, S., Grimaud, J.A., et al. (1991) Immunization of mice and baboons with the recombinant Sm28GST affects both worm viability and fecundity after experimental infection with Schistosoma mansoni. *Parasite Immunol.*, **13**, :473-90

- Boulanger, D., Warter, A., Sellin, B., Lindner, V., Pierce, R. J., Chippaux, J.-P., and Capron, A. (1999) Vaccine potential of a recombinant glutathione S-transferase cloned from *Schistosoma haematobium* in primates experimentally infected with an homologous challenge, *Vaccine* **17**, 319-326.

- Bushara, H.O., Bashir, M.E., Malik, K.H., Mukhtar, M.M., Trottein, F., Capron, A., Taylor, M.G. (1993) Suppression of Schistosoma bovis egg production in cattle by vaccination with either glutathione S-transferase or keyhole limpet haemocyanin. *Parasite Immunol.*,**15**, 383-90.

- Brouwers, J.F., Smeenk, I.M., van Golde, L.M. and Tielens, A.G. (1997) The incorporation, modification and turnover of fatty acids in adult *Schistosoma mansoni*, *Mol. Biochem. Parasitol*, **88**, 175-185

- Brindley, P.J., Strand, M., Norden, A.P., Sher, A. (1989) Role of host antibody in the chemotherapeutic action of praziquantel against Schistosoma mansoni: identification of target antigens. *Mol Biochem Parasitol.*, **34**, 99-108

- Brito, C.F., Oliveira, G.C., Oliveira. S.C., Street, M., Riengrojpitak, S., Wilson, R.A., Simpson, A.J. and Correa-Oliveira, R. (2002) Sm14 gene expression in different stages of the Schistosoma mansoni life cycle and immunolocalization of the Sm14 protein within the adult worm, *Braz. J. Med. Biol. Res.* **35**, 377-81

- Capron, A., Capron, M., Dombrowicz, D., and Riveau, G. (2001) Vaccine strategies against schistosomiasis: from concepts to clinical trials, *Int. Arch. Allergy Immunol.* **124**, 9-15.

- Christ-Hazelhof, E., Nugteren, D. H., and Van Dorp, D. A. (1976) Conversions of prostaglandin endoperoxides by glutathione-S transferases and serum albumins, *Biochim. Biophys. Acta* **450**, 450-461.

- Cioli, D. and Pica-Mattoccia, L. (2003) Praziquantel. *Parisitol. Res.*, **90**, S3-S9

- Collaborative Computational Project, Number 4 (1994) The CCP4 Suite: Programs for Protein Crystallography, *Acta Crystallogr*. **D50**, 760-763.

- Dietze, E. C., Wang, R. W., Lu, A. Y. H., and Atkins, W. M. (1996) Ligand effects on the fluorescence properties of tyrosine 9 in alpha 1-1 glutathione S-transferase, *Biochemistry* **35**, 6745-6753.

- Eads, J., Sacchettini, J.C., Kromminga, A. and Gordon, J.I. (1993) Escheria coli-derived rat intestinal fatty acid binding protein with bound myristate at 1.5Å resolution and I-FABP Arg106→Gln with bound oleate at 1.74Å resolution, *J. Biol. Chem.* **268**, 26375-26385

- Eastman, S.J., Wilschut, J., Cullis, P.R. and Hope, M.J.(1989) Intervesicular exchange of lipids with weak acid and weak base characteristic: influence of transmembrane pH gradients, *Biochim. Biophys. Acta* **981**, 178-184

- Esteves, A., Joseph, L., Paulino, M. and Ehrlich R. (1997) Remarks on the philogeny and structure of fatty acid binding proteins from parasitic platyhelminths, *Int. J. Parasitol.* **27**, 1013-1027

- Fallon, P.G., Tao, L.F., Ismail, M.M., Bennett, J.L. (1996) Schistosome resistance to praziquantel: Fact or artifact? *Parasitol Today.*, **12**, 316-20.

- Foster, J. M., and Johnston, D. A. (2002) Helminth genomics: from gene discovery to genome sequencing, *Trends Parasitol*. **18**, 241-242.

- Furlong S.T. and Caulfield J.P. *Schistosoma mansoni*: synthesis and release of phospholipids, lysophospholipids, and neutral lipids by schistosomula (1989), *Exp Parasitol* **69**, 65-67

- Fusco, A.C., Salafsky, B., Delbrook, K., (1986) Schistosoma mansoni: production of cercarial eicosanoids as correlates of penetration and transformation. *J Parasitol.*, **72**, 397-404

- Grzych, J.-M., Grezel, D., Xu, C. B., Neyrinck, J.-L., Capron, M., Ouma, J. H., Butterworth, A. E., and Capron, A. (1993) IgA antibodies to a protective antigen in human schistosomiasis mansoni, *J. Immunol*. **150**, 527-535.

- Gu, Y., Singh, S. V., and Ji, X. (2000) Residue R216 and catalytic efficiency of a murine class alpha glutathione S-transferase toward benzo[a]pyrene 7(R),8(S)-diol 9(S),10(R)-epoxide, *Biochemistry* **39**, 12552-12557.

- Fan, J., Gan. J., Yang, W., Liying, S., McManus, D. and Brindley, P.J. (2003) A *Schistosoma japonicum* very low-density lipoprotein-binding protein, *Int. J. Biochem. & Cell. Biol*. **35**, 1436-1451

- Ferrari, M.L.A., Coelho, P.M.Z., Antunes, C.M.F., Tavares, C.A.P. and da Cunha, A.S. (2003) Efficacy of oxamniquine and praziquantel in the treatment of *Schistosoma mansoni* infection: a controlled trial, *Bull. World Health Organ.* **83**, 190-196

- Foster, R., (1987) A review of clinical experience with oxamniquine. *Trans R Soc Trop Med Hyg*; **81**, 55-9

- Halawany, A., (1952) The centenary of the discovery of Bilharzia. *J. Egypt. Med. Assoc*., **35**, 89-93.

- Harn, D.A., Gu, W., Oligino. L.D., Mitsuyama, M., Gebremichael, A., Richter. D. (1992) A protective monoclonal antibody specifically recognizes and alters the catalytic activity of schistosome triose-phosphate isomerase. *J Immunol*., **148**, 562-7.

- Harnett, W. and Kusel, J.R. (1986) Increased exposure of parasite antigens at the surface of adult male Schistosoma mansoni exposed to praziquantel in vitro. *Parasitology*; 93, 401-5.

- Herr, F.H., Aronson, J. and Storch, J. (1996) Role of the portal lysine residues in electrostatic interactions between heart fatty acid binding protein and phospholipid membranes, *Biochemistry* **35**, 11840-11845

- Hervè, M., Angeli, V., Pinzar, E., Wintjens, R., Faveeuw, C., Narumiya, S., Capron, A., Urade, Y., Capron, M., Riveau, G., Trottein, F. (2003) Pivotal roles of the parasite PGD_2 synthase and the host D prostanoid receptor 1 in schistosome immune response. *Eur. J. Immunol.*, **33**, 2764-72

- Hitchens, T. K., Mannervik, B., and Rule, G. S. (2001) Disorder-to-order transition of the active site of human class pi glutathione transferase, GST P1-1, *Biochemistry* **40**, 11660-11669.

- Hodson, M.E. and Cistola D.P. (1997) Ligand binding alters the backbone mobility of intestinal fatty acid binding protein by ^{15}N NMR relaxtion and ^{1}H exchange, *Biochemistry* **36**, 2278-2290

- Hubatsch, I., and Mannervik, B. (2001) A highly acidic tyrosine 9 and a normally titrating tyrosine 212 contribute to the catalytic mechanism of human glutathione transferase A4-4, *Biochem. Biophys. Res. Commun*. **280**, 878-882.

- Ibarra, C., Nieslanik, B. S., and Atkins, W. M. (2001) Contribution of aromatic-aromatic interactions to the anomalous pKa of tyrosine-9 and the C-terminal dynamics of glutathione S-transferase, *Biochemistry* **40**, 10614-10624.

- Jakobson, I., Warholm, M., and Mannervik, B. (1979) The binding of substrates and a product of the enzymatic reaction to glutathione S transferase A, *J. Biol. Chem*. **254**, 7085-7089.

- Ji, X., von Rosenvinge, E. C., Johnson, W. W., Tomarev, S. I., Piatigorsky, J., Armstrong, R. N., and Gilliland, G. L. (1995) Three-dimensional structure, catalytic properties, and evolution of a sigma class glutathione transferase from squid, a progenitor of the lens S-crystallins of cephalopods, *Biochemistry* **34**, 5317-5328.

- Ji, X., Von Rosenvinge, E. C., Johnson, W. W., Armstrong, R. N., and Gilliland, G. L. (1996) Location of a potential transport binding site in a sigma class glutathione transferase by X-ray crystallography, *Proc. Natl. Acad. Sci*. U.S.A. **93**, 8208-8213.

- Johnson, K.A., Angelucci, F., Bellelli, A., Herve, M., Fontane, J., Tsernoglou, D., Capron, A., Trottein, F., Brunori, M. (2003) Crystal structure of the 28 kDa glutathione S-transferase from *Schistosoma haematobium*. *Biochemistry*, **42**, 10084-94.

- Jordan, P., (1972) Epidemiology and control of schistosomiasis *Br. Med. Bull.,* **28**, 55-9.

- Jowsey, I. R., Thomson, A. M., Flanagan, J. U., Murdock, P. R., Moore, G. B. T., Meyers, D. J., Murphy, G. J., Smith, S. A., and Hayes, H. D. (2001) Mammalian class sigma glutathione S-transferases: catalytic properties and tissue-specific expression of human and rat GSH-dependent prostaglandin D2 synthases, *Biochem. J.* **349**, 507-516.

- Kaikaus, R.M., Bass, N.M. and Okner R.K. (1990) Functions of fatty acid binding protein, *Experientia* **46**, 617-630

- Kanaoka, T., Ago, H., Inagaki, E., Nanayama, T., Miyano, M., Kikuno, R., Fujii, Y., Eguchi, N., Toh, H., Urade, Y., and Hayaishi, O. (1997) Cloning and crystal structure of hematopoietic prostaglandin D-synthase, *Cell* **90**, 1085-1095.

- Kennedy, M.W., Scott, J.C, Lo, S., Beauchamp, J. and McManus, D.P. (2000) Sj-FABPc fatty-acid binding protein of the human blood fluke Schistosoma japonicum: structural and functional characterization and unusual solvent exposure of a portal-proximal tryptophan residue, *Biochem. J*. **349**, 377-384

- Kirk, R.W., Kurian, E. and Prendergast, F.G. (1996) Characterization of the sources of protein-ligand affinity: 1-sulfonato-8-anilinonaphtalene binding to intestinal fatty acid binding protein, *Biophys. J.* **70**, 69-83

- Koehler, R. T., Villar, H. O., Bauer, K. E., and Higgins, D. L. (1997) Ligand-based protein alignment and isozyme specificity of glutathione S-transferase inhibitors, *Proteins* **28**, 202-216.

- Kurian, E., Kirk, W.R. and Prendergast, F.G. (1996) Affinity for rRat intestinal fatty acid binding protein: Further examination, *Biochemistry* **35**, 3865-3874

- Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature* **227**, 680-685

- LaLonde, J.M., Levenson, M.A., Roe, J.J., Bernlohr, D.A. and Banaszak, L.J. (1994) Adipocyte lipid-binding protein complexed with arachidonic acid, *J. Biol. Chem.* **269**, 25339-25347

- Laskowski, R. A., McArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr*. **26**, 283-291.

- Liang, J., Edelsbrunner, H.,Woodward, C. (1998) Anatomy of Protein Pockets and Cavities: Measurement of Binding Site Geometry and Implications for Ligand Design, *Protein Sci.*, **7**, 1884-1897

- Lien, L., Gustafsson, A., Andersson, A.-K., and Mannervik, B. (2001) Human glutathione transferase A1-1 demonstrates both half-of-the-sites and all-of-the-sites reactivity, *J. Biol. Chem.* **276**, 35599-35605.

- Lim, K., Ho, J. X., Keeling, K., Gilliland, G. L., Ji, X., Ruker, F., and Carter, D. C. (1994) Three-dimensional structure of Schistosoma japonicum glutathione S-transferase fused with a six-amino acid conserved neutralizing epitope of gp41 from HIV, *Protein Sci*. **3**, 2233-2244.

- Liu, S., Ji, X., Gilliland, G. L., Stevens, W. J., and Armstrong, R. N. (1993) Second-sphere electrostatic effects in the active site of glutathione S-transferase. Observation of an on-face hydrogen bond between the side chain of threonine 13 and the -cloud of the tyrosine 6 and its influence on catalysis, *J. Am. Chem. Soc*. **115**, 7910-7911.

- Lo Bello, M., Nuccetelli, M., Chiessi, E., Lahm, A., Mazzetti, A. P., Battistoni, A., Caccuri, A. M., Oakley, A. J., Parker, M. W., Tramontano, A., Federici, G., and Ricci, G. (1998) Mutations of Gly to Ala in human

glutathione transferase P1-1 affect helix 2 (G-site) and induce positive cooperativity in the binding of glutathione, *J. Mol. Biol*. **284**, 1717-1725.

- Mannervick, B., and Guthenberg, C. (1981) Glutathione transferase (human placenta), *Methods Enzymol.* **77**, 231-235.

- Matsumoto, Y., Perry, G., Levine, R.J., Blanton, R., Mahmoud, A.A., Aikawa, M. (1988) Paramyosin and actin in schistosomal teguments. *Nature*;**333**, 76-8

- Matthews, B. W. (1968) Solvent content of protein crystals, *J. Mol. Biol.* **33**, 491-497

- McTigue, M. A., DeWight, R. W., and Tainer, J. A. (1995) Crystal structure of a schistosomal drug and vaccine target: glutathione S-transferase from Schistosoma japonicum and its complex with the leading anti-schistosomal drug praziquantel, *J. Mol. Biol.* **246**, 21-27.

- McRee, D. E. (1999) XtalView/Xfit: A Versatile Program for Manipulating Atomic Coordinates and Electron Density, *J. Struct. Biol.* **125**, 156-165.

- McKerrow, J. and Salter, J. (2002) Invasion of skin by schistosoma cercariae, *Trends Parasitol*. **18**, 193-195

- McNair, A.T., Dissous, C., Duvaux-Miret, O., Capron, A. (1993) Cloning and characterisation of the gene encoding the 28-kDa glutathione S-transferase of Schistosoma mansoni. *Gene*, **124**, 245-9

- Milhon, J. L., Thiboldeaux, R. L., Glowac, K., and Tracy, J. W. (1997) Schistosoma japonicum GSH S-transferase Sj26 is not the molecular target of praziquantel action, *Exp. Parasitol.* **87**, 268-274.

- Moser, D., Tendler, M.,Griffiths, G. and Klinkert M-Q. (1991) A 14-kDa *Schistosoma mansoni* polypeptide is homologous to a gene family of fatty acid binding proteins, *J. Biol. Chem* **266**, 8447-8454

- Mountford AP, Trottein F. (2004) Schistosomes in the skin: a balance between immune priming and regulation. *Trends Parasitol.*, **20**, 221-6

- Murphy, E.J., Barcelo-Coblijn, G., Binas, B. and Glatz, J.F. (2004) Heart fatty acid uptake is decreased in heart-fatty acid binding protein gene-ablated

mice, *J. Biol. Chem.* Jun 11

- Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method, *Acta Crystallogr*. **D53**, 240-255.

- Navaza, J. (1994) AMoRe: an Automated Package for Molecular Replacement, *Acta Crystallogr*. **50**, 157-163.

- Oakley, A. J., Bello, M. L., Battistoni, A., Ricci, G., Rossjohn, J., Villar, H. O., and Parker, M. W. (1997) The structures of human glutathione transferase P1-1 in complex with glutathione and various inhibitors at high resolution, *J. Mol. Biol.* **274**, 84-100.

- Orozco, M., Vega, C., Parraga, A., Garcia-Saez, I., Coll, M., Walsh, S., Mantle, T. J., and Luque, F. J. (1997) On the reaction mechanism of class pi glutathione S-transferase, Proteins: *Struct., Funct., Genet.* **28**, 530-542.

- Ory, J.J. and Banszak, L. (1999) Studies of ligand binding reaction of adipocyte lipid binding protein using the fluorescent probe 1,8-anilinonaphtalene-8-sulfonate, *Biophys. J.* **77**, 1107-1116

- Otwinowski, Z., and Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode, *Methods Enzymol*. **276**, 307-326.

- Polekhina, G., Board, P. G., Blackburn, A. C., and Parker, M. W. (2001) Crystal structure of maleylacetoacetate isomerase/glutathione transferase zeta reveals the molecular basis for its remarkable catalytic promiscuity, *Biochemistry* **40**, 1567-1576.

- Pinzar, E., Miyano, M., Kanaoka, Y., Urade, Y., and Hayaishi, O. (2000) Structural basis of hematopoietic prostaglandin D synthase activity elucidated by site-directed mutagenesis, *J. Biol. Chem.* **275**, 31239-31244.

- Ramaswamy, K., Kumar, P., He, Y.X. (2000) A role for parasite-induced PGE2 in IL-10-mediated host immunoregulation by skin stage schistosomula of Schistosoma mansoni. *J Immunol.*, **165**, 4567-74

- Ramos, C.R., Figueredo, R.C., Pertinhez, T.A., Vilar, M.M., do Nascimento, A.L., Tendler, M., Raw, I., Spisni, A. and Ho, P.L. (2003) Gene structure and M20T polymorphism of the *Schistosoma mansoni* Sm14 fatty

acid binding protein, *J. Biol. Chem.* **278**, 12745-12751

- Richieri, G.V., Low, P.J., Ogata, R.T. and Kleinfeld, A.M. (1998) Thermodynamics of fatty acid binding protein to engineered mutants of the adipocyte and intestinal fatty acid binding proteins, *J. Biol. Chem.* **273**, 7397-7405

- Richieri, G.V., Ogata, R.T., and Kleinfeld A.M. (1992) A fluorescently labeled intestinal fatty acid binding protein, *J. Biol. Chem.* **267**, 23495-23501

- Remoué, F., Mani, J.-C., Pugnière, M., Schact, A.-M., Capron, A., and Riveau, G. (2002) Functional specific binding of testosterone to Schistosoma haematobium 28-kilodalton glutathione S-transferase, *Infect. Immun*. **70**, 601-605.

- Richieri, G.V., Ogata, R.T., Zimmerman, A.W., Veerkamp, J.H., and Kleinfeld, A.M. (2000) Fatty acid binding proteins from different tissues show distinct pattern of fatty acid interactions *Biochemistry* **39**, 7197-7204

- Ricci, G., Caccuri, A. M., Bello, M. L., Rosato, N., Mei, G., Nicotra, M., Chiessi, E., Mazzetti, A. P., and Federici, G. (1996) Structural flexibility modulates the activity of human glutathione transferase P1-1. Role of helix flexibility in the catalytic mechanism, *J. Biol. Chem*. **271**, 16187-16192.

- Ricci, G., Del Boccio, G., Pennelli, A., Lo Bello, M., Petruzzelli, R., Caccuri, A. M., Barra, D., and Federici, G. (1991) Redox forms of human placenta glutathione transferase, *J. Biol. Chem.* **266**, 21409-21415.

- Richieri, G.V. and Kleinfeld, A.M. (1995) Unbound free fatty acid levels in human serum, *J. Lipid Res*. **36**, 229-240

- Riveau, G., Poulain-Godefroy, A. P., Dupre, L., Remoue, F., Mielcarek, N., Locht, C., and Capron, A. (1998) Glutathine S-transferases of 28 kDa as major vaccine candidates against schistosomiasis, *Mem. Inst. Oswaldo Cruz* **93** (Suppl. 1), 87-94.

- Rumjanek, F.D. and Simpson A.J. The incorporation and utilization of radiolabelled lipids by adult *Schistosoma mansoni* (1980) *Mol. Biochem. Parasitol.* **1**, 31- 44

- Sacchettini, J.C., Gordon, J.I. and Banaszak, L.J. (1989) Crystal sructure of

rat intestinal fatty acid binding protein. Refinement and analysis of the *Escheria coli*-derived protein with bound palmitate, *J. Mol. Biol*. **208**, 327-339

- Samuelson, J.C., Caulfield, J.P., David, J.R. (1980) Schistosoma mansoni: post-trasformational surface changes in schistosomula grown in vitro and in mice. *Exp. Parisitol.*, **50**, 369-83

- Salafsky, B., Wang, Y.S., Kevin, M.B., Hill, H., Fusco, A.C. (1984) The role of prostaglandins in cercarial (*Schistosoma mansoni*) response to free fatty acids. *J Parasitol.*, **70**, 584-91

- Scott, J. C., and McManus, D. P. (2000) Molecular cloning and enzymatic expression of the 28-kDa glutathione S-transferase of Schistosoma japonicum: evidence for sequence variation but lack of consistent vaccine efficacy in the murine host, *Parasitol. Int*. **49**, 289-300.

- Sheehan, D., Meade, G., Foley, V. M., and Dowd, C. A. (2001) Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily, *Biochem. J.* **360**, 1-16.

- Simpson, A.J., (1992) Tegumental proteins of Schistosoma mansoni: complex biomolecules and potent antigens. *Mem Inst Oswaldo Cruz.*;**87** Suppl 4:11-7.

- Stella, L., Caccuri, A. M., Rosato, N., Nicotra, M., Lo Bello, M., De Matteis, F., Mazzetti, A. P., Federici, G., and Ricci, G. (1998) Flexibility of helix 2 in the human glutathione transferase P1-1. Time-resolved fluorescence spectroscopy, *J. Biol. Chem.* **273**, 23267-23273.

- Storch, J. and Thumser, A.E. (2000) The fatty acid transport function of fatty acid binding proteins, *Biochem. Biophys. Acta* **1486**, 28-44

- Taylor, J. B., Vidal, A., Torpier, G., Meyer, D. J., Roitsch, C., Balloul, J. M., Southan, C., Sondermeyer, P., Pemble, S., and Lecocq, J. P., (1988) The glutathione transferase activity and tissue distribution of a cloned Mr28K protective antigen of Schistosoma mansoni, *EMBO J.* **7**, 465-472.

-Taylor, M.G. (1994) Schistosomiasis vaccines: Farewell to the God of Plague? *J. Trop. Med. Hyg.,* **97**, 257-68

- Thomson, A. M., Meyer, D. J., and Hayes, J. D. (1998) Sequence, catalytic properties and expression of chicken glutathione-dependent prostaglandin D2 synthase, a novel class sigma glutathione S-transferase, *Biochem. J.* **333**, 317-325.

- Tendler, M., Brito, C.A., Vilar, M.M., Serra-Freire, N., Diogo, C.M., Almeida. M.S., Delbem, A.C., Da Silva, J.F., Savino, W., Garratt, R.C., Katz, N., Simpson, A.S. (1996) A Schistosoma mansoni fatty acid-binding protein, Sm14, is the potential basis of a dual-purpose anti-helminth vaccine, *Proc. Natl. Acad. Sci. U S A* **93**, 269-73

- Thom, R., Dixon, D. P., Edwards, R., Cole, D. J., and Laphorn, A. J. (2001) The structure of a zeta class glutathione S-transferase from Arabidopsis thaliana: characterization of a GST with novel active-site architechture and a putative role in tyrosine catabolism, *J. Mol. Biol*. **308**, 949-962.

- Trottein, F., Vaney, M.-C., Bachet, B., Pierce, R.-J., Colloc'h, N., Lecocq, J.-P., Capron, A., and Mornon, J.-P. (1992) Crystallization and preliminary X-ray diffraction studies of a protective cloned 28 kDa glutathione S-transferase from Schistosoma mansoni, *J. Mol. Biol.* **224**, 515-518.

- Trottein, F., Godin, C., Pierce, R. J., Sellin, B., Taylor, M. G., Gorillot, I., Silva, M. S., Lecocq, J. P., and Capron, A. (1992) Inter-species variation of schistosome 28-kDa glutathione S-transferases, *Mol. Biochem. Parasitol*. **54**, 63-72.

- Todd, A. E., Orengo, C. A., and Thornton, J. M. (2002) Plasticity of enzyme active sites, *Trends Biochem. Sci*. **27**, 419-425.

- Urade, Y., Fujimoto, N., Ujihara, M., and Hayaishi, O. (1987) Biochemical and immunological characterization of rat spleen prostaglandin D synthetase, *J. Biol. Chem.* **262**, 3820-3825.

- Vilar, M.M., Barrientos, F., Almeida, M., Thaumaturgo, N., Simpson, A., Garratt, R. and Tendler, M. (2003) An experimental bivalent peptide vaccine against schistosomiasis and fascioliasis, *Vaccine* **22**, 137-144

- Veerkamp, J.H, Peeters, R.A. and Maatman, R.G. (1991) Structural and functional features of different types of cytoplasmatic fatty acid-binding proteins, *Biochem. Biophys. Acta* **1081**, 1-24

- Xiao, B., Singh, S. P., Nanduri, B., Awasthi, Y. C., Zimniak, P., and Ji, X. (1999) Crystal structure of a murine glutathione S-transferase in complex with a glutathione conjugate of 4-hydroxynon-2-enal in one subunit and glutathione in the other: evidence of signaling across the dimer interface, *Biochemistry* **38**, 11887-11894.

- Xu, C. B., Verwaerde, C., Grzych, J. M., Fontaine, J., and Capron, A. (1991) A monoclonal antibody blocking the Schistosoma mansoni 28-kDa glutathione S-transferase activity reduces female worm fecundity and egg viability, *Eur. J. Immunol.* **21**, 1801-1807.

- Weisiger, R.A. (2002) Cytosolic fatty acid binding proteins catalyze two distinct steps in intracelleular transport of their ligands, *Mol. Cell. Biochem.* **239**, 35-42

- Widersten, M., Kolm, R. H., Bjornsted, R., and Mannervick, B. (1992) Contribution of five amino acid residues in the glutathione binding site to the function of the human glutathione S-transferase P1-1, *Biochem. J.* **285**, 377-381.

- Wilce, M. C., and Parker, M. W. (1994) Structure and function of glutathione S-transferases, *Biochim. Biophys. Acta* **1205** (1), 1-18.

- World Health Organization (1993) Expert Comitee on the Control of Schistosomiasis. The control of schistosomiasis, WHO Technical Report 830, Geneva.

- World Health Organization (1996) Fact sheet on schistosomiasis, World Health Organization, Geneva.

- World Health Organization (2002) Prevention and Controlof Schistosomiasis and soil-transmitted helminthiasis, WHO technical report series 912, Geneve

- Wootan, M.G., Bernlohr.D.A. and Storch, J. (1993) Mechanism of fluorescent fatty acid transfer from adypocite fatty acid binding protein to membranes, *Biochemistry* **32**, 8622-8627

- Zanotti, G., Scapin, G., Spadon, P., Veerkamp, J.H. and Sacchettini, J.C. (1992) Three-dimensional structure of recombinant human muscle fatty acid-binding protein, *J. Biol. Chem.* **267**, 18541

CHAPTER 7

7. PUBLICATIONS

The experimental work I carried on during the three years of Ph.D., mainly deals with the study of the structure and function relationship of proteins directly implicated in the pathogenesis of schistosomiasis, the second most prevalent parasitic disease after malaria. This work arise from a close collaboration with the group of Prof. Donato Cioli at the CNR Institute of Cell Biology of Monterotondo, with the group of Prof. François Trottein at the Institute Pasteur of Lille (France) and resulted in two articles, which may be found in the next pages.

Johnson, K.A., Angelucci, F., Bellelli, A., Herve, M., Fontane, J., Tsernoglou, D., Capron, A., Trottein, F., Brunori, M. (2003) Crystal structure of the 28 kDa glutathione S-transferase from Schistosoma haematobium. *Biochemistry*, **42**, 10084-94.

Angelucci, F., Johnson, K.A., Baiocco, P., Miele, A.E., Brunori, M., Valle, C., Vigorosi, F., Troiani, A.R., Liberti, P., Cioli, D., Klinkert, M.Q., Bellelli, A. (2004) Schistosoma mansoni fatty acid binding protein: specificity and functional control as revealed by crystallographic structure. *Biochemistry,* **43**, 13000-11