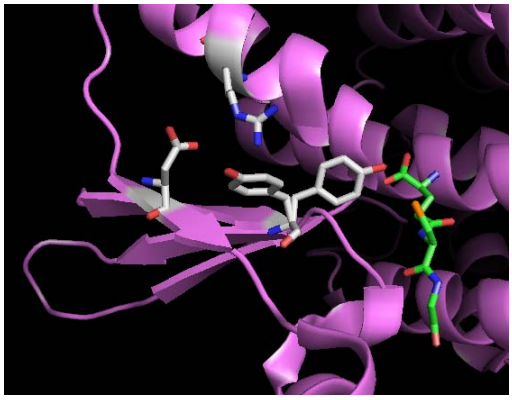
Università degli Studi di Roma "La Sapienza"

Dottorato di Ricerca in Biochimica
XVIII Ciclo (A.A. 2002-2005)

STRUCTURE-FUNCTION ANALYSIS OF PATHOGENICALLY IMPORTANT PROTEINS FROM SCHISTOSOME PARASITES

Paola Baiocco



Docente guida
Prof. Demetrius Tsernoglou

Coordinatore
Prof. Paolo Sarti

Dicembre 2005

ACKNOWLEDGEMENTS

I am grateful to Prof. Maurizio Brunori for giving me the opportunity to learn biochemistry and to progress as a researcher.

Special thanks to Prof. Demetrius Tsenorglou, who helped me in the crystallographic techniques and to Prof. Andrea Bellelli who supervised me during my Ph.D.

I am sincerely grateful to Dr. Adriana E. Miele who introduced me to molecular biology and to crystallography and has always supported me.

A very special thank to my laboratory colleagues: Dr. Francesco Angelucci and Dr. Louise J. Gourlay for helpful scientific (and not) discussions.

I want to thank all the crystallography group: Adele Di Matteo, Andrea Ilari, Beatrice Vallone, Daniele Bonivento, Fabiana Renzi, Giorgio Giardina, Linda Savino and Luca Federici for sharing the hard work at the synchrotrons.

Thanks to Prof. Anna Tramontano, Prof. Paolo Sarti, Alessandro Arcovito, Alessandro Borgia, Fabio Centola, Gianna Panetta, Nicoletta Calosci, Stefano Gianni and Serena Rinaldo for helping me in any occasion.

Experimental work was carried out in collaboration with the group of Prof. D. Cioli and Dr. P. Liberti, at the Institute of Cell Biology of the CNR (Monterotondo, Rome) and with Prof. F. Trottein's group, at the Institute Pasteur of Lille (France). This project was gratefully funded by the Electron Synchrotron Elettra, Trieste.

CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Schistosomiasis

In 1851 Theodor Bilharz, a young German pathologist, discovered the organisms responsible for a disease called Schistosomiasis, and later to be known also as Bilharziasis from his name (Halawany, 1952).

Schistosomiasis, recognized since the time of Egyptians, is a parasitic disease that leads to chronic illness. At present, it is distributed in 75 countries of Africa, southeast Asia, and northwest South America. Approximately 200 million people are infected with schistosomes, resulting in 1000000 deaths each year. 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 damage (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*.

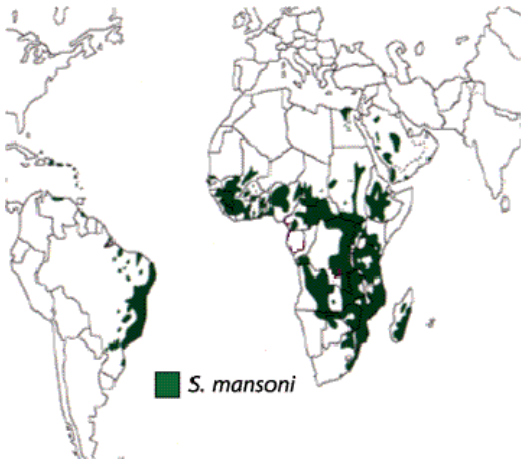


Figure 1.1 *Geographical distribution of S. mansoni.*

S. mansoni causes intestinal schistosomiasis and is prevalent in 52 countries and territories of Africa, Caribbean, the Eastern Mediterranean and South America (Fig. 1.1).*S. japonicum* causes intestinal schistosomiasis and is present in 7 African countries and in the Pacific region

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

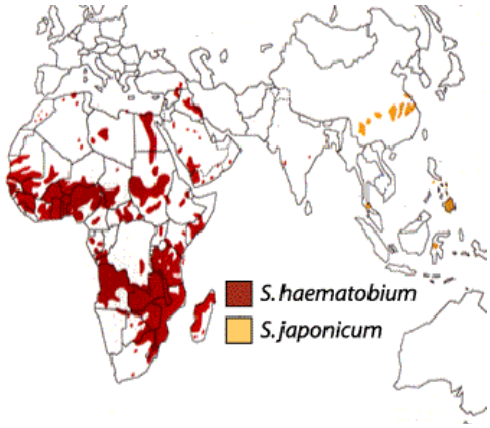


Figure 1.2. *A geographical distribution of S. mansoni and S. haematobium*

The transmission of the parasite is strictly linked to water supply and people become infected because they do not have access to safe water, and maintain transmission because of the absence of proper excreta disposal systems. Infection is acquired during the course of routine domestic, agricultural or occupational duties.

Transmission occurs when schistosome eggs from human excreta hatch into miracidia upon contact with water. The miracidia swim around until they infect into the appropriate species of water snail of the genus Biomphalaria, which is the intermediate host of the schistosomes. The stages in the snail include two generations of sporocysts and the production of cercariae. The

growth in the snail leads to exponential multiplication of the parasite. From a single miracidium a few thousand cercaria, each one capable of infecting men, are produced. 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 (Fig. 1.3). Adult worms in humans reside in the mesenteric veins in various locations, which at times seem to be specific for each species. For instance, *S. japonicum* is more frequently found in the mesenteric veins draining the small intestine (A), and *S. mansoni* occurs more often in the superior mesenteric veins draining the large intestine (B). However, both species can occupy either location, and they are capable of moving between sites, so it is not possible to state unequivocally that one species only occurs in one location. *S. haematobium* most often occurs in the venous plexus of bladder (C), but it can also be found in the rectal venules. The females deposit eggs in the small veins of the portal and periversical systems. The eggs are moved progressively toward the lumen of the intestine (*S. mansoni* and *S. japonicum*) and of the bladder and ureters (*S. haematobium*), and are eliminated with faeces or urine, respectively (1).

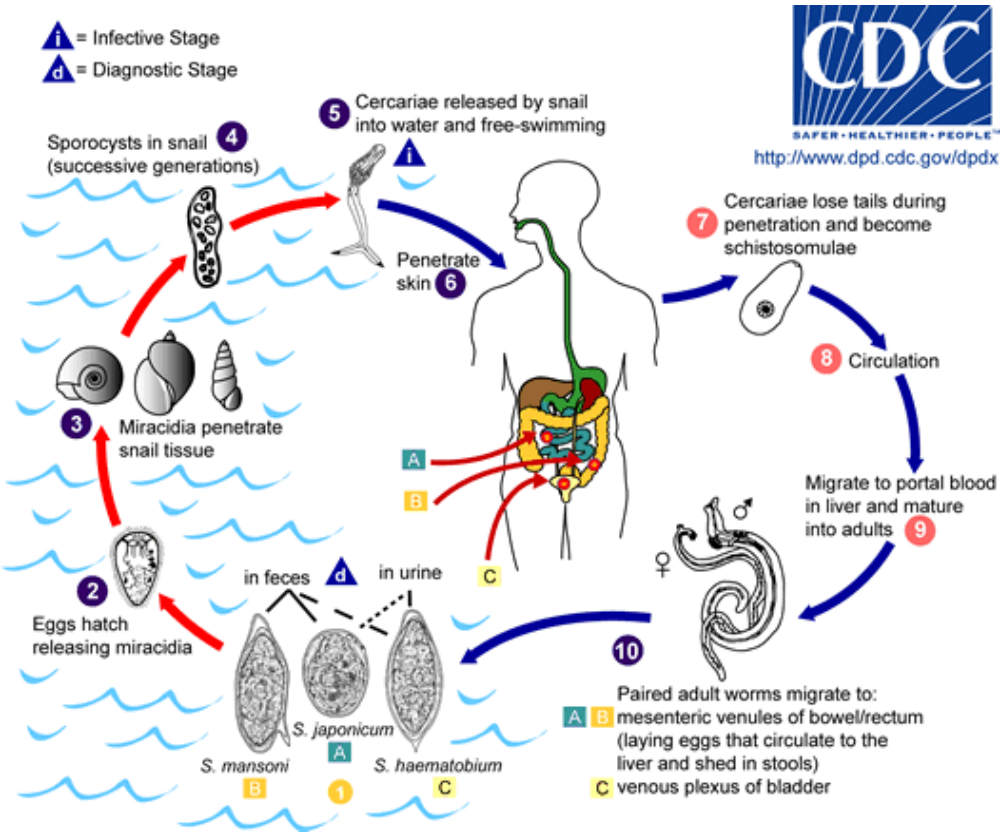


Figure 1.3. ***The life cycle of schistosome worms.*** *Eggs are eliminated with faeces or urine (1). Under optimal conditions the eggs hatch and release miracidia (2), which swim and penetrate specific snails (3). In the snail the miracidia transform into sporocysts (4) and the infective cercariae (5) are released in the water. The cercariae swim and penetrate the skin of the human host (6), and shed their forked tail, becoming schistosomulae (7). The schistosomulae migrate through several tissues to their residence in the veins (8, 9). 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, or more, according to the species. Adult worms (10) in humans reside in the mesenteric veins in various locations, which at times seem to be specific for each species.*

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 and it is estimated that 200 million people are already infected. Extreme poverty, unawareness of the risks, inadequacy or total lack of public health facilities, especially with reference to water supply and disposal, 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 growth delay (Fig. 1.4).

Figure 1.4 *25 year-old man with advanced schistosomiasis*

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 the disease. The increase in "off-track" tourism has led to increasingly serious infections with previously unknown effects, 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 the most 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 on 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 agriculture.

1.2.2 Life cycle and infection behaviour

The eggs pass through the walls of 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 by peristaltic action of the gut, help the egg moving into and passing through the tissues. Biochemical factors, such

as proteolytic enzymes released by the eggs, may help their way through the host tissue. The host inflammatory response also seems to be essential for successful migration of the egg to the lumen of the intestine. This response is caused by the lodged eggs, leading to granuloma formation. 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 them 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. Three main factors affect the hatching of schistosome eggs: temperature (25-30°C), light and osmotic pressure. On entering a hypotonic environment, such that of fresh water, the increase in the osmotic pressure and the activation of the enzyme leucine amino peptidase, inhibited by NaCl, result in rupture of the egg's shell. 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 produce, in the definitive host, adults worms which are either all male or all female. The miracidium seeks and penetrates a snail as 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 substances released by the snail, including

mucus, long chain fatty acids and even amino acids, attract the miracidia. When miracidia are in contact with a snail, a period of exploratory behaviour occurs before the 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 the penetration is complete. The location of 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* prefers organs such as 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. Approximately 4 weeks pass from the penetration of the snail until the release of the cercariae which are released every 24 hours only during daylight. 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 an animal 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 humans and rodents acted as hosts, shedding was found to be

intermediate between the previous two. Schistosomes' cercariae are characterised by long tails, bifurcated at the end (the furcal rami), and a tegument all around the body, made by a trilaminate plasma membrane, and with a glycocalyx on the outer surface. The cercariae of *S. mansoni* are released 25 to 30 days after the snail has been infected and are sexually differentiated. They are a non-feeding stage of the lifecycle, and their energy requirements are met by stored glycogen in both the tail and the body. As they are non-feeding, their energy stores will become depleted, resulting in a reduced infectivity since they become incapable of swimming towards their host, due usually to depletion of the tail reserves as these are used first. Cercariae can remain infective *in vivo*, under optimal conditions, for about 5-8 hours after shedding, although *in vivo* this is probably much less, due to factors such as variation in water temperature. Similarly to the miracidium, the cercariae exhibit a number of behavioural features that enable them to locate their definitive host, such as the bursts of upwards swimming to reach the surface of the water, followed by periods of passive sinking and also sensitivity to other stimuli, such as shadows on the water, turbulence and chemicals secreted by the hosts skin. Once the cercaria has reached its host, it attaches to and penetrates the human skin, using secretions from glands in the head region, and when inside it digests its tail to become a schistosomulum larva. This latter starts migrating through the body and developing at the same time, until it reaches maturity as an adult worm in the liver.

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 esterases on skin

triglycerides). The 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 occurring during the transformation from a free-living infective stage to a parasitic larval schistosomulum.

This metamorphosis takes place in less than one hour and involves: (i) changes in the outer membrane: the cercarial tegument surrounded by a trilaminate plasma membrane with a thick glycocalyx, is lost together with the tail at the time of penetration; (ii) the emptying of the various penetration glands; (iii) release of vesicles from the tegumental cell bodies beneath the muscular tissue, into the tegumental cytoplasm to form a multi-laminate tegumental surface in replacement of the trilaminate membrane; (iv) release of contents from oesophageal glands into the lumen starting the 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 sensitivity to osmosis.

Both physical and biochemical changes take place during the metamorphosis. 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 molecules with low immuno-reactivity. By 24-48 hours 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 has also 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 defense systems. The parasites enter the initial epidermal layer of the skin very rapidly (less than 30 min) and, once reached the dermis, they rest for about 40h, making this compartment a temporary barrier to further penetration. Then they require 10 more hours to locate a venule and a further 8h to penetrate the venule wall. Once in the blood capillaries the schistosomula are carried to the lungs, where they become lodged and double in size over the next few days. This so called “lung phase” lasts for 3-8 days. Following this period the larvae make 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 (Fig. 1.5).

Figure 1.5. *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. Within such an intimate association, the female receives products, such as glucose, across the membrane of 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.

Figure 1.6: *Scanning electron micrograph of a pair of Schistosoma mansoni parasites.*

In the mouse model system only about 20% of the initial cercarial inoculation 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. These lodged eggs trigger 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 wherever they are present, in the intestine or, in the case of eggs swept back by the blood, in the liver and occasionally in other tissues including brain and lungs. In case of a single sex

infection there's no egg production, therefore this infection cause only little pathology.

The clinical signs of the disease pass through two different phases: the acute and the chronic. The acute stage of the infection is often asymptomatic, but when symptoms do occur they include fever, nausea, headache, an irritating cough and, in extreme cases diarrhoea accompanied with blood, mucus and necrotic material. This phase is 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 a response to the sudden high level of antigen exposure and is usually associated with the onset of eggs deposition. Frequently, heavy infections can lead to fibrotic chronic schistosomiasis or to the death of the patient.

The chronic phase of infection is the more important aspect of *S. mansoni* pathology. The most common symptoms are diarrhoea and fever but the infection can also depress children's growth rate and lead to enlargement of the liver and spleen. Fibrosis of the liver can result in portal hypertension, ascites formation, and oesophageal varices leading to fatal hematemesis. In *S. haematobium*, the most frequently affected organ is the urinary bladder, where calcification of eggs trapped in the tissues often occurs. The disease is characterised by blood in the urine (haematuria), hence the infection is often referred to as 'Urinary Schistosomiasis'. Cancer of the bladder is an important complication of infection with *S. haematobium*.

The granuloma formation occurs over about a 16 day period after the deposition, predominantly in the periportal area. Each egg becomes surrounded by a dense infiltrate composed mainly by a number of different

type of immune cells, including both T and B lymphocytes, macrophages, giant cells, epitheloid cells, mast cells, plasma cells, fibroblasts and eosinophils (Fig. 1.7).

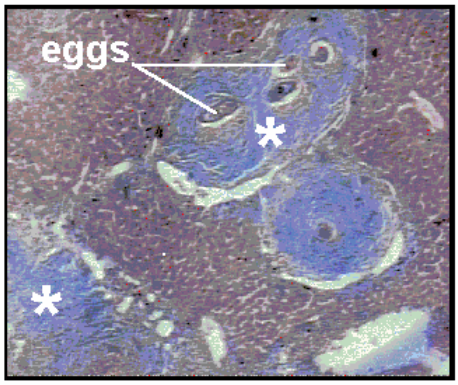


Figure 1.7 *S. mansoni eggs in the liver of an experimentally infected mouse. Two of the several eggs in this section are labeled, and the granulomas (*) that have formed around the eggs are typical of the damage that occurs in the livers of infected hosts.*

These cells all play a significant role in the formation of the granuloma as they interact with each other in a highly complex fashion, involving both cell-cell interactions and indirect interaction through a wide variety of cytokine chemical messengers. It is important to note that these cellular responses are dynamic processes, and different cellular responses are important at different stages of granuloma formation. Because of their size, eggs cannot pass through the capillary beds as the blood flows through the liver. 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. In any case, the main factor is how much immunoresponsive is the host to the schistosome egg antigen (SEA). Although some of the tissue responses are due to the physical presence of the

eggs and to their damage, most of the outcome is due to the host response to the 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 principal 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 cases of no immunodepression and complete immunosuppression result in high host mortality.

1.3 Control of Infection

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 many developing countries. Control of Schistosomiasis aims to prevent new infections, usually by interruption of the parasites life cycle. This may be achieved by a number of methods including: 1) action to eliminate the intermediate host; 2)

elimination of the parasite from the definitive host; 3) prevention of infection of the definitive host; 4) prevention of infection of the intermediate host; 5) the eventual use of a vaccine. Unfortunately, it is doubtful whether an effective vaccine will be ready in the next future for use within national control programmes, the main reason for pessimism being that schistosomiasis does not confer immunity and multiple infestations 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 egg production, most of which fails to be excreted from the host. Therefore the main clinical forms of the 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 clinical signs will not appear until a sufficient number of egg-associated granulomas have been generated. This process is decelerated by the diverse degrees of immunity developed by endemic populations, but, because a balance must be struck to reduce the risk of over-stimulating 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.

There is as yet no effective vaccine for human use. The identification of schistosome antigens is a major priority of the research. To be effective against schistosomiasis, the vaccine must have the following features: i) It must be cheap, because the countries affected by schistosomiasis include some of the poorest in the world; ii) It must be effective in terms of protection, because efficient protection of a large percentage of vaccinated people is essential to interrupt transmission of the parasite; iii) It must be stable, this is linked to i): in fact in order to be used in many of the affected countries, it must be able to survive storage as it is transported to isolated populations where the disease is endemic; iv) It should ideally be effective with a single application, which is of outmost importance to keep the costs of the mass vaccination schemes needed to a minimum.

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. The Special Programme for Research and Training in Tropical Diseases of the World Health Organization (WHO/TDR) has developed a programme to assist selection of promising antigens for further development (Table 1.1).

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.

Table 1.1 *Major Schistosome antigens selected by WHO.*

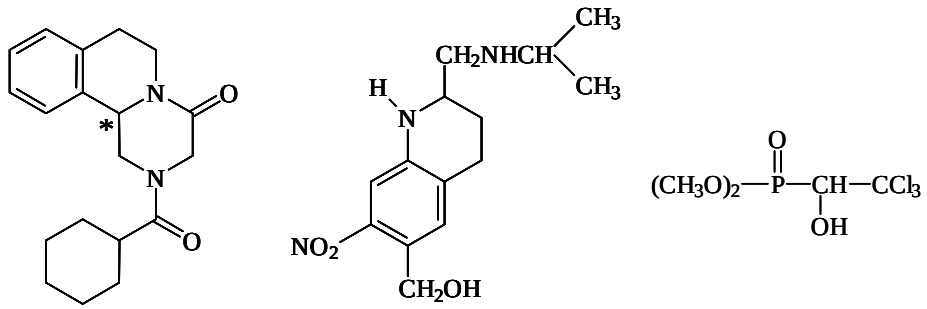
With the great majority of published work carried out with *S. mansoni*, it is likely that a vaccine against this species will be developed first. The selection of vaccine candidates favoured antigens showing high rates of protection, together with additional advantages, such as the role of glutathione S-transferase (GST) in 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.

The antigens listed in Table 1.1 constitute a perplexing array of proteins and it is unclear how proteins that are predicted to be cytosolic, such as Sm14 (Moser *et al.,* 1991), GST (Balloul *et al.,* 1987), triose-phosphate isomerase (TPI) (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 possibility 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 acting as antigenic disguise (Simpson, 1992). If this were the case, however, the antigens would be exposed by damaged or dying parasites and would perhaps offer scarce protection against live and virulent schistosomes. 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 the areas of the 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 investigations on possible mechanisms of protein adsorption to the tegumental surface and the interaction of protein species within the tegumental membrane.

An alternative to vaccination is chemotherapy that can reduce morbidity and lethality but must be reiterated periodically, mainly because of multiple infection. However, morbidity could be possibly suspended permanently, even with continued transmission, if chemotherapy were followed by vaccination. Effective drugs, belonging to trivalent antimonials were first 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* as reviewed by Taylor (1998); this approach to the control of schistosomiasis is, however, hampered by its high impact on the environment.

Currently, diagnosis followed by appropriate 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 effective against the adult worms (Fig. 1.8).



a) b) c)

Figure 1.8 *Molecular structures of* **a)** *Praziquantel: a chiral centre is indicated;* **b)** *Oxaminquine;* **c)** *Metrifonate.*

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. Therefore there is 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 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).

Oxamniquine has an excellent record of efficacy and safety for the treatment of infections caused by *S. mansoni* (Foster 1987), but is not active against the other human schistosomes.

Metrifonate, is active only against *S. haematobium*; it was rarely employed and is no longer available as a brand product, although it can still be found as a generic drug.

Early diagnosis and treatments remain the principal target in the control strategy, however, the major disadvantages to the use of drugs to control schistosomiasis is the need to repeat drug treatment at relatively short intervals.

1.4 Present research directions

Two main areas of research can be identified, one aiming to produce effective vaccines and the other to increase the limited number of anti-schistosomal 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 re-infection. 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. Combining chemotherapy and vaccine strategies could represent the main possibility for controlling schistosomiasis.

Previous and present research efforts by the WHO aim to control schistosomiasis rather than eliminate it, therefore it is a long-term commitment that subsequently requires a continuous supply of funds.

Schistosomiasis research has been hindered by reduced funding input over the last 20 years. The World Health Organisation (WHO) has expressed its concern on this matter and is keen to promote scientific interest to develop new therapeutic strategies.

The resources necessary for a structural biology project on Schistosomiasis 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 life stages of the schistosomes; (ii) being possible drug targets; (iii) having a demonstrated antigenicity. [1]

[1] For figures see: www.path.cam.ac.uk/~schisto/ and "Topics in International Health – Schistosomiasis

CHAPTER 2

THESIS PLAN

2.1 GST as a vaccine candidate against Schistosomiasis

The World Health Organization (WHO) proposed the Glutathione S-transferase as one of the most promising antigens for the development of anti-schistosomal vaccines and chemotherapies (W.H.O., 1993-1996-2002). As helminths contain very low levels of other detoxification enzymes, such as catalase, superoxide dismutase and cytochrome P450, GST may prevent toxin accumulation in schistosomes and may represent the parasite's primary defence against oxidative damage.

Despite the extensive study of these enzymes, their catalytic mechanism is still incompletely understood and remains controversial. All GSTs operate via activation of GSH to GS^-; this requires to lower the pK_a of GSH by at least two or three pH units. In the GSTs belonging to the classes alpha, mu, pi and sigma, GSH activation proceeds via the interaction with a tyrosine at H-bonding distance from the sulphur of GSH; in the enzymes from other classes the catalytic residue is either cysteine or serine (Armstrong *et al.*, 1997; Sheehan *et al.*, 2001).

The first crystal structure of the 28kDa GST from *S. haematobium* (Sh28GST) has been resolved in our laboratory (Johnson *et al.*, 2003). Based on structural considerations, we have recently presented a novel mechanism for GSH activation (Angelucci *et al.*, 2005 and attachment 2), that may possibly be extended to other GST classes and may be relevant to the search of new anti-schistosomal therapies. Sh28GST belongs to the sigma class and the catalytic Tyr10 adopts two alternative positions with respect to its orientation in the G-site: an activating conformer ($Tyr^{in}10$), where the phenoxyl group of Tyr10 projects towards the G-site forming a hydrogen bond with the thiolate of GSH; and a previously undescribed non-activating

conformer ($Tyr^{out}10$), where the phenoxyl ring is positioned out of the G-site and is stabilised by a π-cation interaction with the guanidinium group of Arg21.

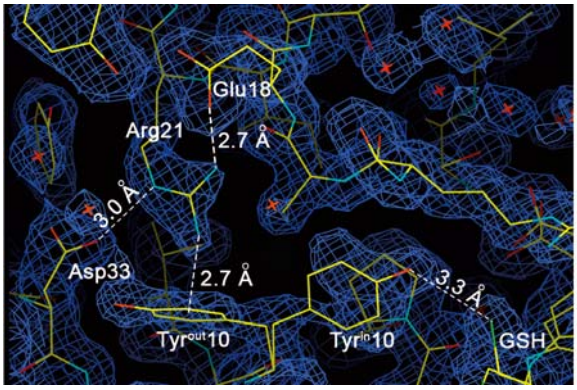


Figure 2.1 *Electron density map showing the double conformation of Tyr10 in Sh28GST and the distances between guanidinium group of Arg21 and Asp33 and Glu18. The π-cation interaction between Arg21 and $Tyr^{out}10$ is calculated with respect to the centre of the phenolic ring.*

Electrostatic interactions between electropositive groups and the electron-rich π-clouds of aromatic rings are common in proteins and have been assigned both structural and functional roles. The functional role of such an interaction has recently been elucidated for Sh28GST (Angelucci *et al.*, 2005) where Tyr10 is proposed to switch alternatively towards the bound GSH and towards the solvent where the ionization of the catalytic tyrosine is favoured, due to the pK_a decrease caused by the π-cation interaction. The role of Tyr10 is fundamental for the catalysis, as demonstrated by the functional studies carried out on the inactive Y10F mutant, which presents the $Phe^{out}10$ conformer stabilised by the π-cation interaction. The role of Arg21, a strictly conserved residue in the GST enzymes containing a catalytic tyrosine, is

therefore suggested to be linked to interaction with Tyr10. Stenberg *et al.*, (1991) carried out basic kinetic studies suggesting the importance of arginines in GSH and inhibitor binding. To date, Arg21 mutants have yet to be crystallised. Therefore, in order to prove the influence of Arg21 in catalysis and to investigate about its suggested structural role, we have mutated arginine 21 in leucine and in glutamine.

Moreover, GSTs are related to various human pathologies and the structure-function studies of their interactions with ligands have a strong pharmaceutical relevance (Armstrong *et al.*, 1997), as demonstrated by the fact that new GST inhibitors are presently being searched for by several pharmaceutical companies.

The structural and functional characterisation of arginine mutants of Sh28GST and of the complex with one inhibitor, S-hexyl glutathione (GTX), were the aim of this thesis. All proteins were heterologously expressed in *E. coli* cells. Proteins were purified to homogeneity and suitable crystallisation growth conditions were determined. The structural characterisations were achieved through X-ray crystallography, while functional studies employed static and time resolved spectroscopy. High resolution diffraction data were collected and the resulting electron density maps were used together with molecular modelling techniques to build and refine the final protein structures. The structural data obtained imply new chemical properties of this enzyme which may help to provide a better insight into the overall enzymatic mechanism and may help to select new portions of the polypeptide chain as possible targets for drug design.

2.2 Other target proteins:

As an effective drug without serious side-effects, praziquantel (PZQ) is the single major treatment for schistosomiasis and several other trematode and cestode infections (Andrews *et al*., 1983), although its target and mechanism of activity are not known (Shekhar, 1991). Recent reports of schistosome strains resistant to PZQ and other anti-helminths indicate that new therapies must be identified to fight this debilitating disease (McTigue *et al.*, 1995).

2.2.1 Fatty acid binding protein

The World Health Organization selected Sm14, a fatty acid binding protein (FABP), as one out of six anti-schistosome vaccine candidates (Table 1.1). The physiological role of FABPs includes protection of cell membranes and enzymes from the effect of high concentrations of free fatty acids (FAs) and of their acyl-CoA derivatives, storage of FAs, lipid trafficking, and regulation of cell growth and differentiation. These proteins have been intensively studied because schistosomes lack the metabolic pathways required for the biosynthesis of sterols and lipids, and completely depend on the host for these substances. Uptake and transport of fatty acids and other lipids in *S. mansoni* depend on the fatty acid binding protein (Sm14). Sm14 is present in all the stages of the life cycle and is localized in the external cell layer, i.e., near the interface of the parasite/host contact.

During my Ph.D., I have been involved in the project regarding the functional and structural study of Sm14FABP, which is not discussed in this thesis (see attachment 1).

2.2.2 Cyclophilin: a binding protein of an antischistomal drug

A new protein physiologically relevant in the development of schistosomes, recently proposed as a possible vaccine candidate antigen (Al-Sherbiny *et al.*, 2003), was identified by our collaborators of the Institute of Cellular Biology (CNR, Monterotondo, Roma): a cyclophilin-like protein from *Haemonchus contortus*, a parasite affecting wild ruminants.

Cyclophilins are interesting proteins, implicated in many biological processes, including binding to Cyclosporin A (CsA) and have an enzymatic activity as prolyl isomerases. The immunosuppressive drug CsA, is widely used in the prevention of graft rejection and in the treatment of immune disorders. It has also antiparasitic effects as reported for schistosome, plasmodia, cestodes and toxoplasma. Treatment with CsA resulted in a drastic reduction in the number of schistosomes, especially in immature worms (Klinquert *et al.,* 1995). The presence of some cyclophilin in schistosomes has been widely reported (Argaet *et al.*, 1992, Klinkert *et al.*, 1995, Klinkert *et al.*, 1996, Kiang *et al*., 1996) but no structure is available, despite of the fact of a structure could reveal some new information on the role of this proteins in that disease. Therefore, during this research project, I have been involved in cloning the cyclophilin-like protein from the adult worm cDNA of *S. mansoni* (SmCyp), and in the expression, purification and functional characterisation of the rotamase activity of a cyclophilin-like protein from *H. contortus* (HcCyp).

CHAPTER 3

GLUTATHIONE S-TRANSFERASE

3.1 Target metabolic pathway: detoxification

Living organisms are continuously exposed to non-nutritional foreign chemical substances. Such xenobiotics may interact deleteriously with the organism, causing toxic and sometimes carcinogenic effects (Sheehan *et al.,* 2001). However, cells possess an impressive array of enzymes capable of bio-transforming a wide range of compounds, altering their chemical structures and creating novel functional groups that may either reduce their toxicity or facilitate their excretion. The enzymatic detoxification of xenobiotics is a complicated metabolic pathway which includes several enzymes involved in the conversion of lipophilic, non polar xenobiotics into more water-soluble metabolites, which are readily eliminated from the cell. In 1947 R.T. Williams defined the field of detoxification and proposed that these non-reactive compounds could be bio-transformed in two phases: functionalization, which uses oxygen to form a reactive site, and conjugation, which results in addition of a water-soluble group to the reactive site. These two steps, functionalization and conjugation, are termed Phase I and Phase II, respectively. Phase I of the detoxification system, mainly carried out by the cytochrome P450 supergene family of enzymes, is generally the first enzymatic defence system against foreign compounds. Phase II conjugation reactions generally follow Phase I activation, resulting in the transformation of the xenobiotic into a water-soluble derivative, excreted in urine or bile. There are several types of conjugation reactions including glucuronidation, sulfation, and glutathione and amino acid conjugation. Therefore, detoxification is not a single reaction, but rather a process that involves multiple reactions and multiple components.

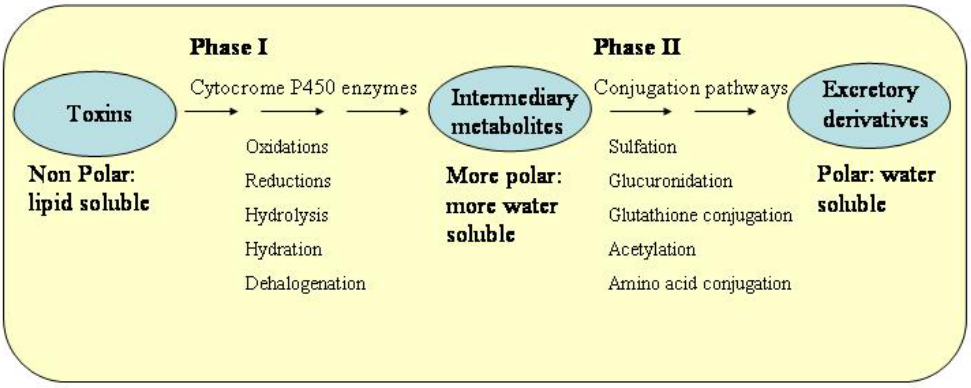


Figure 3.1 *The consequence of detoxification is the bio-transformation of lipophilic compounds into water-soluble compounds, readily excreted in the urine.*

A crucial detoxification enzyme is Glutathione S-transferase (GST) which catalyses the nucleophilic addition of GSH to endogenous and xenobiotic electrophilic toxins, that are subsequently eliminated by several transport mechanisms existing in the cell, specific for glutathione conjugates, including an ATP-dependent GS-X pump. The pharmacological inhibition of enzymes involved in detoxification reactions may induce cell damage either directly or indirectly, via the action of unmodified xenobiotic compounds

3.2 Introduction to Glutathione S-transferases

GSTs have been subdivided into an ever-increasing number of classes based on a variety of criteria, including amino acid/nucleotide sequence identity, immunological and kinetic properties, and tertiary/quaternary structures. Human GSTs are divided into three distinct super families: the membrane bound microsomal, cytosolic and mitochondrial family members. (Townsend *et al*., 2003). GST are of interest to pharmacologists and toxicologists because they provide targets for anti-asthmatic and anti-tumor drug therapies. GSTs catalyse the conjugation of γ-

glutaminylcysteinylglycine (GSH), to a wide variety of endogenous and exogenous electrophilic compounds, such as chemical carcinogens, environmental pollutants and antitumor agents, being implicated in a variety of resistance phenomena involving cancer chemotherapy agents, insecticides and microbial antibiotics. These transferases inactivate endogenous α,β-unsaturated aldehydes, quinones, epoxides, and hydroperoxides formed as secondary metabolites during oxidative stress. These enzymes are also involved in the biosynthesis of leukotrienes, prostaglandins, testosterone and progesterone (Hayes *et al*., 2004).

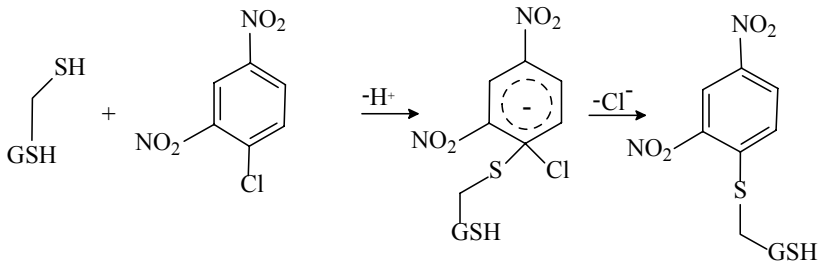
The GSTs catalyze the general reaction shown in Scheme 3.1:

Scheme 3. 1

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

A general assay to determine enzymatic activity of GST, RX is 1-chloro-2,4-dinitrobenzene (CDNB).

Scheme 3. 2



A great contribution to the comprehension of GST catalytic mechanisms has been given from structural analysis which helped to understand how the enzyme recognizes and activates glutathione for nucleophilic attack and how,

or even if, the enzyme specifically recognizes electrophilic substrates (R-X). Representative crystal structures are available for most classes and, despite limited overall sequence identity, these follow generally similar folds, with structural differences concentrated especially around the active site and the inter-subunit interfaces. Each subunit is composed of two domains: an N-terminal domain (domain I) that adopts a thioredoxin-like fold and an all α-helical C-terminal domain (domain II). There are at least two ligand-binding sites per subunit: the glutathione-binding site (G-site), which is very specific for GSH and the hydrophobic substrate-binding site (H-site), which can bind a large variety of different electrophiles. The G-site is constructed mainly from residues of the N-terminal domain, whereas the H-site has major contributions from the C-terminal domain. Whilst the functional properties of the amino acid residues making up the G-site of a GST are generally conserved among different classes, the residues forming the hydrophobic substrate-binding pocket vary considerably among different GSTs. Since the structure of the H-site governs the substrate specificity of a particular GST, diversity in the H-site gives the GST family the ability to catalyze reactions toward a large number of structurally diverse substrates. GST from mammalian sources have been well characterized, but studies of GSTs from non-mammalian sources have revealed the existence of several new classes and thus greatly extended our knowledge of the structural and functional diversity of these proteins (Sheehan *et al*., 2001).

3.3 GST antigenic properties in schistosomiasis

In 1987, Balloul *et al.*, suggested the 28 kDa glutathione-S-transferase from *Schistosoma mansoni* (Sm28GST) as a potential vaccine candidate,

demonstrating that antibodies against Sm28GST antigen have moderate effect on the reduction of worm burden, its very strong influence on the female worm fecundity and egg viability. Sm28GST actively participates in the elimination of toxic products through its GST activity and in parasite fertility probably through its PGD_2 synthase (PGDS) activity. Therefore, inhibition of the enzymatic (GST and PGDS) activities of Sm28GST by neutralizing antibodies generated after immunization with the protein might be detrimental for the parasite and could explain why this antigen is presently considered as an important vaccine candidate against schistosomiasis. Another particularly interesting regulatory phenomenon caused by *S. mansoni* infection is the retention of activated Langerhans cells (LCs), a subpopulation of epithelial dendritic cells (DCs), in the epidermis for at least 48h post-infection (Herve *et al.,* 2003).

It is noticeable that a significant anti-fecundity effect has also been observed in the case of 28kDa GST from *S. japonicum* (Taylor *et al.,* 1998). These observations suggest that schistosome 28 kDa GSTs could be considered as efficient candidates to reduce morbidity. Pre-clinical studies have therefore been initiated with the 28 kDa GST from *S. haematobium*.

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 schistosome GST as vaccine candidate has been strengthened by vaccination experiments performed in Sudan against cattle schistosomiasis due to *S. bovis* (Bushara *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 worm fertility 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* (Bushara *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 (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.

3.4 The 28kDa Glutathione S-transferase from S. haematobium

The information 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 hypothesize the effect of mass vaccination in man using Sh28GST. Moreover, the *S. haematobium* protein was 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.

Schistosomal 28 kDa GSTs belong to the sigma class of GSTs based on sequence comparison. These GSTs were shown to have high transferase activity with model compound (CDNB), fatty acid hydroperoxide-GSH peroxidase activity (Taylor *et al*., 1988) and eicosanoid synthesis (Hervé *et al.*, 2003). Thus, Sh28GST is a multifunctional enzyme which plays key role in the host-parasite interactions.

3.5 Materials and Methods

3.5.1 Cloning, Expression and Purification of Wild Type and Mutant Sh28GSTs

The wt gene was amplified by PCR using KOD HotStart DNA polymerase (Novagen) using Sh28GST held in the pCR 3.1 vector (a kind gift from Prof. F. Trottein, Lille, France) as the template for the reaction, and sub-cloned into the bacterial expression vector pET23b (Novagen). The primers used in the reaction were as follows: forward 5'-ATA T**CC ATG G**CT GGT GAT CAT ATC AAG G-3'; and reverse, 5'-G**GC TAGC** CTA GAA GGG AGT TGC AGC CC-3'. *Nco*I and *Nhe*I recognition sites (indicated in boldface) were incorporated into the forward and reverse primers, respectively, to facilitate directional cloning. Successful amplification was confirmed by agarose gel electrophoresis.

Site-directed mutagenesis of wt/pET-23d to generate the two Arg21 mutants (R21L and R21Q) was carried out using the QuickChange Site-directed Mutagenesis kit (Stratagene) according to the manufacturer's instructions. The primers for the R21L mutant were: forward 5'-CGC GGA CGA GCT GAA TCG ATC **CTG** ATG ACA CTT GTG GC-3'; and reverse 5' –GCG CCT GCT CGA CTT AGC TAG **GAC** TAC TGT GAA CAC CG-3'. The primers for R21Q were as above but with the substituted codon **CAG** in 5' forward primer and **GTC** in 5' reverse primer. PCR amplification was carried out at annealing temperatures of 55°C (R21L) and 45°C (R21Q). The fidelity of the amplification reactions was confirmed by DNA sequencing at the BioMolecular Research Sequencing Service (University of Padova, Italy).

Escherichia coli BL21 (DE3) pLysS bacterial cells (Novagen) were transformed with the wild type and mutant plasmids and grown at 37°C in 500 ml Luria broth supplemented with 50 μg/ml ampicillin until an A^{600} of 0.5 was reached. Protein expression was induced upon addition of 0.2mM isopropyl-β-D-thiogalactopyranoside (Sigma-Aldrich), incubating for a further 5h. Bacterial cells were harvested by centrifugation at 11000g for 15 min. The bacterial pellet (wt and R21L) was resuspended in 30 ml phosphate-buffered saline (PBS), pH 7.4 containing 10 mM β-mercaptoethanol (β-ME). Cell lysis was achieved by sonication in bursts of 4s at 9s intervals for 5 min. The sonicated extract was centrifuged at 13800g for 20 min to recover the supernatant. All proteins were filtered across a membrane with a pore size of 0.20 μm and were subsequently purified through a 5ml GSTrap™ HP column (Amersham Biosciences), pre-equilibrated with PBS pH 7.4, according to the manufacturer's instructions. GSH-bound wt and R21L proteins were eluted with 50 mM Tris-HCl, pH 7.3, 0.1 M NaCl containing 10mM GSH. Peak fractions containing pure protein (determined by SDS-PAGE analysis) were pooled and extensively dialyzed vs. PBS, pH 7.4, containing 10 mM β-ME, prior to concentration by ultrafiltration using Amicon Ultra-15 centrifugal filter units (Millipore) (membrane cut off 10000 Da).

The R21Q bacterial pellet was resuspended in 20mM Tris, pH 8.3 and sonicated as previously described. The cell extract was centrifuged as above and purified by a Sepharose 15Q ion-exchange column (Amersham Biosciences), pre-equilibrated with 20 mM Tris, pH 8.3. Bound protein was eluted over a salt gradient (0-0.5M NaCl) with the same buffer and subsequently dialysed and concentrated as above. The R21Q was also purified by a GSH-trap column, as described for R21L.

The wild type in complex with S-hexyl glutathione (GTX), was prepared using the purification procedure described above but eluting with 20 mM GTX instead of GSH and was concentrated to 10 mg/ml. To obtain GSH-free proteins, Sh28GST and R21L were eluted with 20 mM glycine in PBS pH 10.

Protein concentrations were routinely determined spectroscopically, measuring the A_{280} using a HP 845X UV-visible spectrophotometer (Hewlett-Packard) using an extinction co-efficient of 0.96 mg/ml for all proteins. Proteins purification were carried out by FLPC using an Akta Prime purification system.

3.5.2 Crystallisation

Crystallisation of wild type and mutant GSTs were generated by the hanging drop vapour diffusion method at 21°C.

Tiny crystals of R21L grew within three days in a drop containing a 1:1 ratio of protein (15mg/ml) to well solution (2.3 M ammonium sulphate, PBS pH 7.4), 5 mM β-ME and 10% PEG 200). They were used in microseeding trials to increase the crystal size according to standard hanging drop protocols. The hanging drop consisted of a two parts R21L (15mg/ml), one part crystal seeds (diluted 1/1000 with well solution) and one part well solution (1.8 M ammonium sulphate, 100 mM PBS pH 7.4, 5 mM β-ME and 10% PEG 200).

R21Q (10mg/ml) crystals were obtained in similar conditions as for R21L; the well solution contained 2.5 M ammonium sulphate, 0.2 M MES pH 6.0 and 5 mM β-ME. Many small crystals grew after few days. The quality of R21Q crystals was improved by microseeding techniques using 0.5 µl of

protein, 0.5 µl of seeds (diluted with well solution 1/10) and 1 µl of well solution.

Crystals of GTX-bound wt were obtained by co-crystallisation using a protein solution of Sh28GST 10mg/ml in PBS, pH 7.4, 5 mM β-ME, 10 mM GTX and a well solution of 20% PEG 3350, 0.2 M MES pH 6.0 and 5 mM β-ME. Very large crystals grew after a few days.

Crystals of GSH-bound wt were grown at pH 6.0 at a 1:1 ratio of protein (15mg/ml contained in PBS pH 7.4, 5 mM β-ME and 10 mM GSH to well solution of 0.2 M imidazole/malate buffer pH 6.0 containing 10% PEG 4000.

Crystals of GSH-bound R21L were grown using a 1:1 ratio of protein (10mg/ml) in PBS containing 5mM β-ME and 10mM GSH to a well solution of 0.1M sodium acetate pH 5.5 containing 25% PEG 5000 MME.

Crystals of all proteins were cryo-cooled by stepwise soaking in the well solution containing increasing concentrations of glycerol until a final glycerol concentration of 18% was achieved.

3.5.3 Data Collection and Processing

Diffraction data for all crystals were collected at ~100 K at resolutions ranging from 2.0-2.5 Å at the Elettra Synchrotron facility (Trieste, Italy), at the European Synchrotron Radiation Facility (ESRF; Grenoble, France) and at Deutsches Elektronen Synchrotron (DESY; Hamburg, Germany). Space groups were determined by autoindexing and processed using DENZO and SCALEPACK in HKLsuite (Otwinowski *et al.,* 1997). The statistics of crystallographic data collection for R21L, R21Q and Sh28GST in complex with GSH (pH 6.0) and GTX are given in Table 1.

3.5.4 Molecular Replacement, Model Building, and Refinement

The 3-D structure of mutant and wild type proteins were determined by molecular replacement methods. The structure of Sh28GST was used as a model (PDB entry 1OE7), resulted in a clear rotation and translation solution using the program AMoRe (Navaza, 1994) from the CCP4suite. Further refinement was carried out using alternate steps of least-squares refinement using REFMAC5 (Murshudov *et al.,* 1996) and fit to generated electron density maps using Xtalview/Xfit (McRee, 1999). Water, PEG and β-mercaptoethanol molecules were added according to peaks observed in the electron-density maps. The free R-factor, based on 5% of the data, was used throughout to guide the refinement procedure. All data were refined to give satisfactory final R factors as reported in Table 1.

Interestingly, the autoindexing of R21Q data set produced a list containing two likely choices for the space group: $C222_1$ with a monomer in the unit cell, and $P2_1$ with a dimer in the unit cell. The same problem was found in the case of wild type (Johnson *et al.*, 2003). The final refinement using $C222_1$ has good statistic parameters, but it was impossible to generate a proper dimer in the unit cell using the symmetry operators. Therefore, this data set was indexed and processed again as $P2_1$ space group, with two molecules in the asymmetric unit. The statistic parameters are satisfactory as showed in Table 1.

3.5.5 Quality of the Structural Data

The electron density for the final models of R21L and wt in complex with GTX and GSH are well-defined for residues 4–211 in both chains A and B. The electron density for the models of R21Q and GSH-bound R21L, instead,

are not well-defined in the C-terminal region and lack residues 207-211 in the monomeric chain A. All the structures display good geometry and no residues in the disallowed region of the Ramachandran plot (Table 1).

Table 1. *Summary of crystallographic data. R_{free} was calculated from 5% of reflections taken from the raw data.*

	R21L	R21L+GSH	R21Q	Sh28GST+GTX	Sh28GST+GSH
Data collection statistics					
Space Group	I23	$P222_1$	$P2_1$	$P3_2$	$P322_1$
Unit Cell Dimensions (a, b, c in Å)	148.7 148.7 148.7	53.1 53.1 141.8	53.5 77.4 53.5	52.8 52.3 141.4	53.3 53.3 142.3
No. of unique reflections	24373	15440	29520	18295	8261
I/σ	10.2	19.9	14.9	11.3	10.2
Completeness (%)	99.8	99.3	98.2	99.0	99.1
Average Redundancy	22.0	5.6	4.6	3.5	8.3
Rmerge	0.063	0.055	0.053	0.122	0.070
Refinement Statistics					
Resolution (Å)	20-2.3	20-2.0	20-2.0	20-2.2	20-2.5
R	0.206	0.189	0.219	0.178	0.194
R_{free}	0.270	0.239	0.292	0.243	0.252
Rms deviations in bond length	0.013	0.009	0.012	0.011	0.013
Rms deviations in bond angles	1.875	1.208	1.3	1.432	1.483
Ramachandran Plot					
Residues in allowed regions (%)	99.5	100	99.3	99.7	99.6
No. Molecules/Asymmetric unit	2	1	2	2	1

3.5.6 Steady State Determinations

The catalytic activities of wild type and mutant GSTs were determined spectroscopically using a HP845x UV-visible spectrophotometer (Hewlett-Packard) under steady state conditions, using the chromogenic substrate 1-chloro-2,4-dinitrobenzene (CDNB; Sigma-Aldrich). Experiments were carried out measuring the absorbance increase at 340nm at 20°C over 120s. The assay was initiated upon addition of the enzyme to a 1 ml quartz cuvette containing GSH and CDNB in 0.1 M potassium phosphate pH 7.0. Experiments were carried out, both at constant GSH (2mM) and CDNB concentrations (2mM). Wild type and R21L protein concentrations used were 0.1 μM and 2.4 μM, respectively. Data were fitted using the Origin 7 Server Software.

The ability of GTX to inhibit the enzymatic activity of Sh28GST and R21L was also measured using the above steady state conditions at constant CDNB concentration (2mM), over a range of GSH concentrations (0.5mM, 1mM, 2mM and 4mM), for three diverse GTX concentrations (150μM, 250μM and 500μM). Assays were initiated upon addition of the wt (0.1μM) or R21L (2.4μM).

The influence of pH on catalytic activity was investigated for Sh28GST and R21L, repeating the above steady-state experiments over a pH range of 6.0 to 9.0. The appropriate pH conditions were prepared by mixing 50mM sodium phosphate with 50mM sodium borate. Reactions were carried out as above at GSH concentrations ranging from 200μM to 4mM**.** Observed reaction velocities were corrected for spontaneous reaction rates prior to addition of

the enzyme, when necessary. The pH dependence of kinetic parameters was analyzed using MATLAB 5.3.

3.5.7 Fluorescence Measurements

The binding of GSH to wt and R21L was followed measuring intrinsic fluorescence quenching of the enzyme at 20°C, upon ligand addition, using a Spex Fluoromax spectrofluorimeter (excitation wavelength was 280 nm, emission range 300÷400 nm). 1 μM of protein was added to a 2 ml cuvette containing 0.1M potassium phosphate pH 7.0. Similar experiments were carried out for GTX using protein concentrations of 1μM and 0.5μM for wt and R21L, respectively. All experimental data were analysed using MATLAB 5.3.

3.5.8 Stopped-Flow Rapid Mixing Experiments

Rapid mixing experiments were carried out at 20°C using an Applied PhotoPhysics stopped-flow syringe drive. The dead time of the instrument is approximately 2ms. The binding of GSH to wt and R21L was measured, following the decrease in tryptophan fluorescence (excitation wavelength 280nm, 32 nm filter). Asymmetric volumes of protein and GSH, were mixed to give a final protein concentration of 5μM (wt) and 10μM (R21L) and a range of GSH concentrations (15-50μM). Wild type experiments were carried out in 0.1M potassium phosphate buffer pH 7.0 whereas R21L experiments were carried out in 0.1M potassium phosphate buffer pH 8.0. Reactions were recorded over 0.2s. Averages of multiple traces collected for each GSH concentration were fitted to a single exponential and used to determine the k_{off} of GSH for the protein.

The ability of wt and R21L to deprotonate GSH was also assessed by rapid mixing experiments, under the above conditions, following the absorbance increase at 240nm at diverse GSH concentrations (15-50μM for wt and 125-500μM for R21L).

3.5.9 Experimental Determination of the pK_a of Tyr10

The ionization of Tyr10 was measured spectrophotometrically using a HP845x UV-visible spectrophotometer (Hewlett-Packard) as a function of pH, measuring the tyrosinate concentration from the intrinsic absorbance at 293nm ($\Delta\varepsilon$ of 2400 M^{-1} cm^{-1}). Mixtures of 50mM sodium borate and 50mM sodium phosphate were prepared to cover a pH range from 6 to 10. Difference spectra at each pH were recorded, subtracting the spectra generated for the protein at pH 6.0. The concentrations of R21L and R21Q ranged from 15 to 35 μM.

3.6 Results

3.6.1 Cloning, Expression and Purification of Wild type and mutant Sh28GSTs

The cloning and expression of soluble wtGST, R21L and R21Q was carried out in bacterial cells according to materials and methods. Successful expression was confirmed by SDS-PAGE under reducing conditions. A band was observed at 24 kDa, corresponding to the predicted M_r value of the GST monomer. Purification of wt, R21L and R21Q was routinely carried out yielding 20-30 mg/500 ml culture of pure protein, as judged by SDS-PAGE analysis.

3.6.2 Structural Analysis of wtGST in complex with GSH and GTX

Structures of wtGST containing only Tyr^{in}10 were obtained, both with GSH and the competitive inhibitor GTX at pH 6.0. Only GTX-bound wt will be further discussed as all structural features are similar for the GSH-bound enzyme.

In the inhibitor complex the hydrophobic site, which binds the electrophilic substrates, lies adjacent to the G-site and at the interface between the N- and C-terminal domains. GTX inhibits GSTs by inserting it's hexyl chain in the H-site, preventing the conjugation of GSH to its electrophilic substrate. In GTX-bound wt, the H-site is characterised by non-specific hydrophobic interactions. The typical polar interaction between the sulphur of GTX and the hydroxyl group of Tyr10 (3.1 Å) was present as well as the other characteristic interactions between the GSH backbone and the region of the polypeptide chain that forms the G-site (Fig. 3.2).

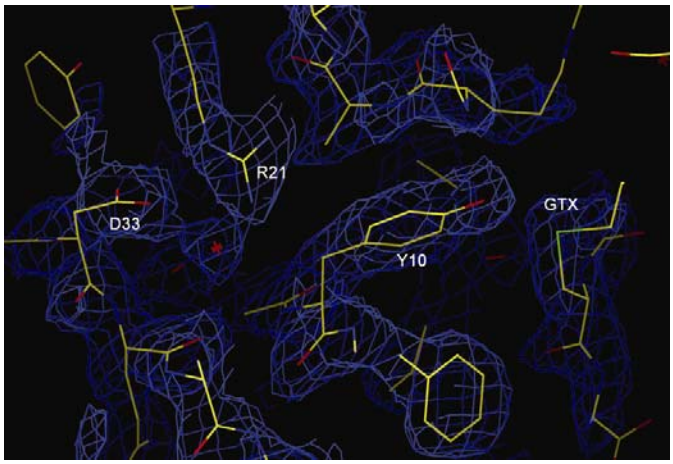


Figure 3.2. *Electron density map ($2F_o$-F_c at 1σ) of the GTX-bound wtGST showing the polar interaction between the hydroxyl group of the Tyr^{in}10 and the sulphur atom of GTX.*

The presence of the inhibitor stabilises the overall structure as confirmed by the well-defined electron density map. In the G-site there is a cluster of aromatic residues namely Tyr10, Phe11, Phe38 and Phe211. The C-terminal region was observed to interact with the N-terminal region specifically at the G-site, and hydrogen bonds were identified between Arg206-Glu18, Thr209-Arg14 and Tyr202-Glu18. These findings are in agreement with the data presented in the literature (Ibarra *et al.,* 2003 and Ibarra *et al.,* 2001).

A peculiarity in the GTX-bound wt structure was evident, regarding the position of Arg35, which is interestingly located in the cavity where the $Tyr^{out}10$ is found, forming a hydrogen bond with Asp33 and Asn12 (Fig. 3.3a). Arg35 is in close proximity to the G-site and it adopts a different position according to Tyr10 orientation.

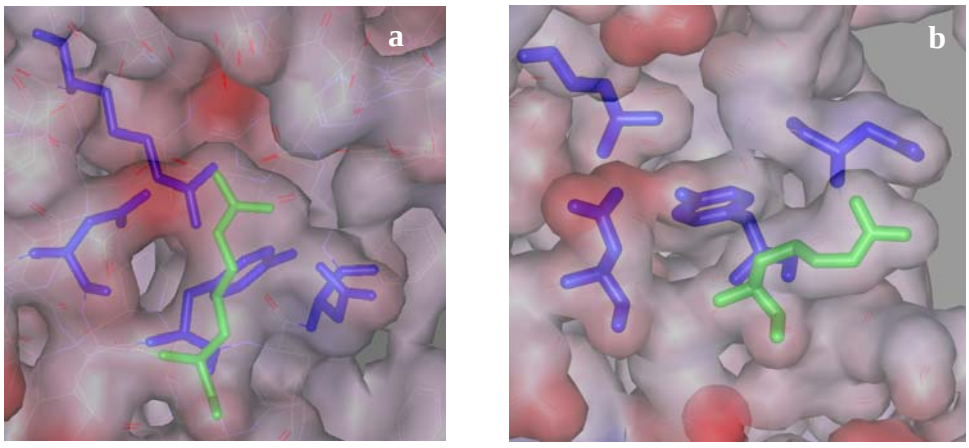


Figure 3.3 *Comparison between Arg35 positions in the Arg21 pocket in (**a**) GTX-bound GST ($Tyr^{in}10$) and (**b**) R21L ($Tyr^{out}10$). Tyr10, Asn12, Arg21 and Asp33 are represented as blue sticks. Arg35 is represented as green sticks. Both figures were generated using MSViewer.*

In order to extend the possible role of Arg35 to other GST classes, the high conservation of Arg35 was established, searching the entire SwissProt database, using the query motif: [Y] 7X [E,Q,H] 2X [R] 10-11X [E,D] 1-2X

[R,K]. This motif was found in 65 out of 95 alpha, mu, pi and sigma class GSTs, containing a tyrosine in the catalytic site. This motif was not found in the 83 other GSTs containing a catalytic serine or cysteine residue. Structural analysis of all existing crystallographic data for sigma class GSTs (only $Tyr^{in}10$) also highlighted the conservation of the interaction between Arg35 and Asp33 which, from our structure, was found to be responsible for the closure of the $Tyr^{out}10$ pocket.

Furthermore, the GTX- and GSH-bound states of wt have structured C-terminal regions that interact with the N-terminal region. Principal contacts are made between the residues presented in Table 3.2

Contact pairs		Distance
Phe211(O)	Arg14(CB)	3.17 Å
Phe211(O)	Gly15(N)	2.77 Å
Tyr202(OH)	Glu18(CD)	3.30 Å
Thr209(O)	Arg14(NH2)	3.15 Å

Table 3.2. *Contacts between C- and N-termini in GTX- and GSH-bound wt at pH6.0*

3.6.3 Structural Analysis of R21L and R21Q

There is no significant difference between the tertiary structures of R21L, R21Q and wt, not unexpected given that the incorporated mutation is a single amino acid substitution. Each protein is a symmetric homodimer. The N-terminal regions of each monomer (residues 1-88) have the same GST thioredoxin-like fold constituted by four-stranded β-sheets flanked by three α-helices. The C-terminal region (residues 89-211) is composed of 6 α-helices and differs in mutant and wild type proteins only with respect to the degree of structural constraint and location as further described. The N-

terminus forms the G-site which contains the catalytic tyrosine and binds GSH. The principal structural differences between mutant and wild type structures occur in the flexible N-terminal loop (residues 12-35) located in close proximity to the active site. Moreover in chain B of R21L there is another portion of the N-terminal domain (residues 59-65) which seems to be very flexible since it is characterized by poor electron density. The N-terminus forms the G-site which contains the catalytic tyrosine residue and binds GSH.

Structurally, the R21Q mutant is similar to R21L as revealed by superimposition giving a RMSD value of 0.87 Å. The high value of RMSD is due to the absence of the last five amino acidic residues (206-211) of R21Q which do not appear in the electron density maps, therefore it is not possible to understand where the C-terminal chain is positioned with respect to the G-site.

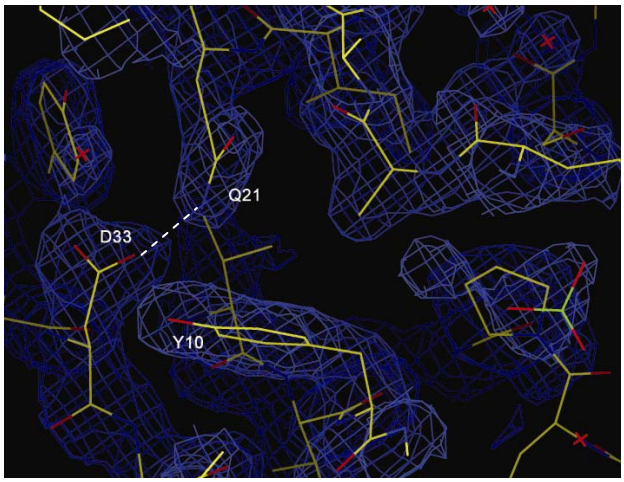


Figure 3.4: *Electron density map of R21Q ($2F_o$-F_c at 1σ) showing Gln21 at hydrogen bond distance from Asp33.*

In the structure of GSH-bound wt at pH 7.4 (Johnson *et al.*, 2003), Tyr10 adopts a double conformation, alternating both 'in' and 'out' of the G-site. A double conformation of the residue in 10 position, is also present in the tyrosine to phenylalanine mutant structure (protein kindly provided by M. Hervè, J. Fontaine and F. Trottein from Pasteur Institute, Lille, France). The interaction between the aromatic ring of Phe10 and Arg21 is clearly evident in the 2.0Å structure solved in our laboratories and not reported in this thesis. The Phe^{out}10 conformer is populated despite exposure of the non polar amino acid to solvent. It was expected that mutating Arg21 would stabilize the Tyr^{in}10 conformation, however Tyr^{out}10 was exclusively observed in both arginine mutants, confirming that the role of the π–cation interaction is not structural (Fig.3.4, 3.5). Mutating Arg21 to Gln results in the loss of the interaction between Glu18(OE2) and Gly13(O), however the hydrogen bond between Asp33(OD1) and Gln21(OE1) is maintained (2.88 Å) (Fig.3.3). Mutating Arg21 to Leu disrupts three important polar interactions that are formed between Arg21(NH1) and Asp33(OD1), Glu18(OE2) and Gly13(O) at distances of 3.12 Å, 2.79 Å and 2.84 Å, respectively (Fig. 3.5).

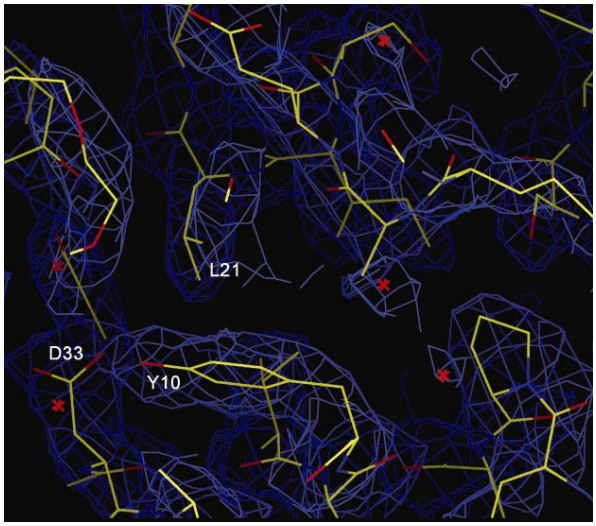


Figure 3.5 . *The electron density map ($2F_o$-F_c at 1σ) of R21L showing the exclusively populated Tyr^{out}10 conformer at hydrogen bond distance with Asp33. Leu21 is at a distance of 4Å from Tyr10.*

A comparison of the contacts between Tyr^{out}10 and other active site residues for the GSH-bound wt reported by Johnson *et al.,* (2003) and R21L was carried out as described in material and methods. In Sh28GST, there are no interactions between Tyr^{out}10 and other residues at a distance shorter than 3 Å except for Arg21. For R21L however, the hydroxyl group of Tyr10 forms a hydrogen bond (2.63Å) with Asp33(OD2). In wt these residues are too distant from one another (4.2Å) to make proper contact. It was unclear whether this interaction accounted for the Tyr^{out}10 conformer position rather than the loss of the π-cation interaction with the guanidinium group of Arg21. Analysis of the structure of R21Q shows that Tyr^{out}10 does not interact with Asp33(OD2), therefore we conclude that this interaction is not responsible for this conformer position.

3.6.4 Structural Analysis of GSH-bound R21L

Overall, GSH-bound R21L appears to be less ordered than ligand-bound wt. There is significantly less-defined electron density particularly in the active site and the C-terminal regions. The electron density corresponding to the last four C-terminal amino acids of GSH-bound R21L is absent, implying that this region is flexible. Tyr10 is found in a double conformation and, the H-bond between the hydroxyl group of Tyr^{out}10 and Asp33(OD2) is maintained (2.63Å) (Fig. 3.6). Tyr^{in}10 interacts with GSH in the G-site similarly to the wild type, however the distance between Tyr10(OH) and GSH(S) is longer (3.88 Å) than in wild type and there are additional interactions between the

guanidine group of Arg16 and GSH. The active site loop appears to be stabilised by polar interactions formed between Asn12-Gly15, Arg14-Asp172, Arg14-Asp168 and Arg14-His169. Finally, several side chain of the site loop (11-18, 33 and 35), display poorly defined electron density, confirming that structural changes occur in this region.

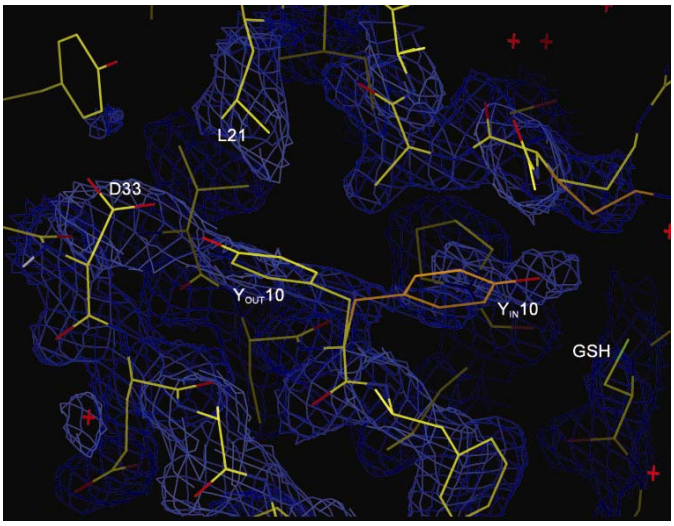


Figure 3.6: *Electron density map ($2F_o$-F_c at 1σ) showing Tyr10 double conformation. The H-bond between $Tyr^{in}10$ and the sulphur atom of GSH and the H-bond between $Tyr^{out}10$ and Asp33 are shown.*

3.6.5 Structural Comparisons of wild type and the arginine mutants

In order to understand the conformational changes occurring in the Tyr^{out} pocket, attempts were made to block Tyr10 in a single conformation and we considered two opposing models for comparison: GTX-bound wt (exclusively $Tyr^{in}10$) and R21L (exclusively $Tyr^{out}10$). Characterisations and comparisons of the $Tyr^{out}10$ pocket were made using the Cast-P program (http://cast.engr.uic.edu/cast/) (Table 3.3). Two positions of the active site are particulalry worth mentioning: Asn12 and Arg35. In GTX-bound wt Arg35

interacts via hydrogen bonds with Asp33 and Asn12, however in R21L such interactions are lost because this cavity is fully occupied by the $Tyr^{out}10$ conformer. For GTX-bound GST, the pocket is closed to the solvent (no mouth openings) and has a molecular surface area of 41 $Å^2$. For R21L the pocket contains one mouth opening and has a larger molecular surface area (169 $Å^2$). Therefore, it can be seen that the catalytic mechanism and flipping of the tyrosine is coupled to a change in the molecular surface area and solvent accessibility of the $Tyr^{out}10$ pocket. These changes are due to a concerted rearrangement of amino acids that form this pocket. Analysis of the B-factors for such pocket amino acids indeed indicates that several key residues exhibit higher mobility than the rest of the molecule. The B value for the side chain of Arg35 in wild type is slightly higher (46$Å^2$) than the B-factor average for the side chains of the entire protein (41$Å^2$). B-factor value for Arg35 in GST+GTX is comparable with the average of the protein, $B_{average}$=20$Å^2$ and 21$Å^2$ respectively.

Comparisons between pocket volumes (Table 3.2) confirm that this pocket is bigger in wild type enzyme and is more accessible to solvent.

Molecule	**N_mth**	**Area_ms ($Å^2$)**	**Vol_ms ($Å^3$)**
Wt + GSH	1	100	86
Wt + GTX	0	74	41
R21L	1	169	153
R21L + GSH	2	280	255
R21Q	2	189	281

Table 3.3 *The table shows the number of mouth openings for each pocket (N_mth), the volume (Vol_ms) enclosed in the molecular surface area (Area_ms) of the $Tyr^{out}10$ cavity of ShGST+GTX, R21L, R21L+GSH and R21Q, as calculated with the program Cast-P using a probe radius of 1.0Å (Liang et al., 1998).*

For R21L and R21Q, pockets were selected by substituting Tyr^{out} with an Ala residue with the same Cα and Cβ coordinates. For WtGST+GTX and R21L+GSH the pockets were selected according to the presence of the Arg21(NH2) and the C-alpha and C-beta of the catalytic Tyr.

3.6.6 Interactions at the Dimer Interface

R21L and R21Q were confirmed to be dimers by sedimentation equilibrium ultracentrifugation (results not shown). To ascertain that no structural rearrangements are present at the dimer interface of R21L, the inter-subunit contacts within a distance of 3.5Å, were compared to those present in GTX-bound wt. In R21L, the inter-subunit contacts occur between residues 68–104 of each monomer. Overall, there seems to be minor reshuffling of interactions coming from the loss of several interactions and the gaining of others. All contacts are lost between Arg52 (α2-helix of domain I) of monomer A with Ile138 and Asp104 in monomer B. The electron density for the side chain of Arg52 in both monomers is not well-defined and it is lost in the solvent. Furthermore, hydrogen bonds (2.85Å) are lost between Lys66A(NZ)-Glu90B(OE1) and between Lys81A(NZ)-Glu89B(OE2) (3.12Å in wt). Two interactions are lost between Glu96A(C)-Arg76B(NE) and NH2 which are now at distances of 3.58Å and 3.53Å, respectively. Newly formed interactions include several between Glu70 and Asp104, of which two are strong polar interactions (Glu70A(OE1)-Asp104B(OD1) (3.21Å) and Glu70A(OE2)-Asp104B(OD2) (2.69Å).

As in the wt, the involvement of Asp104 in the formation of interfacial interactions via a salt interaction with GSH had been suggested in wt (Johnson *et al*., 2003). In GTX-bound wt, however, Asp104 does not participate in any interfacial interaction. Despite these interfacial differences, there is no significant affect on dimerization, as previously mentioned.

Although there are diverse interactions between wt and R21L, it remains to be seen if they are structurally important.

The superimposition of all structures revealed a quaternary rearrangement regarding helices 5a and 5b at the dimer interface.

The superimposition of R21L and GSH-bound R21L structures has been carried out maintaining monomer A and B alternatively fixed revealing a quaternary movement of helices 5a (residues 118-129) and 5b. Upon ligand binding, helix 5a and 5b from each monomer move symmetrically towards each other (movement of approx. 14 Å), closing the V-shape formed at the dimer interface more evident in the ligand-free structures (Fig. 3.7a). This rearrangement involves only helices at the top but not the bottom of the structure (Fig. 3.7c). Superimposition of ligand-bound wtGST and Y10F also illustrated this movement, confirming that it is not a consequence of the Arg21 mutation. Closure of the V-shape upon binding of GSH is also coupled to movement of the last five C-terminal residues, from an unstructured and delocalized state to a more ordered state, relocating towards the N-terminus (Fig. 3.7a).

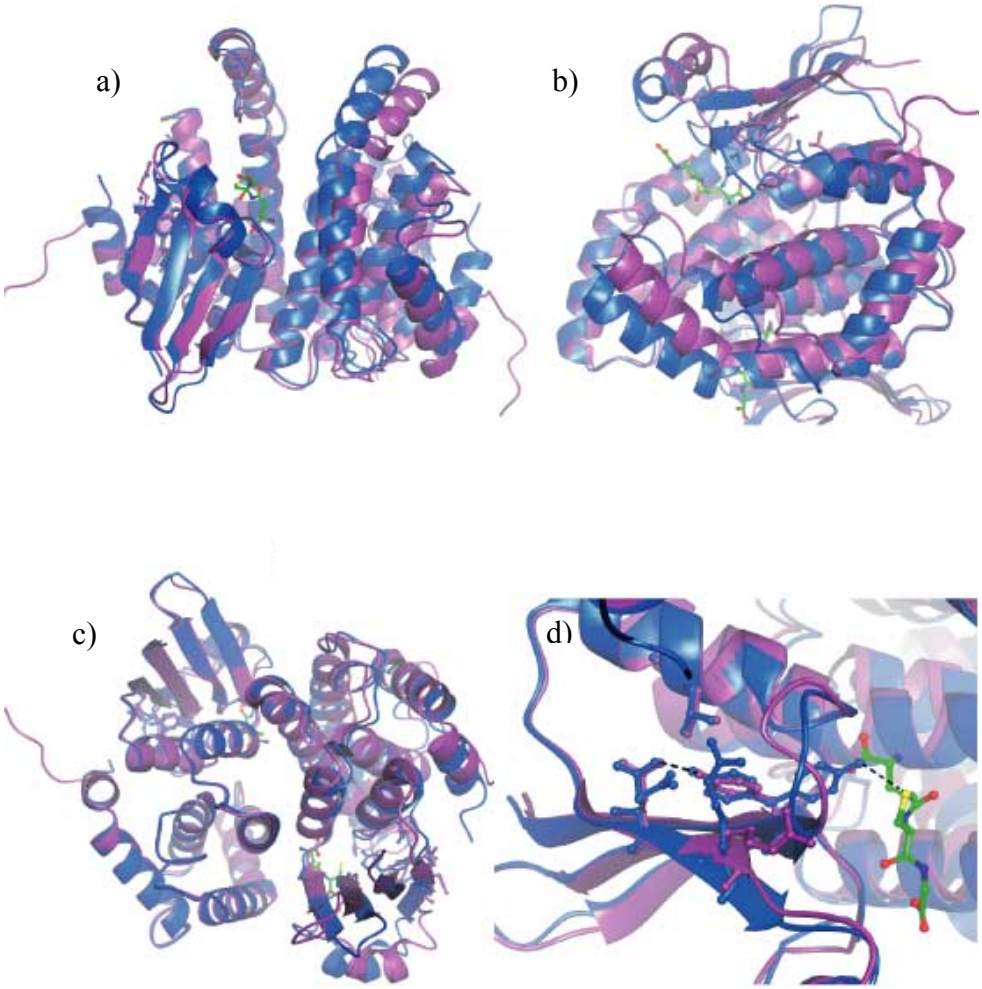


Figure 3.7 *Superimposition of only two structures for clarity is reported: R21L (purple) and GSH-bound R21L (blue). a) overall quaternary structure; b) side view illustrating the movement of helices 5a and 5b; c) bottom view; d) G-site illustrating the H-bond between* $Tyr^{in}10$ *and GSH (green). Figures are carried with Coot program.*

3.6.7 Ligand Binding Experiments

The affinity of all proteins for reduced GSH was measured by tracing the change in tryptophan fluorescence emission with increasing concentration of GSH under equilibrium conditions. Wild type gave a K_d for GSH (22 μM). The K_d of R21Q was calculated to be similar (21 μM) to that of the wild type. In contrast, the K_d of R21L was significantly higher (285 μM), corresponding to a 10-fold lower affinity for GSH than for wild type. Therefore, substitution with Leu influences GSH binding to a greater extent than Gln. Similar experiments were carried out for Y10F, demonstrating that substitution of Tyr10 does not affect GSH-binding, as shown by a K_d value of 21 μM.

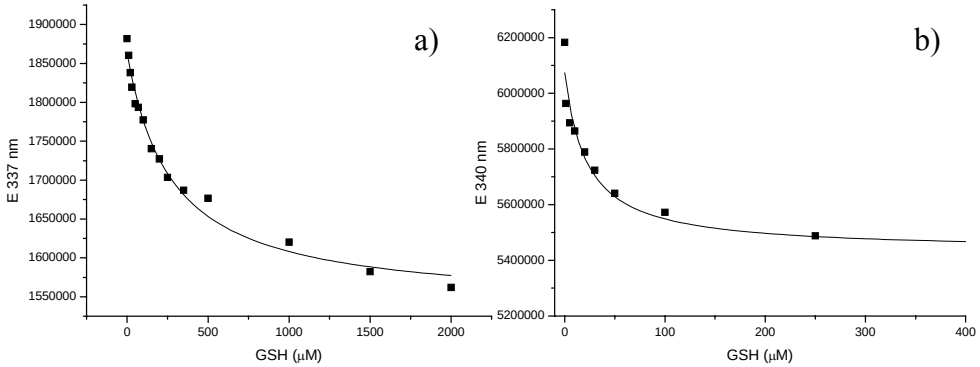


Figure 3.8. *Determination of K_d by fluorimetry for a)R21L and b)R21Q mutants.*

The binding affinity for GTX of wt, R21L and R21Q was also measured. The K_d values were calculated to be 4.7 μM, 11 μM and 16 μM, respectively, indicating that the arginine mutants have binding affinities that are approx. 2-fold lower that that of the wild type enzyme. On comparison of the GSH- and GTX binding affinities (table 3.4), it can be seen that all the proteins bind

GTX with higher affinity. The most obvious difference was seen for R21L which has a 28-fold higher affinity for GTX than for GSH.

Protein	K_d (GSH)	K_d (GTX)
wt	22μM	4.7μM
R21L	285μM	11μM
R21Q	21μM	16μM
Y10F	21μM	n.d.

Table 3.4. *Summary of binding data at equilibrium of wt and mutant Sh28GSTs*

3.6.8 Determination of Catalytic Activity

Steady state kinetic measurements were carried out to determine the effect of the Arg21 mutation on overall catalytic activity. Both arginine mutants had lower k_{cat} values in comparison with wt (48 s^{-1}). k_{cat} were: 4 s^{-1} (R21L) and 0.4 s^{-1} (R21Q).

3.6.9 Stopped-Flow Rapid Mixing Experiments

Rapid mixing experiments by measuring fluorescence changes upon GSH binding were made to ascertain if the velocity of this process is affected by the Arg21 mutation. Reactions of wt and R21L were fitted to a single or double exponential, depending on the trace (materials and methods). The K_d of GSH binding to wt was calculated to be 50 μM. Due to the poor signal for R21L, higher GSH concentrations were required (125μM-500μM) in comparison with wild type experiments (15-50 μM). The K_d for GSH binding to R21L was calculated to be 250 μM and in good agreement with the K_d calculated in fluorimetry experiments (285 μM). The binding affinity of R21L for GSH is approximately 5-fold weaker than for our values for wt.

The ability of wt and R21L to deprotonate GSH was also assessed by following the absorbance increase at 240nm at higher GSH. The K_d for the wild type enzyme was calculated to be approximately 50 μM indicating that we are measuring the same process in both absorbance and fluorescence. There was no absorbance signal for R21L corresponding to its inability to deprotonate GSH. This finding is not unexpected given the role of Arg21 in promoting the deprotonation of $Tyr^{out}10$ by lowering its pK_a.

3.6.10 Measure of the pKa of Tyr10 in wt and mutants

Overall, wtGST contains nine tyrosine residues per monomer and to prove that one of these residues has a lower-than-usual pK_a of 7.2, titrations of wt and mutants were carried out to determine the pK_a value of Tyr10 and to assess the ability of Arg21 to lower the pK_a of the catalytic tyrosine. The curve of absorbance change at 300nm as a function of pH for wtGST shows clearly two titration points at 7.2 and 9.5. In order to assign the first pK to Tyr10, the same experiment has been performed on the inactive mutant Y10F, which indeed showed only one titration point at high pH (Fig. 3.8). Similarly, the low pK_a of Tyr10 was not observed for R21L and R21Q, both displaying high pK_a values of approx. 10 (Fig. 3.8). Similarly, the low pK_a of Tyr10 was not observed for R21L and R21Q, both displaying high pK_a (Fig. 3.9).

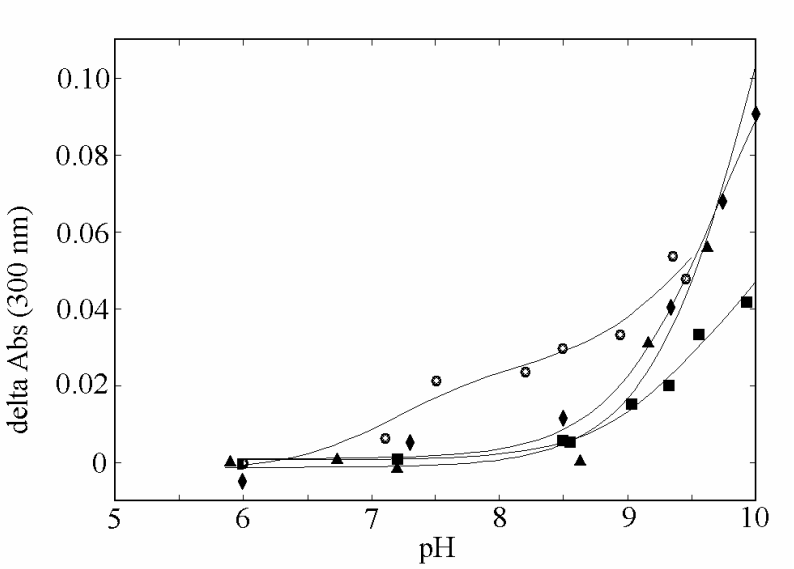


Figure 3.9 *Spectroscopic acid-base titration of wt (open circles), Y10F (squares), R21L (triangles) and R21Q (diamonds) mutants. The difference spectra were calculated using the spectrum recorded at pH = 6.0 as a reference. The differential absorbance readings at 300 nm are plotted as a function of pH according to the least squares fit of the experimental data to one (mutants) or two (wild-type) transitions assigned to the ionizable groups with the pK_as reported in the text.*

3.7 Discussion

In the previously reported structure of *S. haematobium* GST in complex with GSH, the catalytic Tyr10 was found in a double conformation (Johnson *et al.,* 2003). Only the $Tyr^{in}10$ conformer is catalytically competent, being located at short distance from GSH; however the role of the $Tyr^{out}10$ conformer was presumed to be important because the π-cation interaction with the conserved Arg21 residue can lower the pK_a of Tyr10 to approx. 7.2 (Angelucci *et al.,* 2005). Structure-function analysis of two arginine mutants (R21L and R21Q) and one catalytically inactive mutant (Y10F) reported in this paper, revealed important effects of the substitution of Arg21 upon ligand binding and catalytic activity. It was initially proposed that

substitution of Arg21 and hence removal of the π-cation interaction, would stabilise only the $Tyr^{in}10$ conformer; however both structures of R21L and R21Q showed that only the $Tyr^{out}10$ conformer was populated, implying that the π-cation interaction has no effect on the position of Tyr10 in the empty G-site. Once this site is filled with GSH, the Tyr10 is again found in two conformations, at least as far as R21L is concerned. Although the role of Arg21 is clearly not structural, it was demonstrated to be crucial in the GSH activation step of catalysis (Stenberg *et al.*, 1991). As suggested in several reports, this process is mediated by the lowering of the average pK_a of the catalytic residue, thus promoting GSH deprotonation (Armstrong *et al.*, 1997; Xiao *et al.*, 1996; Dietze *et al.*, 1996). Titration experiments, as carried out for Y10F and wtGST (Angelucci *et al.,* 2005), were repeated for our mutants proving that Arg21 is essential to lower the pK_a of Tyr10 by >3 units.

The essential role of Arg21 was further supported in steady state experiments which illustrated that both arginine mutants were significantly less active than the wild type enzyme even if to a different extent. The interaction of Tyr10 with Asp33 may possibly account for the residual activity of R21L with respect to the almost inactive R21Q, since aspartate and glutamate residues have been shown to participate in several catalytic mechanisms as proton shuttles (Nyquist et al., 2003; Frank et al., 2004). Therefore, in R21L but not in R21Q, Asp33 could act as a proton acceptor and partially deprotonate Tyr10. This interpretation is also in agreement with the observed increase in catalytic activity of R21L at higher pH.

It has been previously reported for the human class alpha GST that mutation of R20A (equivalent to Arg21) results in a reduced binding affinity for GSH and a significant loss of catalytic activity (Stenberg *et al.,* 1991). Our GSH

binding experiments showed that R21Q has an affinity similar to that of the wild type enzyme, whereas R21L binds GSH with a 10-fold lower affinity. Although Arg21 does not directly participate in GSH binding, the differences in affinity between the two mutants are likely attributed to the more polar character of glutamine which is stabilised by a H-bond with Asp33(OD2), absent in R21L. Similarly to R21Q, Y10F binds GSH with an affinity similar to the wild type enzyme, in agreement with with the fact that all the point mutations leave the extended G-site practically unaffected.

The binding of the competitive inhibitor GTX to wt and mutants showed similar affinities, implying that GSH and GTX interact differently with the enzyme. In fact, in the crystal structure of GTX-bound wt, the $Tyr^{in}10$ conformer is blocked and the inhibitor occupies both the G-site and the H-site with its hexyl group.

The structures of wtGST crystallized at pH 6.0 with GSH or GTX, show that only the $Tyr^{in}10$ conformation is stabilised, in contrast to the previously reported double conformation (Johnson *et al.*, 2003), likely attributed to diverse crystallisation conditions used for crystal growth. Inhibitor binding experiments illustrated that the binding affinity for the competitive inhibitor GTX was higher than for GSH. This is in agreement with GTX-binding experiments carried out for other GSTs (Ortiz-Salmerón *et al.,* 2003). Rapid mixing experiments were also carried out to calculate the velocity of GSH binding, generating a K_d value approximately two-fold lower that that calculated by equilibrium fluorimetry, but nevertheless in line with it. Thiolate formation was also measured by its absorbance at 240nm, demonstrating that GSH-binding and thiolate formation occur simultaneously. The velocity of GSH binding, measured for R21L, was found

to be 5-fold lower than for the wild type enzyme. In contrast to the wild type enzyme, thiolate formation was not observed once more proving the role of Arg21 in deprotonation of the bound GSH.

Comparison of 3D structures with the wild type enzyme shows, as expected, that the overall fold is maintained. Differences lie in few important features: i) in the absence of ligand Tyr10 is blocked only in one conformation; ii) in the presence of ligand (at least GSH) the double conformation is restored; iii) Arg35 moves concertedly with Tyr10; iv) these movements account for the difference in shape and solvent accessibility of the G-site.

Moreover, by comparing R21L-GSH with R21L, R21Q and wt-GTX we found that GSH or GTX binding i) promotes a previously undescribed quaternary change the interface between the two GST monomers; ii) promotes a series of little structural changes in the active site loop; iii) makes the C-terminal region less flexible and moves it towards the N-terminal region. In detail, the entry of a cofactor or an inhibitor, causes minor rearrangements in the G-site, but closes its aperture to the solvent by structuring the C-terminal loop towards the N-terminus; somehow this information is transferred to the helices 5a and 5b, which undergo a rigid body joint-like movement, ending with the closure of the V-shaped interface, and producing a more compact and stable dimer.

In contrast, in the $Tyr^{out}10$ conformer, populated in the Arg21 mutants, Arg35 moves away from the $Tyr^{out}10$ pocket to permit entry of the catalytic residue, breaking the H-bond with Asp33. In this situation, the active site loop is stabilised by interactions with Asn12. We propose that the movement of Arg35 contributes to minor rearrangements within conserved residues in

the N-terminal domain, thereby regulating the various protein conformations adopted upon ligand binding. The C-terminal tail is flexible, unstructured and lost in the solvent in both unliganded R21L and R21Q.

The structure of R21L/GSH is overall more mobile in the active site loop and C-terminal regions than the ligand bound wild type structures, reflected by less defined density. As already mentioned, Tyr10 is found in a double conformation, as several other active site loop residues, indicating that their movements are connected to Tyr10 movement. Such alterations in the location of the C-terminus with respect to the N-terminal loop region, may account for the observed relationship between the ionisation of the catalytic tyrosine and C-terminus location, seen for the human alpha class GST (Ibarra *et al.,* 2003).

Analysis of the structure of the catalytically inactive mutant, Y10F, shows that the Phe10 occupies a double conformation and the π-cation interaction with Arg21 is maintained. Despite the substitution of Tyr10, GSH still binds although a greater distance between Arg16 and the sulphur atom of GSH is observed. It is known that Arg16 stabilizes the thiolate in alpha class GSTs (Bjornestedt *et al.,* 1995); such stabilization would not be required in Y10F as the thiolate is not formed. The absence of close contacts between Phe10-GSH and Arg16-GSH is in agreement with the low residual activity observed in the steady-state experiments.

Dimerization of GSTs, which is highly specific and class-dependent, is thought to be important in the stabilisation of each individual monomer and for full catalytic activity. The two monomers of the sigma class enzymes interact at the interface via electrophilic interactions (Stevens *et al.*, 2000). An involvement in dimerization of the active site region, which is highly

flexible and undergoes local unfolding, has been suggested (Stevens *et al.*, 1998). To ensure that no structural rearrangements are present at the dimer interface of both R21L and R21Q, the inter-subunit contacts were compared with those present in GTX-bound wtGST. All the variants are dimeric in spite of some rearrangement of several interactions at the dimer interface in R21L. The GST monomers interact via helices 5a and 5b that together to form a V-motif that is open in the ligand-free state and closed in the bound state. Despite reports referring to helix flexibility in the various GST classes, to our knowledge, specific movement of helices 5a and b coupled to GSH binding, has never been reported (Fig. 3.7). Superimposition of our crystallographic structures enabled clear visualisation of the opening and closing of this "V", mediated by the bending of helices 5a and 5b. The extent of closure of the V-motif was found to be correlated to the degree of GSH saturation and not influenced by the Arg21 or Tyr10 mutations. In addition to closure of the V-motif, there is an overall alteration in dimer packing. Upon ligand binding the monomers close around the ligand, resulting in a more compact dimer. These rearrangements are concerted with the movement of the C-terminal domains that become more structured and relocate next to the active site in the N-terminal domain. Comparing the ligand-bound wild type structures with the arginine mutants it appears that our results are in agreement with findings reporting the involvement of α9-helix of the C-terminal of GST A1-1 (Nilsson *et al.*, 2002). Together, our findings raise the question of whether the quaternary conformational change occurring upon GSH binding can significantly account for positive cooperativity (Wongsantichon and Ketterman, 2005).

To summarize, our structural and functional studies carried out on three Sh28GST mutants in comparison with the wild type protein provide a better insight into the catalytic mechanism of this sigma class GST. It is clear that Arg21 has a fundamental role in proton transfer and GSH activation, but it does not determine Tyr10 position in the G-site. Furthermore, the movement of Tyr10 in and out of the G-site is accompanied to the concerted movement of Arg35 in and out of the $Tyr^{out}10$ pocket, alterations in the structure and localisation of the C-terminus and a never-previously-reported quaternary movement at the dimer interface. Given the conservation of Arg21, Asp33, Arg35 and Tyr10, we propose that the concerted movement of such residues during ligand binding occurring in *S. haematobium* GST may be common to other GSTs of the same class as a general mechanism.

CHAPTER 4

CYCLOPHILIN-LIKE PROTEINS

4.1 Cyclosporin A in Schistosomiasis

The curative and preventive effect of Cyclosporin A (CsA) against schistosomiasis has been proposed by Thomson and Chappell back in 1988. CsA is an immunosuppressive agent of clinical relevance and its anti-parasitic effect in schistosomiasis has been evaluated in mice infected with *S. mansoni* at different stages (Cioli *et al.,* 1995): i) topical application of CsA to the skin site of cercarial penetration, prior to infection, resulted in no reduction in worm burden (Munro *et al.*, 1990); ii) infection of mice with cercariae which had been exposed to CsA in vitro, resulted in only a slight reduction of the worm burden for the highest concentration of CsA tested (100 micrograms/ml) (Nillson *et al.*, 1985); iii) administration of CsA at the time of infection or during the schistosomulum stage resulted in failure of the larvae to develop into adult worms; iv) administration of CsA during the establishment of the adult worm stage resulted in a reduction of the worm burden, as compared to non-treated mice (Munro *et al.*, 1990), and preferentially targeted female worms (Millership *et al.*, 1996); moreover, the established worm pairs, seemed to be made sterile since no eggs were laid in the liver.

CsA has two distinct modes of action against *S. mansoni* infection: as an anti-helminthic drug and as immunomodulator. The combined effects of drug action and cellular cytotoxicity, presumably, account for the very significant levels of worm killing achieved by CsA treatment of the host. Some changes in worm morphology have been examined by light and electron microscopy for drug-induced and it has been reported a damage by massive bolus formation and subsequent herniation of the gut. This event was attributed to the abnormal accumulation of crystals. In some drug-treated

worms, the caecal epithelium had exploded, thereby releasing luminal contents throughout the worm body. In addition, herniations of the gut were seen protruding through the tegument causing surface deformation and disruption of tegumental and parenchymal tissues. The structural integrity of the worm was ultimately compromised allowing access to cytotoxic effector cells of the host.

Finally, CsA is also known to act on T-lymphocytes via the combination with the cyclophilin protein, which has been structurally identified in a number of parasites (Bell *et al.*, 1996), and the identification of the targets of this drug in parasites may lead the development of novel chemotherapeutic agents.

4.2 New target for a schistosomiasis vaccine

In 1984, the first cyclophilin (CypA) was co-discovered by Fischer and co-workers, who developed a protease-coupled assay to demonstrate that it catalyses proline-limited protein folding reactions, and by Handschumacher and co-workers, who employed affinity chromatography to identify a receptor for CsA. In 1989, it became clear that the proteins identified by the two groups were identical (Fischer *et al.*, Takahashi *et al.*,) and it was immediately speculated that the function of cyclophilins in immunosuppression might be linked to prolyl isomerase activity. Subsequent works, however, demonstrate that inhibition of Cyp activity is not the cause of parasite death (Khattab *et al.*, 1998).

The immunosuppressive effect of CsA is believed to be mediated through the interaction of CypA/CsA complex with the phosphatase calcineurin, inhibiting its calcium- and calmodulin-dependent activities

required for the activation of genes encoding interleukin-2 and other cytokines (Walsh *et al.*, 1992).

The peptidyl-prolyl *cis-trans* isomerase activity (PPIase activity) is fundamental in the protein folding process, because it accelerates *cis-trans* amide isomerisation (Fischer & Schmid, 1990). Thermal isomerisation is much slower than protein folding and the occurrence of proline amides may complicate protein folding when a particular prolyl amide must isomerise before a protein can reach its native folded structure. One of the major unresolved issues concerning this class of enzymes is their mechanism of catalysis. Several mechanisms have been proposed (Kofron *et al.*, 1991): 1) catalysis by formation of a covalent tetrahedral carbon of the prolyl imidic bond for the transition state of reaction, 2) catalysis by distortion, in which PPI binds and stabilises a transition state characterized by partial rotation about the C-N amide bond, 3) protonation of the amide nitrogen that lowers the barrier for the *cis-trans* rotation, 4) catalysis by desolvation, based on the observation that the rate of the *cis-trans* isomerisation was significantly accelerated in non-polar solvents and 5) a solvent-assisted mechanism that assumes two steps in catalysis: desolvation by binding to the hydrophobic pocket and stabilisation of the intermediate by a solvent molecule.

As the biological importance of PPIases is gradually revealed, an increasing number of inhibitors have been synthesized as potential drugs for various diseases. The largest family of cyclophilin inhibitors are related to the cyclic peptide cyclosporin A and some non-peptide compounds including ethyl-1-piperidine glyoxylate (ETPIPG) (Kallen *et al.,* 1998). Piperidine ligands are described as the first example of non-peptide ligand structures for the cyclophilin family of proteins. These structures may be of relevance in

the field of drug design, as they suggest starting points for the design of larger tighter-binding analogues (Kontopidis *et al.*, 2004).

4.3 Material and Methods

4.3.1 Cloning of S. mansoni cyclophilin

The *S. mansoni* gene was amplified by PCR using PFu DNA polymerase (Stratagene) from genomic DNA of *S. mansoni* (kind gift from Dr. Raymond Pierce, Lille, France) as the template for the reaction, and sub-cloned into the bacterial expression vector pGEX-4T-1 (Novagen). The primers used in the reaction were as follows: forward 5'-C**GG GAT CC**A TGC GAA CCA AAA AAC AAA AAC GAA ATC TTC CTC GGG-3'; and reverse, 5'-CCG **CTC GAG** CTA AAT TAA CTC TCC GCA TCG AGA AAT AAT AAC GGG-3'. *BamHI* and *Xho*I recognition sites were incorporated into the forward and reverse primers, respectively, to facilitate directional cloning (in bold face). Successful amplification was confirmed by agarose gel electrophoresis. PCR amplification was carried out using a PTC-100™ programable thermocycler at annealing temperature of 55°C. The fidelity of the amplification reactions was confirmed by DNA sequencing at the Bio Molecular Research Sequencing Service (University of Padova, Italy).

4.3.2 Cyclophilin from Haemonchus contortus

The plasmid encoding for a cyclophilin from the nematode *Haemonchus contortus* (HcCyp) has been given to us by Prof. Cioli and co-workers (Institute of Cell Biology, CNR, Monterotondo). It displays a high sequence homology with human cyclophilin as reported in Fig. 4.1. It

contains a RNA recognition domain (RRM) at the N-terminus of the PPIase domain, therefore, a putative role in transcription/translation regulation may be inferred (Valle *et al.*, 2005).

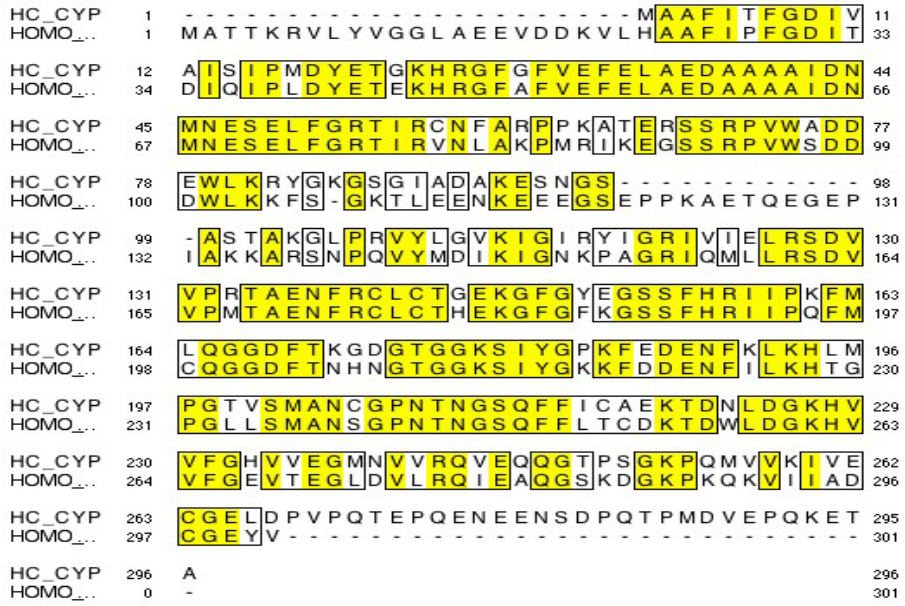


Figure 4.1: *Alignment of HcCyp and human cyclophilin: homology and identity sequence are indicated by rectangular box and in yellow, respectively.*

4.3.3 Protein Expression and Purification

The HcCyp gene was cloned into the pDEST-17 expression plasmid as previously described (Valle *et al.*, 2005), incorporating a non-cleavable histidine tag at the N-terminus of the recombinant protein.

HcCyp has been expressed in *E. coli* cells grown at 37°C in Luria broth using 34 μg/ml of chloramphenicol and 100 μg/ml of ampicillin as antibiotics. When an A^{600} of 0.4 was reached protein expression was induced by addition of 0,3 mM isopropyl-β-D-thiogalactopyranoside (Sigma-Aldrich) incubating for a further 5 hours. Bacterial cells were harvested by centrifugation at

13000 rpm for 30 min. The cells were resuspended in 20 mM HEPES pH 7.4, 200 mM NaCl and 5 mM $MgCl_2$, 10 mM 4-(2-aminoethyl)benzenesulfonyl fluoride (AEBSF) and 2mM β-mercaptoethanol (β-ME). Cell lysis was achieved by sonication in bursts of 4s at 9s intervals for 5 min. The sonicated extract was centrifuged at 13000 rpm for 30 min. The extraction of HcCyp, expressed in inclusion bodies was carried out following a standard protocol. The pellet has been homogenized with 1% sodium deoxycholate in 20 mM Tris pH 8.0, 0.2 M NaCl and centrifuged for 20 min at 13000 rpm. The pellet has been washed three times, by 20 minutes centrifugation, with 0.25% sodium deoxycholate in 10 mM Tris pH 8.8, 1mM EGTA, 1mM β-ME, then once with 20 mM HEPES pH 7.4. Finally, the pellet was resuspended in 6M urea, 20 mM HEPES pH 7.4 and 2mM β-ME stirring overnight at 4°C. The solution was centrifuged for 40 min at 13000 rpm and protein of high purity was obtained from the supernatant. HcCyp was refolded by dialysis vs. 20 mM sodium phosphate pH 7.4 and 2 mM β-ME and concentrated through Amicon-plus concentration tubes using a membrane cut off of 10 kDa.

4.3.4 Biochemical characterisation of HcCyp

Protein concentration was routinely determined spectroscopically, measuring the A_{280} using a HP 845X UV-visible spectrophotometer (Hewlett Packard) using an extinction coefficient of 0.58 mg/ml.

The far-circular dichroism spectra has been measured using a Jasco circular dichroism spectrophotometer (Hewlett Packard). Experiments were carried out at 10°C over a range of 190-300 nm using a 0.1 mg/ml concentration of HcCyp in 20 mM sodium phosphate buffer pH 7.4.

Light scattering experiment was carried out equilibrating HcCyp (2mg/ml) in a 10mM sodium phosphate pH 7.4. HcCyp was previously filtered through 0.2 μm membrane. myoglobin is used as standard protein.

HcCyp was applied on a gel filtration column G3000SW equilibrated in 20 mM sodium posphate pH 7.2 and analysed through HPLC. A calibration curve was plotted using sample at different molecular weight. The standard proteins used were: albumin (67 kDa), ribonuclease A (13.7 kDa) and ovalbumin (45 kDa). Chromatograms were monitored recording the absorbance at 280 nm with a flux of 0.8 ml/min. Each sample has been filtered using a 0.20 μm filter unit before the injection.

Chromatograms in presence of different NaCl and EDTA concentrations have been monitored in order to optimise conditions to alter the polymerisation degree of the protein. Comparison of retention times of the same protein under different solution conditions revealed no salt influence in the aggregation state.

4.3.5 Enzymatic assay and inhibition studies

The PPiase activity is usually determined spectroscopically using a Jasco UV-visible spectrophotometer (Hewlett-Packard) by the α-chymotrypsin-coupled enzymatic assay (Kofron *et al.,* 1991). α-Chymotripsin selectively hydrolyses the C-terminal peptide bond in N-Suc-Ala-Ala-Pro-Phe-4-nitroanilide only when the proline is in the *trans* conformation. This hydrolysis releases the chromogenic product 4-nitroanilide, the accumulation of which is monitored by measuring the absorbance at 390 nm as function of time ($\Delta\varepsilon$=13400 $M^{-1}\cdot cm^{-1}$).

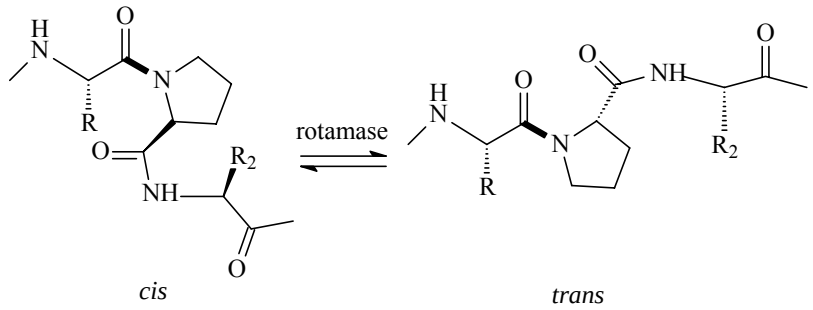


Figure 4.2. *Trans-cis isomerisation of the peptide bond in prolines.*

In aqueous solution, N-succinyl-Ala-Ala-Pro-Phe-4-nitroanilide is in equilibrium between the *cis* and *trans* forms. In order to increase the initial population of the *cis* conformer from 10-12% to 40-60%, the substrate is dissolved in 10% trifluoroethanol (TFE) and 0.45M LiCl (Kofron *et al.*, 1991). In the presence of a PPIase the rate of cis/trans isomerisation of peptides is accelerated which in turn shortens the life time of the blocked peptide.

Experiments were carried out at 15°C measuring the absorbance at 390 nm over 600 seconds equilibrating the substrate in a cuvette containing 20 mM sodium phosphate pH 7.4. HcCyp was freshly prepared before the experiment and added into the cuvette 30 seconds after the beginning of the measure and after further 30 seconds, 50 μM chymotrypsin was added to initiate the reaction. Under these conditions, the burst phase due to the cleavage of the excess isomer is completed during the mixing time and the remaining absorbance change is due to the cis-trans isomerisation.

Time courses at over a range of substrate concentration (20μM, 50μM, 100μM, 300μM, 600μM, 1000μM) were recorded and data were fitted using the Origin 7 Server Software.

The ability of CsA to inhibit the enzymatic activity of HcCyp was also measured in similar steady state conditions previously reported, at constant chymotrypsin and cyclophilin concentrations (50μM and 5nM, respectively), over a range of 4-nitroanilide concentrations (50μM, 75μM, 100μM, 600μM, 1000μM), for three different CsA concentrations (10 nM, 20 nM, 60 nM). CsA has been prepared in a 0.1 M stock solution in 0.5 mM sodium phosphate, pH 7.4 in 40% ethanol because of its insolubility in water solution. Assays were carried out as described above but HcCyp has been preincubated with CsA, 30 min before each experiment.

4.3.6 Crystallisation tests

Crystallisation trials for HcCyp were carried out by the vapour diffusion method using both hanging and sitting drop. Screening experiments were performed with commercial screen kits from Hampton Research, Wizard and Stura. The protein has been dialysed against 20 mM Tris pH 7.5 and 5mM ß-ME. A very large number of crystallisation trials were carried out at different temperatures (4°C and 21°C) and using different protein concentrations. In order to avoid bacterial contamination, 0.3% Na_3N was added to the protein stock solution.

Addition of 8 mM EDTA has been tested during crystallisation trials, in order to avoid possible aggregation due to the presence of His-tag.

Furthermore, diverse concentrations of dimethylsulfoxide (DMSO) has been added both in the protein stock and in the reservoir solution, 0.3% and 5% respectively, as reported in the literature (Mikol *et al.*, 1993).

4.4 Results and Discussion

4.4.1 Cloning, expression and purification of SmCyp and HcCyp

The cloning of *S. mansoni* cyclophilin in pGEX4-T1 plasmids was carried out but expression tests revealed that this protein is present exclusively in inclusion bodies. New attempts for cloning in different plasmid are in progress.

Successful expression of HcCyp in inclusion bodies, carried out in bacterial cells according to materials and methods, was confirmed by SDS-PAGE under reducing conditions. A very pure band, corresponding to the predicted M_r value of the HcCyp, was observed. Further purification through nichel affinity columns was not necessary as judged by SDS-PAGE analysis. Each purification yields approx. 60 mg/500 ml culture of pure protein.

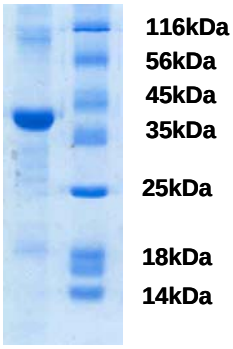


Figure 4.3: *SDS-PAGE of HcCyp after cell lysis.*

4.4.2 Biochemical characterisation

Successful refolding of the denatured protein was confirmed by experiments carried out using a circular dichroism spectrophotometer. The far-UV-CD spectra confirmed the prediction of a protein rich in β-sheets content.

Circular dichroism spectra have also been measured in 20 mM Tris pH 7.5 and 5mM ß-ME, in order to check whether the protein fold was maintained in the crystallisation buffer.

Light scattering experiments result in a gaussian curve centered at 31 kDa confirming the monodispersive phase of a monomeric protein.

Gel filtration analysis carried out by HPLC, interpolated with a calibration curve, gives a higher than expected molecular weight (125 kDa). This result may be due to either to the presence of aspecific interactions with the matrix of the column, or to aggregation of HcCyp monomers.

4.4.3 Prolyl isomerase assay

The PPIase activity of HcCyp was measured according to the coupled enzymatic assay as reported in material and methods. The initial velocity of each time course (panel in Fig. 5) were linearly interpolated and all the slope values were plotted against substrate concentrations in order to measure Michaelis constant (K_M) (Fig. 4.4).

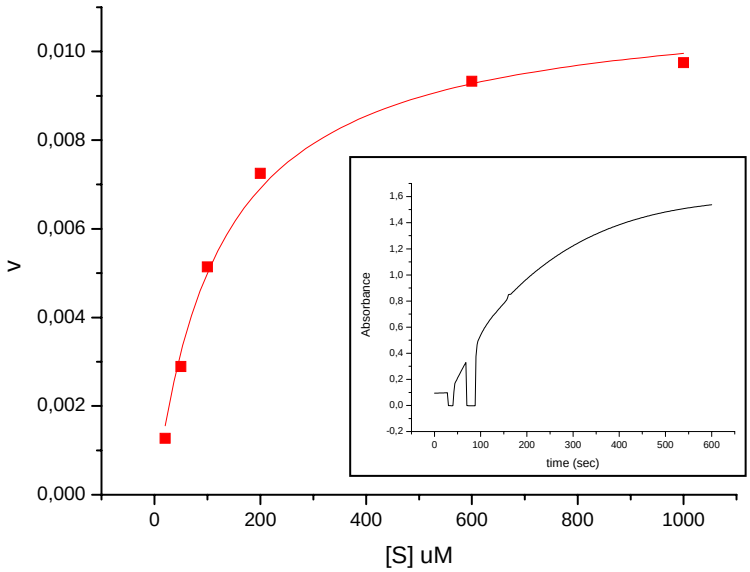


Figure 4.4: *Michaelis Menten plot for a 10 nM HcCyp. In the panel, a time course of the protease coupled prolyl isomerase assay is illustrated in the case of 100 μM substrate concentration.*

K_M and k_{cat} values are calculated at different Cyp concentrations and reported in table 4.1.

HcCyp (nM)	$k_{cat}(s^{-1})$	K_M (μM)
2.5	2.7	81.9
5	1.9	67.2
10	1.1	123.7
20	0.7	257.0

Table 4.1: K_M *and* k_{cat} *values are reported for different concentration of HcCyp.*

Enzymatic activity seems to slightly decrease, increasing cyclophilin concentration used in assays; this is likely due to eventual protein aggregation in solution (see also above, gel filtration experiments). As a

measure of cyclophilin efficiency we have compared k_{cat}/K_M with those from other cyclophilins. HcCyp (k_{cat}/K_M 2.7-8.8 10^4 $M^{-1}s^{-1}$) resulted to be less efficient than human Cyp and plasmodium falciparum Cyp (k_{cat}/K_M 2.2 10^8 M^{-1} s^{-1}) (Berriman *et al.*, 1998, Liu *et al.*, 1990, Bergsma *et al.*, 1991). As a consequence of protein expression in inclusion bodies, we were probably unable to prepare completely active and renatured cyclophilin.

4.4.4 Inhibition assay

Inhibition studies were carried out according to material and methods and K_M and k_{cat} at different CsA concentrations have been calculated, as reported in Table 4.1.

Plotting 1/v versus 1/[S] in the Lineweaver-Burk plot (Fig. 4.5a), it is quite evident that HcCyp inhibition is unaffected by varying inhibitor concentration and that there is a significant inhibition of rotamase activity by cyclosporin A, despite the fact that k_{cat}/K_M values are similar to those obtained in the absence of inhibitor (tables 4.1 and 4.2).

CsA seems to be a very weak competitive inhibitor, even if its role in the catalysis is not completely understood and some example of cyclophilins with the retention of their catalytic efficiency without CsA inhibition has been reported (Liu & Walsh, 1990).

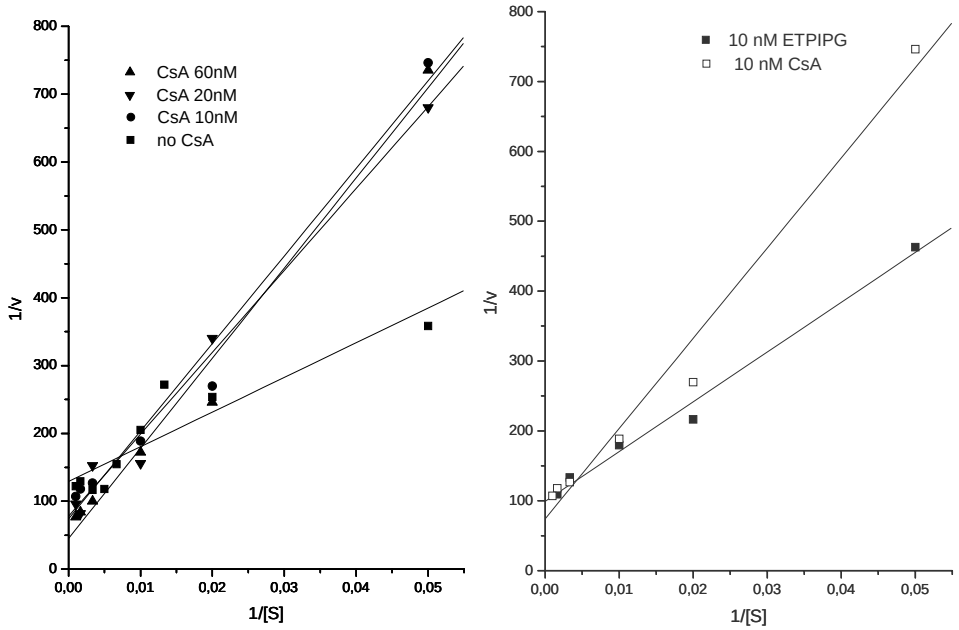


Figure 4.5: *Lineweaver-Burk plots for the inhibition analysis a) at different concentration of CsA b) with ETPIPG. In ascisses are reported 1/[S] where S is substrate concentration.*

Inhibition studies have been carried out also in the presence of a non-peptide inhibitor, ETPIPG (Fig. 4.6 b), and similar calculations have been made resulting k_{cat}/ K_M value similar to those of CsA (table 4.2).

CsA (nM)	K_M (μM)	$k_{cat}(s^{-1})$
10	93.3	2
20	142.8	2.5
60	151.6	2.8
ETPTIPG (nM)		
10	66.8	1.9

Table 4.2: *Michaelis constants and kcat values for inhibition assays*

To summarise, we have studied a potentially pathologically important protein involved in the development of schistosomes and analogue parasites: a cyclophilin from the nematode *H. contortus* that contains an RNA recognition motif. Cyclophilins are typically, expressed in the soluble fraction of bacterial systems, however HcCyp was present only in inclusion bodies. Refolding of proteins isolated from inclusion bodies is critical and diversely folded intermediates may occur; however, we successfully refolded HcCyp which permitted its successful biochemical characterisation. Its peptidyl-prolyl *cis-trans* isomerase activity was measured, revealing that it is not as efficient as human cyclophilin. Since the inhibition of this enzyme may be interesting from a pharmacological point of view, we carried out enzymatic assays in the presence of natural and non–peptide inhibitors. Examination of the primary structure of human cyclophilins revealed an unusual high content of the aromatic residue phenylalanine, many of which have been identified by NMR as participants in the formation of the hydrophobic binding site for CsA (Dalgarno *et al.*, 1986). HcCyp was slightly inhibited by cyclosporin A as revealed by competition experiments. However, it is not completely clear if cyclosporin acts as a weak competitive inhibitor or not. The possible explanation for the low catalytic efficiency and inhibition of cyclosporin could be assigned to the conditions in which refolding was carried out.

In order to express this protein in the soluble fraction, attempts were made to clone cyclophilin from *S. mansoni,* using diverse plasmids. Expression in soluble form, in fact, will hopefully help the crystallisation and the achievement of a crystal structure is fundamental to better understand the

catalytic mechanism and to provide a good template for the design of novel ligands and more efficient inhibitors.

CHAPTER 5

CONCLUSIONS

In this project, two proteins potentially relevant for the survival and the pathogenicity of schistosome parasites were studied: cyclophilin-like protein, a peptidyl-prolyl isomerase enzyme involved in the maintenance of correct folding and a key enzyme of the detoxification mechanisms, the glutathione S-transferase from *S. haematobium* (Sh28GST).

In the first case, I have studied a cyclophilin-like protein from *H. contortus* (HcCyp), a parasite of the same family which affects cows. HcCyp was expressed as a recombinant His-tagged protein in *E. coli*. Despite different growth conditions, it was exclusively found in the inclusion body fraction, however it was successfully solubilised and refolded. Its catalytic rotamase activity has been characterised with respect to its peptide substrate and to a couple of inhibitors: cyclosporin A, which has been used as chemotherapeutic agent in human schistosome infections, and ethyl-1-piperidine glyoxylate, a non-peptide inhibitor known to act on the human homologue of Cyp, which by the way had never been tested on a protein of a parasite species before.

Despite numerous crystallisation trials, we did not manage to obtain any promising conditions. Future work will focus on the expression of *S. mansoni* cyclophilin in soluble form to avoid the critical refolding step, which will hopefully help subsequent structural studies. Moreover it will include the investigation of its RNA binding ability and possibly the assessment of the efficiency of new inhibitors.

As for the detoxification pathway we focused on Sh28GST, which is involved both in the bio-transformation of xenobiotics into more soluble and readily excretable compounds and in the isomerisation of prostaglandin

PGH_2 to PGD_2. Therefore GST is critical in two diverse pathways, essential for defence against the host response.

In this thesis, the high resolution structures of three Sh28GST mutants (R21L, R21Q and Y10F) and the structure of wild type in complex with its competitive inhibitor, S-hexylglutathione are presented, together with the characterisation of their ligand affinities and transferase activities.

The crystal structure of GSH-bound wild type GST (Johnson *et al.*, 2003) revealed a characteristic double conformation of the catalytic residue, namely Tyr10, adopting two alternative positions: an activating conformer ($Tyr^{in}10$), where the phenoxyl group of Tyr10 projects towards the G-site forming a H-bond with the thiolate of GSH and a novel non-activating conformer ($Tyr^{out}10$) where the phenoxyl ring is positioned out of the G-site and is stabilised by a π-cation interaction with the guanidinium group of Arg21. The catalytic role of Tyr10 has been confirmed through an enzymatic assay of the Y10F mutant, which is completely inactive. Arg21 is a strictly conserved residue in GSTs containing a catalytic tyrosine and the functional characterisation of Arg21 mutants demonstrates its fundamental role in catalysis. R21L and R21Q have very low catalytic activities even though their ability to bind glutathione is conserved. The π–cation interaction is responsible for lowering the tyrosine pK_a from approximately 10 to 7.4 and promotes $Tyr^{out}10$ deprotonation, rendering $Tyr^{in}10$ available for subsequent GSH activation (Angelucci *et al.*, 2005).

Structural characterisation of both mutants revealed concerted movements of residues in close proximity to the G-site. These movements are correlated to the Tyr10 position and modify the solvent accessibility of Arg21 pocket. The superimposition of the overall structures of our proteins, revealed novel

tertiary movements involving helices at the dimer interface and the C-terminal regions, resulting in a more compact quaternary structure, upon ligand binding. These quaternary rearrangements were not previously described, although the flexibility of several helices has been extensively reported in the literature (Ibarra *et al.,* 2001) and seems to be correlated to ligand binding more than due to the mutations.

In summary, in the near future, it will be possible to start a compound search for molecules able to bind the G- and/or the H-site, by indirectly influencing the activity of the catalytic residue. The comparison between the parasite enzyme and its human counterpart will help to restrict the development of putative compounds that bind specifically to Sh28GST, therefore reducing the presence of toxic side-effects in the host.

CHAPTER 6

REFERENCES

-Al-Sherbiny M., Osman A., Barakat R., El Morshedy H., Bergquist R., Olds R. (2003) "In vitro cellular and humoral responses to Schistosoma mansoni vaccine candidate antigens" *Acta Trop.*, 88, 117-30.

-Andrews P., Thomas H., Pohlke R., Seubert J., (1983) "Praziquantel" *Med Res Rev.*, **3**, 147-200

-Angeli V., Faveeuw C., Roye O., Fontaine J., Teissier E., Capron A., Wolowczuk I., Capron M., 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., Baiocco P., Brunori M., Gourlay L. J., Morea V., Bellelli A., (2005) "Insights into the catalytic mechanism of glutathione S-transferase: the lesson from Schistosoma haematobium" *Structure* **13**, 1241-6.

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

-Argaet VP, Mitchell GF., (1992) "Cyclophilin of *Schistosoma japonicum*" *J. Parasitol.,* **78**, 660–4

-Armstrong R.N. (1997) "Structure, catalytic mechanism, and evolution of the glutathione s-transferase" *Chem. Res. Toxicol.*, **10**, 2-18

-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

-Balloul J.M., Grzych J.M., Pierce R.J., 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., Pierce R.J., Grzych J.M., Capron A., (1985) "In vitro synthesis of a 28 kilodalton antigen present on the surface of the schistosomulum of Schistosoma mansoni" *Mol Biochem Parasitol.,* **17**, 105-14.

-Bell A., Roberts H.C., Chappell L.H., (1996) "The Antiparasite Effects of Cyclosporin A: Possible Drug Targets and Clinical Applications" *Gen. Pharmacol.*, **27**, 963-71

-Bergsma D.J., Eder C., Gross M., Kersten H., Sylvester D., Appelbaum E., Cusimano D., Livi G.P., McLaughlin M.M., Kasyan K. (1991) *J. Biol. Chem.* **266**, 23204–23214

-Berriman M., Fairlamb AH. (1998) "Detailed characterization of a cyclophilin from the human malaria parasite: Plasmodium falciparum" *Biochem J.*, **334**, 437-45

-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

-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

-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. (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

-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

-Capron A., Riveau G., Capron M., Trottein F., (2005) "Schistosomes: the road from host–parasite interactions to vaccines in clinical trials" *Trends in Parasitology*, **21**, 3, 143-149

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

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

-Cioli D., Pica-Mattocia L., Archer S., (1995) "Antischistosomal drugs: past, present ... and future?" *Pharmacol Ther.*, **68**, 35-85

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

-Dalgarno D.C., Harding M.W., Lazarides A., Handschumacher R.E., Armitage I.M. (1986) *Biochemistry*, **25**, 6778-6784

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

-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

-Frank R.A., Titman C.M., Pratap J.V., Luisi B.F., Perham R.N. (2004) "A molecular switch and proton wire synchronize the active sites in thiamine enzymes" *Science,* **306**, 872-6

-Fischer G., Bang H., Mech C. (1984) "Determination of enzymatic catalysis for the cis-trans-isomerization of peptide binding in proline-containing peptides" *Biomed. Biochim. Acta,* **43**, 1101-1111

-Fischer G., Wittmann-Liebold B., Lang K., Kiefhaber T., Schmid F., (1989) "Cyclophilin and peptidyl-prolyl cis-trans isomerase are probably identical proteins" *Nature*, **337**, 476-478

-Fischer G., Schmid F.X., (1990) "The mechanism of protein folding. Implications of in vitro refolding models for de novo protein folding and translocation in the cell" *Biochemistry*, **29**, 2205-2212

-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.

-Handschumacher R.E., Harding M.W, Rice J., Drugge R.J., Speicher D.W. (1984) "Cyclophilin: a specific cytosolic binding protein for cyclosporin A" *Science,* **226**, 544-547

-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., 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.

-Hayes J.D., Flanagan J.U., Jowsey I.R. (2004) "Glutathione Transferase" *Annual Review of Pharmacology and Toxicology*, **45**, 51-88

-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 PGD2 synthase and the host D prostanoid receptor 1 in schistosome immune response" *Eur. J. Immunol.*, **33**, 2764-72

-Ibarra C.A., Chowdhury P., Petrich J.W., Atkins W.M. (2003) "The anomalous pKa of Tyr-9 in glutathione S-transferase A1-1 catalyzes product release" *J Biol Chem.* **278**,19257-65.

-Ibarra, C., Nieslanik, B.S., and Atkins, W.M. (2001) "Contribution of aromatic-aromatic interactions to the anomalous pK(a) of tyrosine-9 and the C-terminal dynamics of glutathione S-transferase A1-1" *Biochemistry,* **40**, 10614-24

-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.

-Kallen J., Mikol V., Taylor P., Walkinshaw M.D., (1998) "X-ray structures and analysis of 11 cyclosporin derivatives complexed with cyclophilin A" *J. Mol. Biol.*. **283**, 435-49

-Khattab A., Pica-Mattoccia L., Klinkert M.Q., Wenger R., Cioli D. (1998) "Cyclosporins: Lack of correlation between antischistosomal properties and inhibition of cyclophilin isomerase activity" *Exp. Parasitol.*, **90**, 103-109

-Kiang D., El Ghazalie N.E., Medhat A.M., Abdel-Fattah M., Karim A.M., LoVerde PT. (1995) "Identification and characterization of Schistosoma mansoni p17.7, a cyclophilin" *Molecular and Biochemical Parasitology*, **76**, 73-82

-Klinkert M.Q., Bugli F., Engels B., (1995) "Characterization of a *Schistosoma mansoni* cDNA encoding a B-like cyclophilin and its expression in *Escherichia coli*." *Mol Biochem Parasitol.*, **75**, 99–111

-Klinkert M.Q, Bugli F., Cruz J., Engels B., Cioli D., (1996) "Sequence conservation of schistosome cyclophilins" *Mol Biochem Parasitol*. **81**, 239–42

-Klinkert M.Q., Bugli F. Engels B., Carrasquillo E., Valle C., (1995) "Characterization of a Schistosoma mansoni cDNA encoding a B-like cyclophilin and its expression in Escherichia coli" *Molecular and Biochemical Parasitology*, **75**, 99-111

-Kofron J.L., Kuzmic P., Kishore V., Colon-Bonilla E., Rich DH., (1991) "Determination of kinetic constants for peptidyl prolyl cis-trans isomerases by an improved spectrophotometric assay" *Biochemistry*, **30**, 6127-34

-Kontopidis G., Taylor P., Walkinshaw M.W.,(2004) "Enzymatic and structural characterization of non-peptide ligand-cyclophilin complexes" *Acta Cryst*. D**60**, 479-485

-Liu J., Albers M.W., Chen C., Schreiber S.L., Walsh C.T. (1990) *Proc. Natl. Acad. Sci. U.S.A.*, **87**, 2304–2308

-Liu J. & Walsh C.T. (1990) "Peptidyl-prolyl cis-trans-isomerase from Escherichia coli: A periplasmic homolog of cyclophili that is not inhibited by cyclosporin A" *Proc. Natl. Acad. Sci. USA*, **87**, 4028-4032

-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

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

-McTigue M.A., Williams D.R., Tainer J.A., (1995) "Crystal structures of a schistosomal drug and vaccine target: glutathione S-transferase from Schistosoma japonica and its complex with the leading antischistosomal drug praziquantel" *J Mol Biol.*, **246**, 21-27.

-Mikol V., Kallen J., Pflugl G., Walkinshaw M.D. (1993) "X-ray structure of a monomeric cyclophilin A-cyclosporin A crystal complex at2.1 A resolution" *J Mol Biol.* **234**, 1119-30

-Millership J.J., Chappell L.H., Fallon PG. (1996) "Synergistic action of cyclosporin A and polyspecific rabbit anti-sera against murine Schistosoma mansoni" *Parasite Immunol*., **18**, 71-7

-Moser D., Tendler M.,Griffiths G., 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

-Munro G.H., McLaren D.J., (1990) "Schistosoma mansoni: morphology and ultrastructure of adult worms recovered from cyclosporin A-treated mice" *Parasitology*, **1**, 19-28

-Murshudov G., Vagin A., Dodson E., (1996) "Application of Maximum Likelihood Refinement in the Refinement of Protein structures", Proceedings of Daresbury Study Weekend. (REFMAC 5)

-Navaza J. (1994) "AMoRe: An automated package for molecular replacement." *Acta Cryst*. A**50**, 157-163

-Nilsson L.O., Edalat M., Pettersson P.L. Mannervik B. (2002) "Aromatic residues in the C-terminal region of glutathione transferase A1-1 influence the rate-determining steps in the catalytic mechanism" *Biochimica et Biophysica Acta*, **1598**, 199-205

-Nilsson L.A., Lindblad R., Olling S., Ouchterlony O., (1985) "The effect of cyclosporin A on the course of murine infection by Schistosoma mansoni" *Parasite Immunol.*, **7**, 19-27

-Nyquist R.M., Heitbrink D., Bolwien C., Gennis R.B., Heberle J., (2003) "Direct observation of protonation reactions during the catalytic cycle of cytochrome c oxidase" *Proc. Natl. Acad. Sci. U S A.*, **22**, 8715-20

-Ortiz-Salmerón E., Nuccetelli M., Oakley AJ., Parker MW., Lo Bello M., Garcia-Fuentes L. (2003) "Thermodynamic description of the effect of the mutation Y49F on human glutathione transferase P1-1 in binding with glutathione and the inhibitor S-hexylglutathione" *J Biol Chem*.. **278**, 46938-48

-Otwinowski Z., Minor W., (1997) "Processing of x-ray diffraction data collected in oscillation mode" *Meth Enzymol.*, **276**, 307-326

-Sheehan D., Meade G. Foley V.M., 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

-Shekhar KC. (1991) "Schistosomiasis drug therapy and treatment considerations" *Drugs*. **42**, 379-405

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

-Stenberg G., Board PG., Carlberg I., Mannervik B., (1991) "Effects of directed mutagenesis on conserved arginine residues in a human Class Alpha glutathione transferase" *Biochem J.*, **274**, 549-55

-Stevens J.M., Armstrong R.N., Dirr H.W. (2000) "Electrostatic interactions affecting the active site of class Sigma glutathione S-transferase" *Biochem. J.*, **347**, 193-197

-Stevens J.M., Armstrong R.N. Hornby J.A., Dirr H.W. (1998) "Class sigma glutathione transferase unfolds via a dimeric and a monomeric intermediate: impact of subunit interface on conformational stability in the superfamily" *Biochemistry*, **37**, 15534-15541

-Takahashi N., Hayano T., Suzuki M. (1989) "Peptidyl-prolyl cis-trans isomerase is the cyclosporin A-binding protein cyclophilin "*Nature*, **337**, 473-475

-Taylor J.B., Vidal A., Torpier G., Meyer D.J., Roitsch C., Balloul J.M., Southan C., Sondermeyer P., Pemble S., 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. (1998) "Production and testing of Schistosoma japonicum candidate vaccine antigens in the natural ovine host" *Vaccine*, **16**, 1290–1298

-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

-Thomson A.W., Chappell L.H., (1988) "Immunophenotypic analysis of blood and spleen lymphocyte subsets in rats protected against schistosomiasis by cyclosporin A" *Immunol Lett.*, **17,** 169-72

-Towensend D. M., Tew K. D., Tapiero H., (2003) "The importance of glutathione in human disease" *Biomedecine & Pharmacoterapy*, 57, Issue 3-4, 145-155

-Walsh CT., Zydowsky LD., McKeon FD. (1992) "Cyclosporin A, the cyclophilin class of peptidylprolyl isomerases, and blockade of T cell signal transduction**"** *J. Biol. Chem.*, **267,** 13115-13118

-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 Control of Schistosomiasis and soil-transmitted helminthiasis” WHO technical report series 912, Geneve

-Xiao B., Singh S.P., Nanduri B., Awasthi Y.C., Zimniak P., 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., 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

-Valle C., Festucci A., Calogero A., Macrì P., Mecozzi B., Liberti P., Cioli D., (1999) “Stage-specific Expression of a Schistosoma mansoni Polypeptide Similar to the Vertebrate Regulatory Protein Stathmin” *J. Biol. Chem*, **274**, 48, 33869-74

-Valle C., Troiani AR., Lazzaretti P., Bouvier J., Cioli D., Klinkert MQ., (2005) “Molecular and biochemical characterization of a protein cyclophilin from the nematode Haemonchus contortus” *Parasitol Res.*, **96**, 199-205

-Zhao Y., Ke H. (1996) “Mechanistic Implication of crystal structures of the cyclophilin-dipeptide complexes” *Biochemistry*, **35**, 7362-7368

Schistosoma mansoni Fatty Acid Binding Protein: Specificity and Functional Control as Revealed by Crystallographic Structure[†,‡]

Francesco Angelucci,[§] Kenneth A. Johnson,[§] Paola Baiocco,[§] Adriana E. Miele,[§] Maurizio Brunori,[§] Cristiana Valle,[||] Fabio Vigorosi,[||] Anna Rita Troiani,[||] Piero Liberti,[||] Donato Cioli,[||] Mo-Quen Klinkert,[⊥] and Andrea Bellelli*[,§]

Department of Biochemical Sciences "A. Rossi Fanelli", CNR Institute of Molecular Biology and Pathology and Istituto Pasteur-Fondazione Cenci Bolognetti, University of Rome "La Sapienza", Piazzale Aldo Moro 5, 00185 Rome, Italy, CNR Institute of Cell Biology, 32 Via Ramarini, 00016 Monterotondo, Rome, Italy, and Bernhard-Nocht-Institut für Tropenmedizin, 74 Bernhard-Nocht-Strasse, 20359 Hamburg, Germany

Received July 14, 2004

ABSTRACT: *Schistosoma mansoni* fatty acid binding protein (Sm14) was crystallized with bound oleic acid (OLA) and arachidonic acid (ACD), and their structures were solved at 1.85 and 2.4 Å resolution, respectively. Sm14 is a vaccine target for schistosomiasis, the second most prevalent parasitic disease in humans. The parasite is unable to synthesize fatty acids depending on the host for these nutrients. Moreover, arachidonic acid (ACD) is required to synthesize prostaglandins employed by schistosomes to evade the host's immune defenses. In the complex, the hydrocarbon tail of bound OLA assumes two conformations, whereas ACD adopts a unique hairpin-looped structure. ACD establishes more specific interactions with the protein, among which the most important is a π-cation bond between Arg78 and the double bond at C8. Comparison with homologous fatty acid binding proteins suggests that the binding site of Sm14 is optimized to fit ACD. To test the functional implications of our structural data, the affinity of Sm14 for 1,8-anilinonaphthalenesulfonic acid (ANS) has been measured; moreover the binding constants of six different fatty acids were determined from their ability to displace ANS. OLA and ACD exhibited the highest affinities. To determine the rates of fatty acid binding and dissociation we carried out stopped flow kinetic experiments monitoring displacement by (and of) ANS. The binding rate constant of ligands is controlled by a slow pH dependent conformational change, which we propose to have physiological relevance.

Schistosomiasis is the second most prevalent parasitic disease worldwide and affects more than 200 million people in developing countries. Schistosomes are parasitic trematodes whose complex life cycle involves an intermediate host (a freshwater snail) and man as the definitive host. Adult schistosomes live in the mesenteric or perivesical veins of their definitive host and uptake their nutrients directly from the host's blood.

The proteins that participate in the uptake and metabolism of fatty acids and their derivatives may constitute a possible target for therapy or vaccination, given the numerous and important roles played by these compounds. To briefly review this subject, we remark that the parasite lacks the metabolic pathways required for the biosynthesis of sterols and lipids; hence it is completely dependent on the host for these substances (*1*). Besides being nutrients and structural components of the cell membrane, fatty acid derivatives released by schistosomes play a role in the parasite evasion from the host immune response (*2*). Moreover, upon contact with human skin, cercariae (the larval stage released by the intermediate host) respond to chemical stimuli, particularly medium-chain free fatty acids, to start skin invasion (*3*). 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)[1] (*4*). Sm14 is present in all the stages of the life cycle and is localized in the external cell layer, i.e., near the interface of the parasite/host contact (*5*). From this short summary an important conclusion may be drawn: interfering with fatty acid uptake or metabolism may constitute an important therapeutical approach. Accordingly, the World Health Organization selected Sm14 as one out of six antischistosome vaccine candidates for testing (*6*, *7*); and possibly the protein may also be a drug target, since blocking of fatty acid uptake could have dramatic effects on the life

[†] Funding was received by the University of Rome "La Sapienza" (Progetto Ateneo 60%—anno 2003) and by the Ministero dell'Università e della Ricerca Scientifica e Tecnologica of Italy (PRIN 2003). Elettra (Trieste, Italy) provided generous fellowships to K.A.J. and P.B.

[‡] Sm14−ACD and Sm14−OLA have been deposited with PDB codes 1VYF and 1VYG, respectively.

* To whom correspondence should be addressed. Tel: +39 06499 10236. Fax: +39 06444 0062. E-mail: andrea.bellelli@uniroma1.it.

[§] University of Rome "La Sapienza".

[||] CNR Institute of Cell Biology.

[⊥] Bernhard-Nocht-Institut für Tropenmedizin.

[1] Abbreviations: FABP, fatty acid binding protein; Sm14, *Schistosoma mansoni* FABP; SjFABP, *Schistosoma japonicum* FABP; Eg-FABP1, *Echinococcus granulosus* FABP; I-FABP, intestinal FABP; H-FABP, heart/muscle FABP; A-FABP, adipocyte FABP; B-FABP, brain FABP; M-FABP, myelin P2 FABP; L-FABP, liver FABP; CMC, critical micellar concentration; FA, fatty acid; OLA, oleic acid; ACD, arachidonic acid; LA, linoleic acid; PA, palmitic acid; MA, myristic acid; DA, decanoic acid.

10.1021/bi048505f CCC: $27.50 © 2004 American Chemical Society
Published on Web 09/25/2004

cycle of the parasite. This hypothesis is particularly relevant if one takes into account that there are only two available drugs against this disease (*8*).

The FABPs constitute a multigenic family of low molecular mass proteins (14−15 kDa). Analysis of 3D structures of FABPs revealed basic structural similarities among the members of this family despite low sequence identities (*9*). The actual physiological role of FABPs is still incompletely clear, 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 of their acyl-CoA derivatives, storage of FAs, lipid trafficking, and regulation of cell growth and differentiation (*10*, *11*). Ambiguities mainly originate from the observation that these proteins can bind many different hydrophobic substrates (FAs, monoglycerols, diacylglycerol phosphates, lipooxygenase metabolites of arachidonic acid, acetyl Co-A, retinoids, and even heme (*12*)); also, FABPs are able to reversibly associate with artificial phospholipid bilayers, releasing or uptaking ther ligands (*11*). Thus, a very credible hypothesis on the functional role of FABPs suggests that these are involved in the intracellular trafficking of lipids to and from cell membranes; an important piece of evidence is that the heart fatty acid uptake is decreased in heart-fatty acid binding protein gene-ablated mice (*13*). The important metabolic role of FABPs and structural similarity of FABPs in evolutionary related parasites may explain why immunization with Sm14 confers significant levels of cross-protection against infections with *Schistosoma mansoni* and *Fasciola hepatica* in animal models (*6*).

Despite the importance of Sm14 as a vaccine and drug target candidate, there is still little structural and functional data about this protein. Thus, in this work, 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 1.85 Å, respectively. Moreover, we characterized the binding and release reactions of Sm14 and several important FAs using 1,8-anilinonaphthalenesulfonic acid (ANS) as a competitor. The ligand with highest affinity, when water solubility is taken into account (*14*), is ACD, a compound essential to schistosomes. ACD is incorporated by the parasite more readily than other lipids (*15*) and is the immediate precursor of eicosanoid hormones, including the prostaglandins, critical in facilitating the skin penetration process of cercariae (*16*). 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. Indeed, ACD establishes a series of tight specific interactions, among which the most important one is a strong directional π-cation interaction between the guanidinium group of Arg 78 and the C8−C9 double bond. The shape and volume of the binding cavity is practically unchanged in the two complexes, but seems best adapted to bind ACD. Finally, the 3D structures of Sm14 also allowed us to structurally highlight two previously described antigenic determinants of the protein.

EXPERIMENTAL PROCEDURES

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 units/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 was 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′-GTTGAAACATATGTCT-AGTTTCTTGGGA-3′ and the reverse primer 5′-TTGCTC-GAGTTAGGATAGTCGTTTATAATT-3′. The PCR reaction was carried out using 1 μL of template, 50 pmol of each primer, and 2.5 units 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 *Escherichia 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.

Purification of Sm14. Wet cells of a 3 L bacterial growth were resuspended in 40 mL of buffer A (0.5 M NaCl, 50 mM Tris/HCl pH 8.0, 2 mM β-mercaptoethanol, 1 mM EDTA, 2% Triton X-100, and 1 mM protease inhibitor phenylmethylsulfonyl fluoride) (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 $10000\times g$ 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, 20 mM Tris/HCl pH 7.8, 500 mM NaCl, 5 mM imidazole, and 20 mM β-mercaptoethanol and gently stirred overnight at 4 °C. The resulting solution was filtered with a 0.45 μm 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 (20 mM Na-phosphate pH 7.8, 500 mM NaCl) resulted in a properly folded His-tagged Sm14, which was eluted at about 250 mM imidazole by a linear gradient obtained by mixing buffer C with the same containing 500 mM imidazole. SDS−PAGE (*17*) showed the presence of the pure His-tagged Sm14 characterized by a molecular mass of 18.5 kDa. Then, 50 mg of pure His-tagged Sm14 was dialyzed against 100 mM Tris/HCl pH 8.0, 250 mM 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 (2 M urea, 400 mM NaCl, 100 mM Tris/HCl pH 8.0, 10 mM $CaCl_2$, 2 mM β-mercaptoethanol), and 10 units/mL of

Table 1: Summary of Data Collection and Refinement

	Sm14−OLA	Sm14−ACD
space group	$P2_12_12$	$P2_12_12$
no. of unique reflns	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
unit cell dimens		
a (Å)	43.26	42.29
b (Å)	92.39	91.01
c (Å)	36.22	35.33
R_{merge}	0.046 (0.38)[a]	0.067(0.17)[b]
Ramachandran plot		
most preferred	116	110
allowed	7	13
generously allowed	0	0
disallowed	0	0
R	0.19	0.21
R_{free}	0.24	0.28
rms deviations		
bond lengths (Å)	0.02	0.015
bond angles (deg)	1.6	1.4
final model composition		
no. of protein atoms	1071	1057
no. of water molecules	209	76
no. of ligand molecule	1 (OLA)	1 (ACD)

[a] Last shell (1.88−1.85 Å). [b] Last shell (2.49−2.40 Å).

thrombin from bovine plasma (Sigma-Aldrich) was added. The resulting solution was stirred overnight at room temperature, then dialyzed against 0.1 M 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 100 mM salt and verified by SDS−PAGE.

Crystallization of the Complex with FAs. Prior to crystallization the protein was equilibrated in 10 mM MES pH 5.5, 50 mM NaCl, 1 mM 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 1 mg/mL solution of Sm14, so that the ethanol concentration did not exceed 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 10 mg/mL. Crystallization was achieved using the hanging drop vapor diffusion method, employing drops consisting of 1 μL of a well solution of 0.1 M MES pH 6.0, 0.2 M Na acetate, 29−31% PEG 8K (w/v), 5 mM β-mercaptoethanol, and 1 μL of the protein sample. Rodlike crystals of dimensions 0.4 × 0.2 × 0.2 mm grew within a few days at 21 °C. Crystals were cryoprotected by using a solution containing 10% glycerol (v/v), 30% PEG 8K, 0.1 M MES, 0.2 M Na acetate, 5 mM β-mercaptoethanol, pH 6.0 and flash-frozen in liquid nitrogen.

Data Collection and Processing. Data for the two structures were collected at Elettra (Trieste, Italy) at 100 K and then processed using the HKL suite (*18*). Both Sm14−OLA and Sm14−ACD complexes crystallized in the same $P2_12_12$ space group with the unit cell dimensions given in Table 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 (*19*), to be around 45% of total crystal volume, and in both cases there was only one complex molecule per asymmetric unit.

Structure Solution and Refinement of Sm14−OLA Complex. The structure was solved by molecular replacement techniques with the program AMoRe (*20*) of the CCP4 suite (*21*), using the structure of the M-FABP complex (PDB code 2HMB; ref *22*) 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 (*23*). For model building both $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ 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 (*24*), for 13179 reflections between 20.0 Å and 1.85 Å, resulted in a complete model of Sm14−OLA complex, with an *R* value of 0.198 and an R_{free} value of 0.246 (see Table 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.

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 fashion similar to that described above. The difference density located in the binding cavity was modeled as C20:4 fatty acid. Final refinement led to an *R* factor of 0.213 and an R_{free} value of 0.283 for 5746 reflections between 20.0 Å and 2.4 Å. The final structure contained 135 amino acid residues, 77 water molecules, and 1 molecule of ACD (see Table 1) and displayed only the side chains of Met 20 and His 14 in double conformations.

Structural Analysis. The models were monitored for geometrical quality by using PROCHECK (*25*). The *B* factors for several structural elements were calculated using the subroutine BAVERAGE of the CCP4 program (*21*). The contacts between FAs, residues, and waters were considered up to 4.6 Å and measured with the program CONTACT (*21*). The rms deviations between structures were calculated with the program LSQKAB (*21*). Volume and surface of the protein cavity were calculated with Cast-P (*26*). Sm14−OLA and Sm14−ACD structures have been deposited in the PDB with entry codes 1VYF and 1VYG, respectively.

Spectrofluorimetry and Fluorescence-Based FA 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 (*27*). The experiments were carried out using a Spex Fluoromax spectrofluorimeter at 20 °C in 2 mL of 0.1 M MES pH 5.5. Sample solutions of Sm14 were prepared by diluting a stock to concentrations of 0.1 μM for K_d determinations and to 5 μM for stoichiometric experiments. The stock solution (7 mM) of the fluorescent probe was prepared in ethanol. Aliquots of appropiately diluted solutions of ANS in 0.1 M MES were added to the protein and mixed in a cuvette with a magnetic stirrer for 2 min before spectra were collected. ANS and protein concentrations were verified spectrophotometrically (ϵ_{280} = 12750 M^{-1} cm^{-1} for Sm14, ϵ_{280} = 4990

Table 2: Equilibrium Constants for the Dissociation of Six Fatty Acids from Sm14 and Geometrical Parameters of Sm14−OLA and Sm14−ACD

fatty acid		CMC[a] (nM)	K_d (nM)	surfaces (Å^2) cavity	surfaces (Å^2) fatty acid	cavity vol (Å^3)
decanoic acid	(C10:0)		6100			
myristic acid	(C14:0)		88			
palmitic acid	(C16:0)	4000	33			
oleic acid	(C18:1)	6000	9	467.9	521.2	318.9
linoleic acid	(C18:2)	13000	24			
arachidonic acid	(C20:4)	20000	10	441.1	608.2	300.1

[a] Critical micellar concentration (*28*).

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 4.5 μM 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 10 mM, 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 (*28*); 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 2) lie within 0.2%.

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 cutoff ≤ 455 nm. Solutions of Sm14 (0.4 μM) in 0.1 M MES pH 5.5 were mixed with buffered solutions of ANS at concentrations ranging between 3 and 10 μM at 20 °C. The time courses were fitted to a double exponential.

The competition measurements were performed in two ways: Sm14 0.4 μM was first incubated with ANS 50 μM and mixed with different amounts of OLA (1.25−5 μM); alternatively the protein was incubated with OLA (1.5 μM) and mixed with different solutions of ANS (100 to 400 μM). These experiments were carried out under suboptimal 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 (0.4 μM) containing ANS at a concentration ranging from 2 to 3.3 μM at one pH (10 mM Mes pH 5.5 or 10 mM Hepes pH 7.4) is mixed with an equal volume of a concentrated buffer solution at the reciprocal pH (0.1 M Hepes pH 7.4 or 0.1 M 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 (0.4 μM) dissolved in 10 mM HEPES pH 7.4 was mixed with 0.1 M Mes pH 5.5, and the solution was "aged" for different times (ranging from 15 ms to 150 s). After the preset delay, the resulting solution was mixed with an equal volume of a buffer containing 0.1 M Mes pH 5.5 and 4 μM ANS.

RESULTS

Overall Structure of Sm14−FA Complexes. The complexes of Sm14 with ACD and OLA were crystallized under the experimental conditions described in the Experimental Procedures section. We also attempted to crystallize the FA-free Sm14 without success. The 3D structures of Sm14−ACD and Sm14−OLA were solved at 2.4 and 1.85 Å, respectively, by X-ray crystallography with the two models achieving final *R* factors of 21.3% and 19.8%, respectively. As assessed by the program PROCHECK (*25*), the stereochemical parameters of the two structures lie within the mean reference values (see Table 1 for a summary of crystallographic data). The protein main chain atoms and all the protein atoms of the two complexes of Sm14 are superimposable, with an rmsd of 0.32 Å and 0.75 Å, respectively. At the N-terminus both structures have two partially disordered amino acids, belonging to the linker region after the His tag (see Experimental Procedures). Numbering of the amino acids starts from the first native residue (Met 1), and the two non-native residues are called Gly −1 and Ser 0.

Similarly to the heart group of FABPs (H-FABPs), Sm14 folds as a twisted V-shaped β-barrel, formed by 10 antiparallel β-strands named from A to J (following the annotation of Sacchettini et al. (*29*)), closed on one side by interactions between the side chains and capped on the other side by a helix(I)−turn−helix(II) motif (Figure 1, panels A, C).

The two complexes have one single molecule of either ACD or OLA, bound in a large internal cavity. The volume of the cavity (calculated after removal of FA and waters from the models) is ∼300Å^3 (see Table 2), comparable to the cavity volumes of other FABPs belonging to the same family (*7*).

The gateway for entry of the substrate in the internal cavity is formed by the helix(I)−turn−helix(II) motif (residues 16−35) and by two hairpin loops between strands C−D and E−F (residues 56−58 and 74−78, respectively) (Figure 1). In both complexes, the gateway is closed and leaves no room for exit of the bound FAs; we hypothesize that diffusion of FAs into and out of the cavity requires opening of this doorway. In both structures Phe 57 on the C−D loop, previously proposed to be a key residue in the binding process (*30*), closes the entry of the lid, it is oriented toward the interior of the cavity, and its ring is at 3.5 Å to 4.5Å distance from the aliphatic chain of both OLA and ACD, respectively. In Sm14−OLA complex, Phe 57 has a higher *B* factor (41.1Å^2) as compared to the overall *B* factor of the protein (36.1 Å^2), index of certain mobility, whereas in Sm14−ACD, Phe 57 has a *B* factor (53.6Å^2) similar to that of the whole protein.

On the rim of the portal region, Lys 22 at the end of helix-(I) and Lys 58 on the C−D loop have the same relative

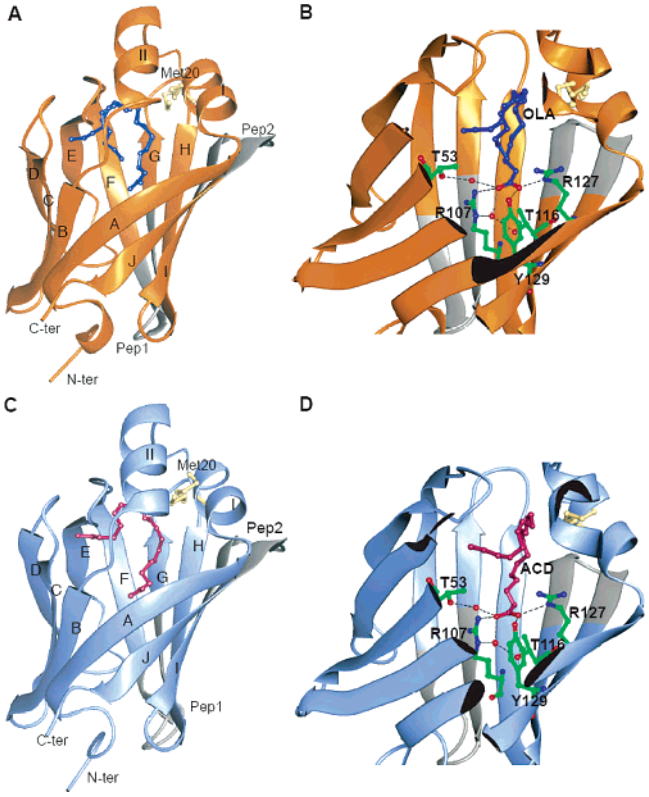


FIGURE 1: 3D structures of Sm14−OLA and Sm14−ACD complexes. (A) Overall structure of Sm14−OLA with secondary structure elements highlighted. Bound OLA (in blue ball-and-stick) is shown in its double conformation. Secondary structure elements and the corresponding residues: Gly 6-His 14 β-strand A, Phe 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. Residues belonging to the antigenic peptides are colored in light gray, and Met 20 is shown in magnolia ball-and-stick. (B) Zoom of the binding cavity with the network of H-bonds fixing the carboxyl headgroup of OLA highlighted. (C) Overall structure of the Sm14−ACD complex. Strands and helices are labeled as in A. ACD is shown in purple ball-and-stick. (D) Same as panel B, for Sm14−ACD complex.

position and orientation as their homologues in H-FABP (Lys 22, Lys 59), where they were shown to interact electrostatically with phospholipid headgroups and supposed to drag the carboxylate of FA toward the internal cavity (*31*). On the back of the α-helix lid, the hydrophobic side chains of Trp 27 and Ile 32 are exposed to the solvent. This unusual orientation allows Trp 27 to insert among the phospholipid tails, thereby helping membrane binding, as shown by Kennedy et al. (*32*) for *Schistosoma japonicum* FABP. In both our structures Trp 27 and Ile 32 have the same relative orientation, consistent with the hypothesis that they interact with the cell membrane.

Structural Features Related to Antigenicity. 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 (*6, 33*). The two peptides, namely Pep1 (from residue 85 to 94) and Pep2 (from residue 118 to 125), are topologically far from each other in Sm14 crystal structure (Figure 1). Pep1, seen at the bottom of the model in Figure 1, 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 αI-helix. Most of the side chains of these peptides 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 structure reveals that both peptides are β-hairpins, i.e., they possess a stable secondary structure motif possibly retained even when the peptides are cleaved from the protein; and their sequence is sufficiently different from the homologous human isoforms to be recognized as "nonself" (at most 33% and 28% sequence identity, respectively).

The residue at position 20 of Sm14 is relevant to antigenicity. In our protein this residue is Met, but a natural allelic variant containing Thr has been described (Sm14-T20). Ramos and co-workers (*34*) studied the effect of mutations of Met 20 to Ala and Thr with respect to antigenicity, and observed a different protective response against *S. mansoni* cercariae in mice. Sm14-M20 displayed greater antigenicity than wild-type Sm14-T20 and artificial mutant Sm14-A20. Met 20 (Figure 1) is at the beginning of αI-helix, but it is far from the two antigenic epitopes; thus its effect can only be indirect. In both our structures, Met 20 presents a double conformation, with 50% occupancy; in one conformation the sulfur atom lies in proximity with the C9−C10 double bond in Sm14−OLA (3.8Å) and of the C11−C12 double bond in Sm14−ACD (3.7 Å), whereas in the other it is found between the two helices (hence out of the FA binding site). The latter conformer may play a critical role in stabilizing the closed conformation of the pocket lid when a FA is bound. It is possible, but it is not obvious from our structures, that the reason of the higher antigenicity of Sm14-M20 over Sm14-T20 lies in its greater thermodynamic stability (*35*).

Structure of Bound Fatty Acids. The greatest energy contribution for FAs binding to FABPs is expected to be due to the extraction of the hydrophobic molecule from the aqueous solvent; specific interactions between the bound FA and amino acid side chains may be responsible for ligand selectivity. These were examined looking at the level of the internal surface of the cavity by comparing the structures of the OLA and ACD complexes. The cavity is lined by polar and hydrophobic amino acids, whose side chains are mostly oriented toward the interior of the protein.

Bound OLA is completely surrounded by protein atoms and by structural waters (Figure 1, panel B). The carboxylate is at H-bond distance with Arg 127 N(ϵ), with Tyr 129 OH and with two structural water molecules, W12 and W78. Although the relative orientation of the carboxylate is well maintained in most of the FABPs, the geometry of H-bonds in the binding site is often different.

Bound OLA is found in two conformations, which differ in the orientation of the last five carbon atoms (Figure 2, panel A). Conformer 1 is found to be in the same U-shaped conformation previously observed for other FABPs (*30, 36*) and occupies the same portion of the cavity. Conformer 2 displays a chain bending similar to that found for ACD in the Sm14−ACD complex (see below), with the last part of the aliphatic chain out of the plane defined by the rest of the molecule. The difference maps clearly show electron density for both conformations of OLA from the C14 to C18 atom of the aliphatic chain. The two conformers are compatible with only slight differences in the positions of the first thirteen carbon atoms (Figure 2, panel A). The

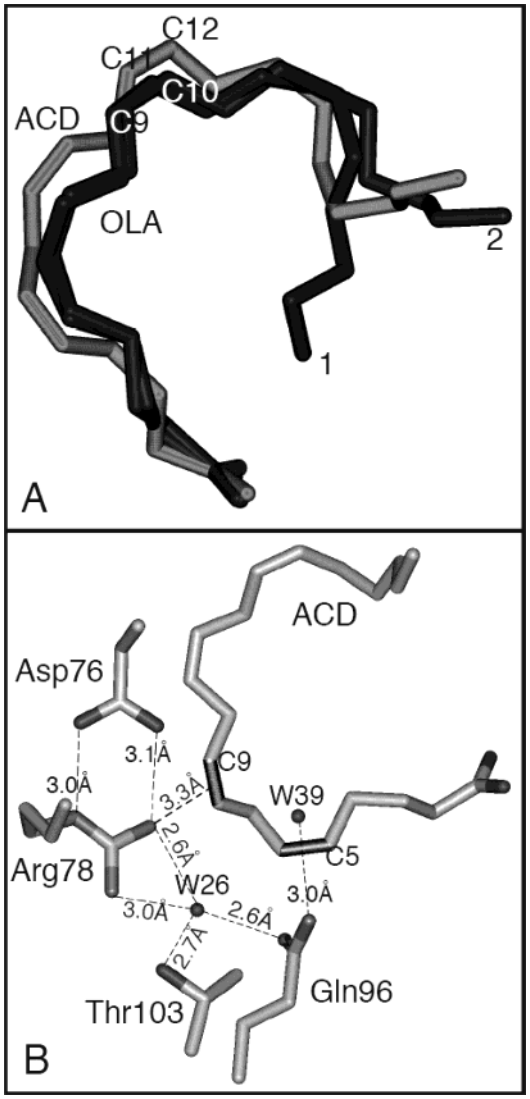


FIGURE 2: Structures of Sm14 bound OLA and ACD. (A) Representation of OLA and ACD bound to Sm14. The partial superimposition of the C9−C10 double bond of OLA (in dark gray) and the C11−C12 double bond of ACD (in light gray) demonstrates a similar configuration of the two FAs in the binding pocket. OLA 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 toward the carboxylic moiety, yielding a V-shaped conformation. (B) 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 the Sm14−OLA complex.

relative occupancy of the two conformers is 50%, and their average *B* factors (43 Å^2) are above the average *B* factor of the side chains of the surrounding cavity (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.

The hairpin conformation of bound OLA is induced by the cis double bond between C9 and C10, and by gauche bonds between C5 and C6, C13 and C14, and stabilized by a large number of van der Waals interactions with the protein side chains and with 7 ordered water molecules in conformer 1 and 6 in conformer 2. In both conformers the C9−C10 double bond interacts with the conserved Phe 16 and Met 20 (only in one of the two conformations of Met20) and with Val 25, Ser 75, Asp 76, Arg 78.

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 (Phe 57 in C−D loop; Ser 75, Asp 76, and Arg 78 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, but it has a single conformer 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 that of the surrounding polypeptide chain (48 Å^2). The carboxylic moiety makes the same types of contacts as OLA (Figure 1, panel D) and 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 although closer in space. Moreover, the long chain of ACD and the presence of four cis double bonds make its structure rigid, force its convex shape, and consequently increase the contact surface between the protein and the lipid, as compared to OLA (Figure 2, panel A). Several specific sets of contacts with 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).

In both ACD and OLA complexes, residue Met 20 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 2, panel A), 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 fit 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, whereas 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 2, panel B; Figure 3, panels A, C). It is interesting to note the continuous contact surface generated by the side chains and by 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.

The mobility of the E−F loop is a relevant structural element of FABPs since it controls the access of the FA to the internal cavity (*37*). In Sm14 and the FABPs belonging to the H family, the role of the E−F loop is even more important since it bears the residues Thr 74, Asp 76, and Arg 78. Our structures demonstrate that Arg 78 establishes a strong, directional π-cation interaction with C8−C9 double bond of ACD (Figure 2, panel B; Figure 3, panel A), whereas

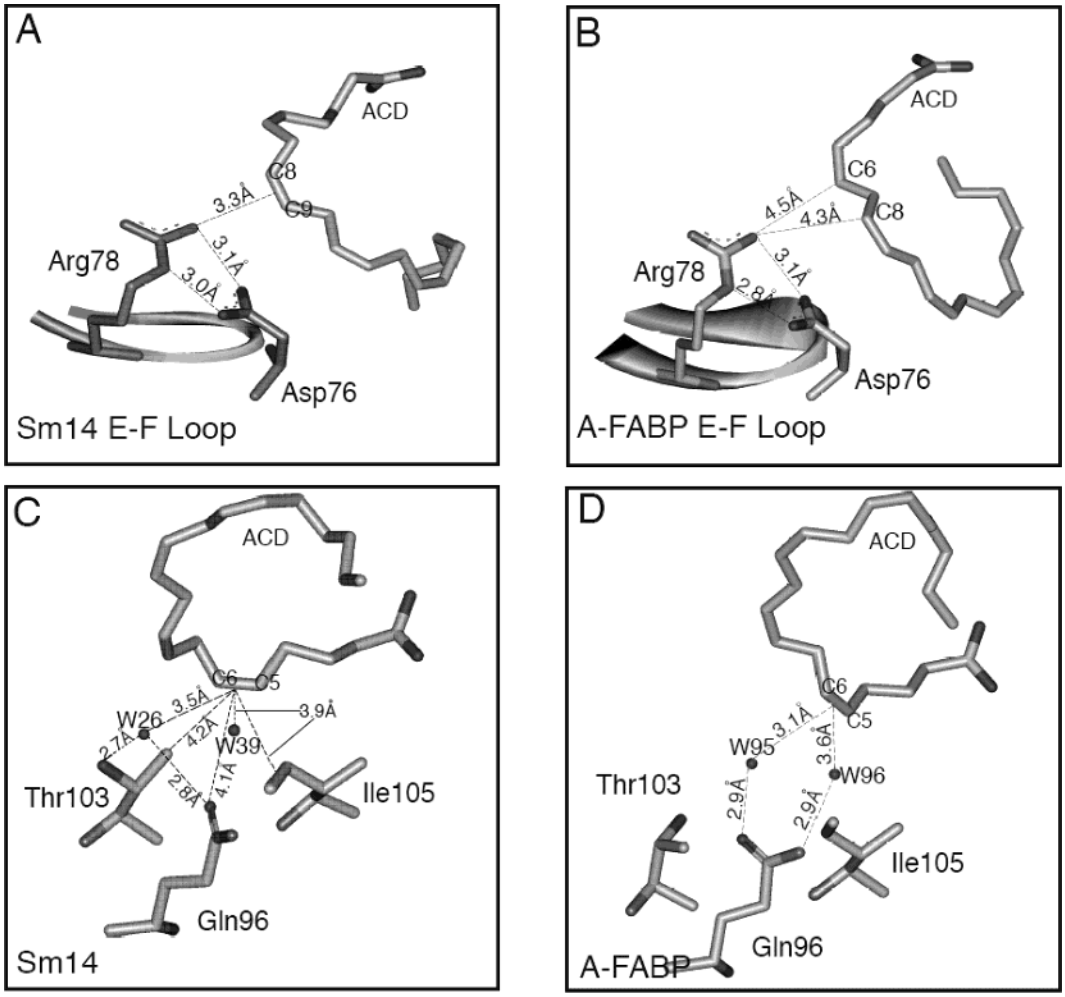


FIGURE 3: Detail of the interactions between Sm14 and ACD. (A) Network of interactions between Sm14 and ACD C8−C9 double bond. Residues belonging to the E−F loop are highlighted. H-bonds between Asp 76 and Arg 78 keep the terminal NH2 of Arg 78 in orbital overlap with C8−C9 double bond. (B) Network of interactions between residues belonging to the E−F loop of murine A-FABP (1ADL, LaLonde et al., 1994) and bound ACD, in the same orientation as in panel A. ACD is bound with a different geometry; hence the conserved Arg 78 loosely contacts the region between C5−C6 and C8−C9 double bonds. (C) The Sm14−ACD complex has been rotated by 180° around the *x* axis with respect to panel A to show the contacts between the pocket and C5−C6 double bond. Interaction distances have been measured between conserved residues and water molecules and the π orbital of the double bond. (D) Same as in C for the A-FABP−ACD complex. In this case only one of the conserved residues contacts ACD via the two water molecules.

Thr 74 and Asp76 are involved in a network of H-bonds that helps in bending the E−F loop toward the interior of the protein and hence toward ACD. Moreover, Asp 76 establishes a double H-bond with Arg 78 that stabilizes the position of the latter in close proximity with the bound FA (Figure 3, panel C). Consistently, in the complex with OLA, which lacks the double bond at C8−C9, the interaction between FA and Arg 78 is looser with the *B* factors of E−F loop greater by about 10 Å^2 with respect to the mean value of the whole protein, whereas in the Sm14−ACD structure they are 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 are accommodated in the pocket with respectively lower or higher shape complementarity, due to their intrinsic geometrical and stereochemical parameters.

Ligand Binding Properties of Sm14. Previous work by Moser et al. (*4*) showed that Sm14 binds ^{14}C-labeled palmitic and linoleic acid at physiological pH, with dissociation equilibrium constants of ~2 μM. Using the environment-sensitive fluorescent probe ANS, we carried out a thorough equilibrium and kinetic investigation of complex formation between Sm14 and six FAs, differing in length and degree of saturation,. Binding of ANS to Sm14 is associated with a substantial increase in fluorescence intensity and a blue shift of the maximum emission, indicative of transfer from water to an apolar environment (*38*). Moreover, the crystallographic structure of the adduct of ANS with murine adipocyte lipid binding protein (A-FABP) shows that this probe is located in the FA's binding site with the sulfonate moiety placed between the two arginines involved in the FA's carboxylate binding site (see ref *39* and above). Given the low solubility of FAs and the high affinity of FABPs for these ligands, a method based on a competitive fluorescent probe seemed ideally suited to determine formation of a complex between Sm14 and various ligands.

The functional properties of Sm14 were initially characterized at pH 7.4 and 5.5. The acidic pH was preferred for three reasons: (i) the equilibrium isotherm of the reaction between ANS and Sm14 at physiological pH (7.4) is complex and suggests a stoichiometry greater than 1; on the other hand, at pH 5.5, Sm14 binds ANS with 1:1 stoichiometry and high affinity (K_d = 0.5 μM); (ii) at this pH the protein is stable for a long time; and (iii) crystals of Sm14 were obtained at

a similar pH (see Experimental Procedures), hence granting a better comparison between functional and structural data. For comparison, the K_d values reported for ANS binding to A-FABP and I-FABP ranged from 1 to 50 μM, under different conditions (*27*, *38*).

Addition of FAs to the Sm14−ANS complex quenches emission at 475 nm, indicating that FAs displace ANS. The best fit analysis of the titration data is compatible with a single FA molecule binding. The values of K_d calculated for 6 FAs are reported in Table 2. The results show that formation of the complex between Sm14 and OLA or ACD has essentially the same dissociation constant: 9 and 10 nM, respectively. The K_d values for decanoic, myristic, palmitic, and linoleic acids are respectively 610-, 9-, 3-, and 2.5-fold greater than for 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 (*28*). This is 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 favorable 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 enthalpic contributions to binding; (iii) given their low pK_a ($\approx$4.7), FAs are predominantly in the anionic form, both at pH 5.5 and at pH 7.4, thus favoring the interactions with the conserved positively charged guanidinium group of Arg 127 (see the section on the structure above). However, these same considerations challenge the data reported in Table 2, because one would expect that the lower the solubility of a given FA, the higher the apparent affinity. Indeed, this occurs in several FABPs that bind FAs with affinities correlated to their solubilities (*14*), i.e., in agreement with the expectation that the greater energy contribution to binding comes from the extraction of the poorly soluble FA from the aqueous phase. The case of Sm14 and unsaturated FAs violates this rule, and ACD is particularly noteworthy since its critical micellar concentration is relatively high, yet its K_d is the lowest in the group (together with that of OLA). This result is explained by the structure of the Sm14−ACD complex, which reveals specific interactions at the level of the two double bonds C5−C6 and C8−C9, thus implying a specificity for this physiologically important ligand.

ANS could not be replaced by prostaglandins, up to μM concentrations, proving that Sm14 does not bind these derivatives of ACD, even at concentrations higher than those prevailing under physiological conditions in the human body.

The differences between the K_ds reported in Table 2 at pH 5.5 and older data (*4*) can be rationalized by the difference in pH. Indeed, when the same ligand titration experiments 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 pH increases is observed for both ANS and FAs and may be of physiological relevance in the release of bound FAs (see below).

Kinetics of Complex Formation. The kinetic rate constants for binding and release of ANS and OLA were determined by fluorescence stopped flow. 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 physiological requirements. At pH 5.5, ANS combination with Sm14 is fast, and its second-order rate constant is $\geq 2 \times 10^8$ M^{-1} s^{-1}, thus approaching the diffusion limit. Given the affinity of ANS ($K_d = 0.5$ μM) and the sensitivity of the instrument, it is impossible to dilute the ligand enough to obtain a more precise estimate. The kinetics of dissociation of ANS bound to Sm14 can be followed by displacement, adding an excess of OLA. The recorded time courses have half-times of $\sim$1 s and are best described by two exponentials, irrespective of the concentration of OLA (data not shown). Although it is obvious that the observed slow fluorescence decrease is associated with the release of ANS, it is unclear if 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.

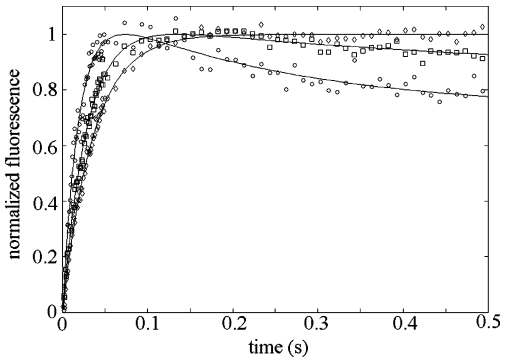


FIGURE 4: Time course of oleic acid replacement by ANS. Sm14 (0.4 μM) incubated with OLA (1.5 μM) in Mes pH 5.5 is rapidly mixed with ANS buffered solutions at concentrations of 100 (◇), 200 (□), and 400 μM (○). Experimental points are fitted with a model including a ternary complex between Sm14−OLA−ANS. 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.

Release of bound OLA was achieved by rapidly mixing the complex with a 70- to 1000-fold excess of ANS (Figure 4). Even at the lowest concentration, ANS is able to displace over 50% of the bound OLA. The observed rate constants for the first process vary from 92 s^{-1} (at the highest ANS: OLA ratio) to 28 s^{-1} (at the lowest ratio), leading to an extrapolated dissociation rate constant for the complex between Sm14 and OLA of 200 s^{-1}. This value seems unreasonable because, given the high affinity of the complex ($K_d = 9$ nM), the calculated combination rate constant would have to be as high as 2×10^{10} M^{-1} s^{-1}. Therefore, we assume that the reaction mechanism is complex and involves a ternary complex of Sm14 with both ANS and OLA to the protein as an unstable intermediate, from which OLA would then dissociate. Accordingly, the main fluorescence increase (Figure 4) could be assigned to the second-order combination between the complex Sm14−OLA and ANS (rate constant of 3×10^6 M^{-1} s^{-1}), whereas the subsequent fluorescence decrease (rate constant of 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 consistent with data for *Echinococcus granulosus* FABP (EgFABP), a carrier with high sequence identity with Sm14, which at pH 7.4 binds more than one ligand (*40*). The rate

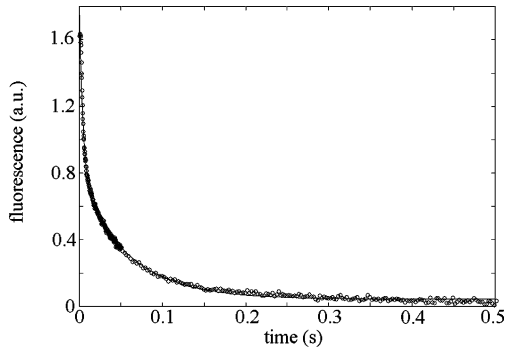


FIGURE 5: Time course of ANS release after a pH jump from 5.5 to 7.4: dissociation of ANS, following the pH jump 5.5→7.4, monitored by a decrease in fluorescence. A solution of 0.4 μM Sm14 containing 2 μM ANS in 0.01 M Mes buffer pH 5.5 was mixed in the stopped flow apparatus with an equal volume of 0.1 M Hepes buffer pH 7.4.

of release of the bound FA from Sm14 is independent of the ionic strength of the medium.

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 the 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 5). It is often observed that pH changes affect the ligand affinity of proteins that function as carriers of small ligands; e.g., this occurs for hemoglobin and transferrin. Since this mechanism helps the protein to upload its ligand even when its concentration is low, and to download it even when its concentration is relatively high, therefore it seemed worth of detailed investigation.

Unexpectedly, after a pH jump the change in ligand affinity occurred at a relatively slow rate. Indeed, when Sm14 equilibrated in diluted Hepes buffer at pH 7.4 was rapidly mixed with a concentrated buffer at pH 5.5 containing ANS (Figure 6, panel A, 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 6, panel A, trace 2). This shows that the low-pH conformation, corresponding to the fast reacting state, was not yet populated immediatly after the acidification to pH 5.5. To determine the rate constant of the pH dependent 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 while the relative amplitude of the two processes (corresponding to the two kinetic phases) was the only parameter that varied with pH. The results of this analysis are shown in Figure 6, panel B; the rate constant of the conformational change is 0.1 s^{-1}, i.e., possibly compatible with physiological requirements.

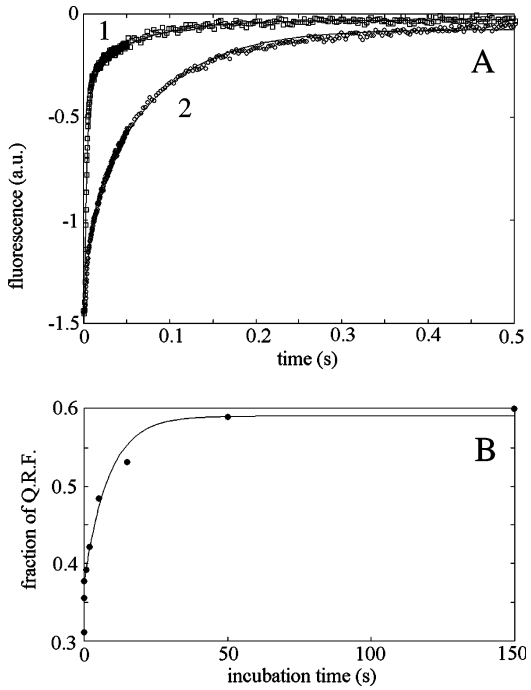


FIGURE 6: Time course of ANS combination after a pH jump from 7.4 to 5.5. (A) Trace 1: Binding of ANS at pH 5.5, monitored by an increase in fluorescence, is fast. A solution of 0.4 μM Sm14 in 0.1 M Mes buffer pH 5.5 was mixed with an equal volume of a solution of 3.3 μM ANS in the same identical buffer. 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 0.4 μM Sm14 in 0.01 M Hepes buffer pH 7.4 with an equal volume of a solution of 3.3 μM ANS in 0.1 M Mes buffer pH 5.5. Other conditions: $T = 20$ °C; excitation wavelength 380 nm, emission wavelength ≥ 455 nm. (B) Fraction of the quickly reacting form (QRF) of Sm14 as a function of the incubation time at pH 5.5 after a pH jump 7.4→5.5. A solution of 0.4 μM Sm14 in 0.01 M Hepes buffer pH 7.4 was first mixed with an equal volume of 0.1 M Mes buffer pH 5.5, and then, after variable incubation time, with a double volume of a solution of 4 μM ANS in 0.1 M 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 $\tau^{-1} = 0.1$ s^{-1}.

DISCUSSION

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* (*6*). The protein is present in all the stages of the life cycle of the parasite which occur in the definitive host, i.e., schistosomulum, adult worm, and eggs (*5*). In the adult worm the basal lamella of the tegument and the gut epithelium are strongly labeled by immunofluorescent probes against Sm14. These tissues have a high flow of lipids, supporting the putative role of Sm14 as an intracellular carrier of FAs (*4*). Therefore a detailed investigation of its structure and function is demanded and is reported in this paper.

Comparison with Other FABPs. FABPs are grouped according to sequence comparison. Within this classification, Sm14 belongs to the family group represented by H-FABP (mammalian heart, muscle, brain, adipose tissue, and myelin FABPs; parasite Sj- and Eg-FABP (*41*)). It seems that all these proteins have higher affinity for polyunsaturated FAs compared to those belonging to the I-FABP group, whose affinity correlates with the water solubility of different FAs

and, hence, decreases as the level of unsaturation increases (*14*).

The equilibrium dissociation constant of H-, B-, M-FABP is in the range 4−7 nM for OLA and 18−27 nM for ACD (14). In general our estimates fall within this range, although the affinity of Sm14 for ACD (K_d = 10 nM) remains the highest of the group. Interestingly, Eg-FABP (expressed by another parasite lacking de novo synthesis of FAs) shows a similar trend in affinities versus OLA, LA, and ACD (*40*). An important difference between Sm14 and the FABPs belonging to the heart group is that the affinity for FAs is maximal at acidic pH, whereas in the group of FABPs from mammals it is greater at physiological pH.

Several structures of FABPs are deposited in the PDB, most of them bound to FAs. The finding that FABPs (including Sm14) can be crystallized almost only in the liganded state implies that binding of FAs stabilizes the protein. This is now understandable by looking at the structures of Sm14 complexes with OLA and ACD (Figure 1), given that the flap constituted by the two helices is less mobile in the complex and seems to contribute to stabilize a more compact state. Only two structures of FABP complexes with polyunsaturated FAs are deposited in the PDB: human B-FABP (*36*) with docosahexanoic acid (DHA, 22:6) and murine A-FABP with ACD (PDB code: 1ADL (*42*)). The dissociation equilibrium constant measured for ACD binding to the latter protein is reported to be 4.4 μM (by means of calorimetry(*42*)) and 0.2 μM (by the acrylodated intestinal fatty acid binding protein ADIFAB method (*14*)). A comparison of the 3D structures of the complexes of Sm14−ACD and murine A-FABP−ACD (Figure 3, panels B, D) shows that the former is characterized by more specific and stronger contacts with the ligand than the latter, since, in spite of the fact that the relevant amino acid residues are conserved, their orientation is not. This shows the crucial role of the precise stereochemistry of residues in the binding site in determining the free energy of complex formation of ACD, as detailed below.

Binding Specificity of FAs to Sm14. The overall structure of Sm14 is a V-shaped β barrel (Figure 1) closed on the top by a lid of helices and on the bottom by side chains, similarly to other FABPs. The homology model of Sm14 without bound FA, presented by Tendler et al. (*6*), is in overall satisfactory agreement with our structures. Our results extend information on the structure and provide the details necessary to account for ligand affinity and specificity. Our data on the affinity of Sm14 for several FAs suggest that the protein is selectively designed to accommodate unsaturated FAs and, in particular, ACD. ACD is not synthesized by schistosomes (as the worm is incapable of FA biosynthesis), but its derivatives play an essential role in the evasion of the immunological defenses of the host (*16*); therefore, the preference of Sm14 for ACD may confer an evolutionary advantage to the parasite, and inactivation of Sm14 may have therapeutic consequences.

ACD is a priori a relatively difficult ligand for a FABP, since it is more soluble and less concentrated in the host's blood than other important FAs. The rigid structure of ACD demands a precise complementarity with the binding site of Sm14, and implies a minor change in entropy upon binding; on the contrary, OLA and saturated FAs are relatively mobile and binding to the protein cavity diminishes the entropy of their hydrocarbon tail. Comparison between the ACD and OLA complexes is instructive, since OLA retains some mobility inside the pocket and binding is favored by its lower solubility (Table 2); consistently, the average *B* factor of ACD is lower than that of the protein, whereas that of OLA is higher.

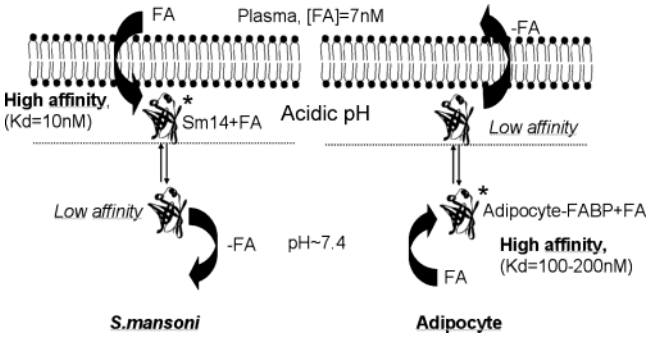


FIGURE 7: A schematic representation of the physiological implications of the pH−affinity dependence of FABPs in *S. mansoni* and in mammalian adipocytes. The scheme depicts the hypothetical principal direction of FA transport in *S. mansoni* tegument (left) and in mammalian adipocyte (right). The main 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 FA trafficking in the two cases.

The crystallographic structure demonstrates a remarkable complementarity between ACD and the binding cavity of Sm14, involving (among others) a strong and highly specific π-cation interaction between the double bond at C8−C9 and the guanidinium group of Arg 78, in turn kept in a favorable stereochemistry by Asp 76. In addition, a specific network of interactions is established between the C5−C6 double bond of ACD and Gln 96, Thr 103, Ile 105, W26, and W39. Finally, the double bonds C5−C6, C8−C9, C11−C12, and C13−C14 are organized around Phe 18 in a optimal manner for π-stacking interactions. Other interactions, including those with the carboxylate head of the FA, are less specific and, although they may contribute to the overall free energy of binding, probably have a minor role in selecting among different FAs.

Structure−Function Relationship and Physiological Role. In view of the high affinity of Sm14 for OLA and ACD, we have addressed the physiologically significant question of the mechanism of release of these FAs. Our experiments demonstrate that the affinity of Sm14 for FAs is higher at acidic pH than at neutral or basic pH, which is opposite to the behavior observed in mammalian FABPs; e.g., A-FABP displays a higher affinity at alkaline pH (*43*), although its isoelectric point is similar to that of Sm14, suggesting a similarity in electrostatic interactions with membranes. Local pH changes inside the cell occur in proximity with the membrane due to the buffering power of the anionic phospolipid headgroups; the microenvironment near the membrane is thus more acidic than the bulk of the cytoplasm (*44*). Since the affinity of Sm14 is higher at lower pH, its main physiological role may be transfer or facilitated diffusion of ACD (and other FAs) from the external tegument membrane (i.e., from the contact interface with the host's blood) to the internal one (i.e., to the worm's body); on the other hand mammalian A-FABP in the adipocytes (whose interior is very rich in FAs) would transport FAs in the opposite direction. This is schematically shown in Figure 7,

where we have highlighted a possible scheme of FA transport in these two cases.

Over and above the pH dependent change in ligand affinity, we have documented that the kinetics are compatible with physiology and that the decrease in affinity at alkaline pH is coupled to a slow ($k \sim 0.1\ s^{-1}$) conformational change of the protein. This unexpected 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. The free FA concentration in human blood is about 7 nM (*45*), compatible with the high affinity of Sm14 for ACD at the acidic pH prevailing in proximity with the cell membrane, whereas in the cytoplasm the release of ACD would be facilitated by the relatively higher pH. If we consider that the schistosome needs to uptake FAs from plasma whereas adipocytes need to discharge them, not only different affinities among different FABP families but also different pH dependences of affinity appear to us relevant to physiology. The reversible association of FABPs with the internal surface of the cell membrane further strengthens our hypothesis and suggests that their isoelectric point is physiologically important. H-FABPs are acidic proteins, L-FABPs are neutral, and M-FABP and murine A-FABP are basic (*10*). Sm14 is, however, unique because, although belonging to the H-FABP family (which is generally acidic), it has a p*I* = 7.8, which is similar to that of M-FABP and A-FABP, promoting the interaction with anionic membranes. Moreover, it has been shown that the uptake of FAs can also occur in close proximity with specific receptors (*11*) and it was recently reported that schistosomes express receptors for human lipoproteins that display acidic properties (*46*). These data support the hypothesis that Sm14 binds FAs from the host's blood and diffuses inside the tegumental cells to enhance their intracellular transport, exploiting the pH gradient between the microenvironment near the membrane and the bulk cytosol, as schematically summarized in Figure 7.

In conclusion, our findings demonstrate how much the lipid affinity, specificity, and pH dependence of FABPs may vary according to subtle differences in the amino acid sequence and especially the details of the 3D structure to suit the needs of the tissue or organism in which they are expressed. Within this range of characteristics, Sm14 is particularly suited to effect the uptake of FAs by the parasite from the host.

ACKNOWLEDGMENT

We thank Dr. Andrea Ilari for several helpful suggestions. We are grateful to Elettra (Trieste, Italy) for data collections and financial support.

REFERENCES

1. 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.
2. Furlong S. T., and Caulfield J. P. (1989) *Schistosoma mansoni*: synthesis and release of phospholipids, lysophospholipids, and neutral lipids by schistosomula, *Exp. Parasitol. 69*, 65−67.
3. McKerrow, J., and Salter, J. (2002) Invasion of skin by schistosoma cercariae, *Trends Parasitol. 18*, 193−195.
4. 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.
5. 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.
6. 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., and 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.
7. Bergquist, N. R., and Colley, D. G. (1998) Schistosomiasis Vaccines: Research to development, *Parasitol. Today 14*, 99−104.
8. 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. W.H.O. 83*, 190−196.
9. Storch, J., and Thumser, A. E. (2000) The fatty acid transport function of fatty acid binding proteins, *Biochim. Biophys. Acta 1486*, 28−44.
10. Veerkamp, J. H, Peeters, R. A., and Maatman, R. G. (1991) Structural and functional features of different types of cytoplasmatic fatty acid-binding proteins, *Biochim. Biophys. Acta 1081*, 1−24.
11. Weisiger, R. A. (2002) Cytosolic fatty acid binding proteins catalyze two distinct steps in intracellular transport of their ligands, *Mol. Cell. Biochem. 239*, 35−42.
12. Kaikaus, R. M., Bass, N. M., and Okner R. K. (1990) Functions of fatty acid binding protein, *Experientia 46*, 617−630
13. 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. 279*, 34481−34488.
14. 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.
15. Rumjanek, F. D., and Simpson, A. J. (1980) The incorporation and utilization of radiolabelled lipids by adult *Schistosoma mansoni*, *Mol. Biochem. Parasitol. 1*, 31−44.
16. Angeli, V., Faveeuw, C., Roye, O., Fontaine, J., Teissier, E., Capron, A., Wolowczuk, I., Capron, M., and Trottein, F. (2001) Role of the parasite prostaglandin D_2 in the inhibition of epidermal Langherans cell migration during schistosomiasis infection, *J. Exp. Med. 193*, 1135−1147.
17. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature 227*, 680−685.
18. Otwinowsky, S., and Minor, W. (1996) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−325.
19. Matthews, B. W. (1968) Solvent content of protein crystals, *J. Mol. Biol. 33*, 491−497.
20. Navaza, J. (2001) AmoRe: an automated package for molecular replacement, *Acta Crystallogr. D57*, 1367−1372.
21. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. D50*, 760−763.
22. 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.
23. McRee, D. E. (1999) XtalView/Xfit: A versatile program for manipulating atomic coordinates and electron density, *J. Struct. Biol. 125*, 156−165.
24. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of the macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. D54*, 1285−1294.
25. Laskowsky, R. A., MacArthur, M. W., Moss, D. S., and Thorntorn, J. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291
26. Liang, J., Edelsbrunner, H., and Woodward, C. (1998) Anatomy of Protein Pockets and Cavities: Measurement of Binding Site Geometry and Implications for Ligand Design, *Protein Sci. 7*, 1884−1897.
27. Kirk, R. W., Kurian, E., and Prendergast, F. G. (1996) Characterization of the sources of protein-ligand affinity: 1-sulfonato-

8-anilinonaphthalene binding to intestinal fatty acid binding protein, *Biophys. J. 70*, 69−83.
28. 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.
29. Sacchettini, J. C., Gordon, J. I., and Banaszak, L. J. (1989) Crystal structure of rat intestinal fatty acid binding protein. Refinement and analysis of the *Escherichia coli*-derived protein with bound palmitate, *J. Mol. Biol. 208*, 327−339.
30. Eads, J., Sacchettini, J. C., Kromminga, A., and Gordon, J. I. (1993) Escherichia 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.
31. 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.
32. 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.
33. 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.
34. 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.
35. 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.
36. 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
37. 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.
38. Kurian, E., Kirk, W. R., and Prendergast, F. G. (1996) Affinity for rRat intestinal fatty acid binding protein: Further examination, *Biochemistry 35*, 3865−3874.
39. Ory, J. J., and Banaszak, L. J. (1999) Studies of ligand binding reaction of adipocyte lipid binding protein using the fluorescent probe 1,8-anilinonaphthalene-8-sulfonate, *Biophys. J. 77*, 1107−1116.
40. Alvite, G., Di Pietro, S. M., Santomè, J. A., Ehrlich, R., and Esteves, A. (2001) Binding properties of *Echinococcus granulosus* fatty acid binding protein, *Biochim. Biophys. Acta 1533*, 293−302.
41. 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.
42. 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.
43. Wootan, M. G., Bernlohr, D. A., and Storch, J. (1993) Mechanism of fluorescent fatty acid transfer from adipocyte fatty acid binding protein to membranes, *Biochemistry 32*, 8622−8627.
44. 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.
45. Richieri, G. V., and Kleinfeld, A. M. (1995) Unbound free fatty acid levels in human serum, *J. Lipid Res. 36*, 229−240.
46. 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.

BI048505F

Structure, Vol. 13, 1241–1246, September, 2005, ©2005 Elsevier Ltd All rights reserved. DOI 10.1016/j.str.2005.06.007

Insights into the Catalytic Mechanism of Glutathione S-Transferase: The Lesson from *Schistosoma haematobium*

Minireview

Francesco Angelucci,[1] Paola Baiocco,[1] Maurizio Brunori,[1] Louise Gourlay,[1] Veronica Morea,[1] and Andrea Bellelli[1,*]
Istituto Pasteur-Fondazione Cenci-Bolognetti and Istituto di Biologia e Patologia Molecolari del CNR
Department of Biochemical Sciences
"A. Rossi Fanelli"
University of Rome "La Sapienza"
00185 Rome
Italy

Summary

Glutathione S-transferases (GSTs) are involved in detoxification of xenobiotic compounds and in the biosynthesis of important metabolites. All GSTs activate glutathione (GSH) to GS^-; in many GSTs, this is accomplished by a Tyr at H-bonding distance from the sulfur of GSH. The high-resolution structure of GST from *Schistosoma haematobium* revealed that the catalytic Tyr occupies two alternative positions, one external, involving a π-cation interaction with the conserved Arg21, and the other inside the GSH binding site. The interaction with Arg21 lowers the pK_a of the catalytic Tyr10, as required for catalysis. Examination of several other GST structures revealed the presence of an external pocket that may accommodate the catalytic Tyr, and suggested that the change in conformation and acidic properties of the catalytic Tyr may be shared by other GSTs. Arginine and two other residues of the external pocket constitute a conserved structural motif, clearly identified by sequence comparison.

Introduction

Glutathione S-transferases (GSTs) are promiscuous enzymes that catalyze several reactions with various substrates. Their essential function is detoxification of xenobiotic compounds that are conjugated to glutathione and then excreted; other functions, not associated with detoxification, include repair of macromolecules oxidized by reactive oxygen species, regeneration of S-thiolated proteins, and biosynthesis of physiologically important metabolites (Armstrong, 1997; Sheehan et al., 2001). GSTs are classified on the basis of their fold, their thermodynamic and kinetic properties, the nature of the residues involved in catalysis, and the specific reaction catalyzed (Sheehan et al., 2001; Wilce and Parker, 1994). In general, GSTs catalyze the nucleophilic attack of the activated thiolate of GSH to electrophilic substrates, and an important step of the catalytic cycle is the activation of GSH to GS^-. However, the precise mechanism of this reaction is still incompletely understood. In the GSTs belonging to the classes alpha, mu, pi, and sigma, glutathione activation proceeds via the interaction with a Tyr at H-bonding distance from the sulfur of GSH; in the enzymes from other classes, the catalytic residue is either Cys or Ser (Armstrong, 1997; Sheehan et al., 2001).

*Correspondence: andrea.bellelli@uniroma1.it
[1]Lab address: http://w3.uniroma1.it/bio_chem/homein.html

Results and Discussion

The *Schistosoma haematobium* GST

Johnson et al. (2003) solved, at 1.65 Å resolution, the 3D structure of the GST from *Schistosoma haematobium* (Sh28GST), a sigma class enzyme active toward several classical substrates and also involved in the biosynthesis of prostaglandin D_2 (PGD_2) from prostaglandin H_2 (PGH_2) (Hervé et al., 2003). PGD_2 is of vital importance to the parasite, being involved in the inhibition of the host immune response and thereby penetration into the host's tissues.

Starting from analysis of this structure, a novel mechanism for GSH activation, which may be relevant to the search for new drugs against schistosomiasis and may possibly be extended to other GSTs, is presented. The structure of Sh28GST (Johnson et al., 2003) shows that the catalytic Tyr occupies two alternative positions, one pointing toward the catalytic site ($Tyr^{in}10$) and contacting the sulfur atom of GSH (as canonical for GSTs) and the other pointing outside ($Tyr^{out}10$) with the phenolic oxygen exposed to water (Figure 1). $Tyr^{in}10$ is stabilized by contacts ($\leq$4.6 Å) with 10 amino acid residues plus GSH, if present, and $Tyr^{out}10$ with 11 different residues, including a remarkable polar interaction with Arg21. The structure of Sh28GST incubated with an excess of GSH shows that the population of $Tyr^{in}10$ is somewhat increased, but never approaches 100%. Although the Tyr^{out} conformer has not been described, examination of other GST structures reveals the existence of a cavity in the position occupied by Tyr^{out} in Sh28GST (Table 1); in the published structures this cavity contains excess electron density, interpreted as water molecules. The cavity is delimited by the amino-terminal region of the polypeptide chain, which forms three β strands and an α helix containing Arg21; this helix packs against the last α helix and is relatively close to the carboxyl terminus.

To investigate whether the peculiar Tyr^{out} conformer seen in Sh28GST is uniquely due to the presence of tyrosine, we solved the structure of the site-directed mutant Tyr10-Phe (Y10F-Sh28GST) expressed in *Escherichia coli* (kindly provided by M. Hervé, J. Fontaine, and F. Trottein, Lille, France). The interaction between the aromatic ring of Phe10 and Arg21 is clearly evident in the 2.0 Å structure of the Y10F mutant, showing that the conformer $Phe^{out}10$ is substantially populated despite exposure of the nonpolar Phe to water. The GSH conjugation and the prostaglandin D_2 synthase activities of this mutant are of course very low to absent (M. Hervé et al., personal communication), confirming the crucial role of Tyr10 in catalysis.

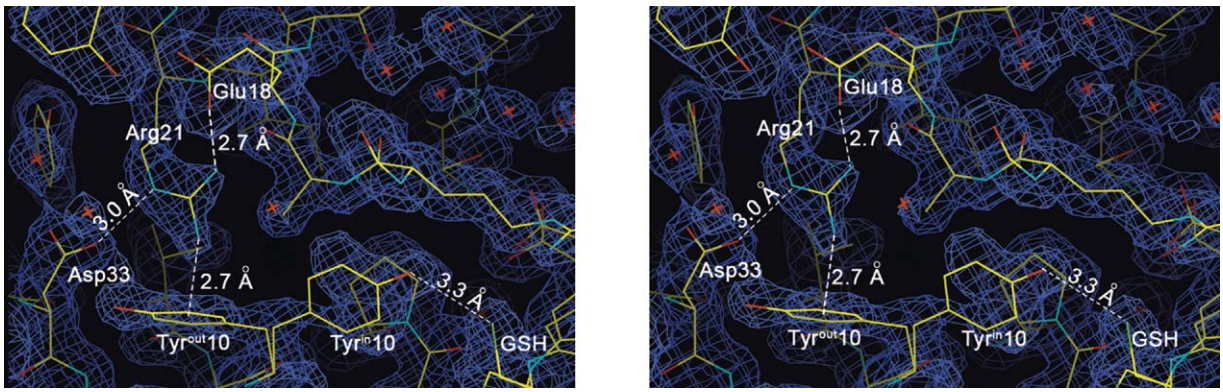


Figure 1. Stereo View of the Electron Density ($2F_o - F_c$ at 1σ; Blue Contours) Showing the Double Conformation of Tyr10 in Sh28GST

Tyr^{in}10 contacts the sulfur atom of GSH by H bonding (3.3 Å). Tyr^{out}10, exposed to water, is stabilized by the π-cation (2.7 Å) interaction with the conserved Arg21, which is held in the proper stereochemistry by Asp33 and Glu18.

A New Functional Role for π-Cation Interaction in Enzymes

Electrostatic interactions between electropositive groups and the electron-rich π-clouds of aromatic rings are not uncommon in proteins and are thought to contribute significantly to the stability of folded proteins (Gallivan and Dougherty, 1999). However, only in rare instances have such interactions been implicated in catalysis; for example, π-cation interactions were shown to be important in substrate and/or transition state recognition at the active site of some enzymes (Zacharias and Dougherty, 2002). Two different geometries are possible: the more common stacked arrangement and the T-shaped (or on-face H-bond) one. Ab initio calculations indicate that, in the gas phase, interaction of the guanidinium group with benzene is stronger in the T-shaped geometry ($\Delta G° = -10.6$ kcal/mol) with respect to the stacked configuration ($\Delta G° = -4$ kcal/mol) (Gallivan and Dougherty, 1999). However, in proteins and when the cation is an amino acid residue containing an sp^2-hybridized nitrogen atom, the stacked geometry is more frequent than the perpendicular H-bond geometry by a 25:1 ratio (Ma and Dougherty, 1997).

In Sh28GST, we observed that the out-conformer is populated not only in the wild-type but also in the Y10F site-directed mutant (see above). The geometry is always T-shaped due to a network of interactions of Arg21 (with Asp33 and Glu18) that forces the NH1 to be perpendicular to the aromatic ring of Tyr10 or Phe10 in the external configuration (Figure 1). In wild-type Sh28GST and in Y10F-Sh28GST, the distance between NH1 of Arg21 and the centroid of the aromatic ring is ~2.7 Å, consistent with a strong π-cation interaction between the NH1 and the π-electron cloud. Interaction with the positively charged Arg side chain is expected to favor ionization of Tyr by 2–3 kcal/mol, thus lowering its pK_a by at least 2 pH units.

In order to test the role of the double conformation of the catalytic Tyr in Sh28GST, the wild-type enzyme and its site-directed mutant were titrated spectroscopically in the pH range 6–11, in the absence of GSH (Figure 2). The UV difference spectrum is diagnostic of the ionization of the phenol ring of Tyr residues and has been compared with that of phenol. The wild-type enzyme displays a distinct transition around neutrality ($pK_a = 7.2$) whose amplitude is compatible with one Tyr residue per GST monomer, and a second one in the alkaline range ($pK_a \geq 9.6$). Y10F-Sh28GST displays a single spectroscopic transition ($pK_a \approx 9.6$). These experiments imply that the catalytic Tyr of Sh28GST can be confidently assigned a $pK_a = 7.2$ in the absence of GSH. Because this low pK_a should result from the weighted average of Tyr^{in} and Tyr^{out}, the pK_a of the latter must be somewhat lower, although we are unable

Table 1. Volume and Area of the Tyr^{out}10 Cavity in Sh28GST and the Homologous Cavities for Two Other GSTs

Protein (PDB Code)	N_mth[a]	Area[b]	Volume[c]	Reference
S. haematobium (1OE7)	1	100.16	86.48	Johnson et al. (2003)
H. sapiens GST P1-1 (16GS)	1	101.00	87.37	Oakley et al. (1998)
H. sapiens GST A1-1 (1GUH)	0	76.69	49.93	Sinning et al. (1993)
O. sloani (2GSQ)	2	110.41	68.34	Ji et al. (1996)
H. sapiens GST M2-2 (1HNA)	0	75.53	49.50	Raghunathan et al. (1994)

Volume and internal surface area of the Tyr^{out}10 cavity of Sh28GST and the homologous water-filled cavities of representative GSTs from various classes as calculated with the program Cast-p using a probe radius of 1.0 Å (Liang et al., 1998). The pockets are selected considering the presence of the NH_2 of the conserved Arg and the C-α and C-β of the catalytic Tyr.

[a] N_mth: The number of mouth openings for the pocket.

[b] Area: The molecular surface area of the cavity in Å^2.

[c] Volume: The volume enclosed in the molecular surface area of the cavity in Å^3. The volume of the pocket for Sh28GST (PDB code 1OE7) was calculated removing the Tyr^{out}10 conformer.

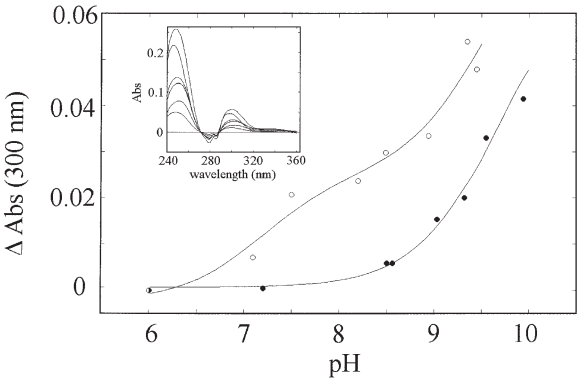


Figure 2. Spectroscopic Acid-Base Titration of Sh28GST and Its Site-Directed Mutant Y10F

Wild-type Sh28GST or Y10F mutant were diluted to a final concentration of 30 μM (per monomer) in a buffer solution containing 50 mM sodium phosphate plus 50 mM boric acid adjusted to the desired pH with NaOH. Absolute absorbance spectra were recorded on a Hewlett Packard 8453 spectrophotometer (Palo Alto, CA), and the difference spectra were calculated using the spectrum recorded at pH = 6.0 as a reference (see inset for the difference spectra of wild-type Sh28GST). The differential absorbance readings at 300 nm are plotted as a function of pH for wild-type Sh28GST (open circles) and the Y10F mutant (closed circles); lines are drawn according to the least squares fit of the experimental data to one (Y10F) or two (wild-type) transitions assigned to the ionizable groups with the pK_as reported in the text.

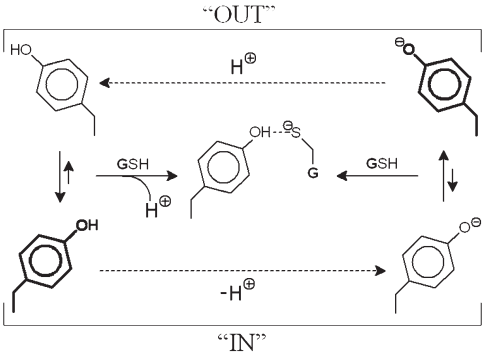


Figure 3. Dual Pathway for GSH Activation

The species in the center represents the activated GS^--enzyme complex. GSH can bind to both conformers of Tyr10 ("IN" and "OUT"), assumed to be in equilibrium. $Tyr^{in}10$ (in bold) is predominantly protonated and works as an H bond donor. $Tyr^{out}10$ (in bold) is mostly ionized and, when GSH binds, it swings in, working as a base. The preferential pathway is dictated by the apparent pK_a of the catalytic residue.

to provide an estimate. Interaction with Arg21 is suggested to be important in promoting the ionization of Tyr^{out}, but is not the only contact stabilizing this conformer.

The Mechanism of GSH Activation in Sh28GST and Its Possible Implications in the GST Superfamily

The active site Tyr of GSTs often has an unusually low pK_a (i.e., <8.5) in the absence of GSH (Ibarra et al., 2001, 2003; Dietze et al., 1996a) compared to that in solution (~10). Tentative explanations of this datum include the on-face H bond with Thr or other Tyr residues and the π-π interactions with neighboring Phe; however, these interactions only partially explain the pK_a shift (Xiao et al., 1996; Dietze et al., 1996b; Ibarra et al., 2001). In our case, the role of the π-cation interaction is expected to be similar to, but more efficient than, that of other interactions of the catalytic Tyr, as described for other GSTs. Additional evidence on the ionization state of the GST-GSH complex is as follows: (1) both the ionized and protonated states of the catalytic Tyr are detected (Dietze et al., 1996a; Ibarra et al., 2003); (2) the protonation of Tyr is dependent on the presence of GSH (Ibarra et al., 2001; Bjornestedt et al., 1995); (3) the pK_a of the bound GSH is ~6.8 (Graminsky et al., 1989) (pK_a in solution is ~9.8); and (4) the catalytic activity (k_{cat}/K_m) decreases when the pH is raised and thus the active form of GSH is unlikely to be deprotonated (Bjornestedt et al., 1995; Kolm et al., 1992).

Assuming that the two conformers are in equilibrium, we propose the hypothetical mechanism depicted in Figure 3. This suggests that Tyr10 switches alternatively toward the solvent where the π-cation bond with Arg21 lowers its pK_a and forces proton release, or toward the active site where the tyrosinate may extract the GSH proton, because of its higher pK_a. This mechanism, whereby Tyr10 works as a proton shuttle, changing its acidic properties through interaction with Arg21, to our knowledge has not been reported before.

This hypothesis may help to describe the binding mechanism of GSH in GSTs belonging to the alpha and pi classes, that is, whether GSH preferentially binds to the protonated or unprotonated GST (see Equations 2 and 3 of Armstrong, 1997). Indeed, if GSH binds with different rate constants to the Tyr^{in} and Tyr^{out} conformers of the enzyme, either path may be more prevalent than one would expect on the basis of the average pK_a of the catalytic residue. Even GSTs with much higher pK_a of the catalytic Tyr and a population of the Tyr^{out} conformer lower than in Sh28GST might still preferentially bind GSH through this conformer (Figure 3), in view of the greater accessibility of the binding site. This would mean that upon rapid mixing with substrate, only the protons bound to the fraction of enzyme presenting the protonated Tyr^{in} conformation would be released. The apparent inconsistency with the observation that proton release is synchronous with GSH binding and thiolate formation (Caccuri et al., 1999) is solved because those experiments were carried out at a pH lower than the pK_a of the catalytic Tyr, and thus even the Tyr^{out} conformation could have been partially protonated; moreover, the enzymes used may have a relatively low population of Tyr^{out}, in view of the relatively high pK_a of their catalytic Tyr.

A different matter is the mechanism of activation of bound GSH, which depends on the precise position of the unique proton shared by Tyr^{in} and glutathione, that is, whether the ES Michaelis complex is more akin to GS thiolate-tyrosine or to GSH-tyrosinate (Equations 4 and 5 of Armstrong, 1997). On the basis of the isotope effect in D_2O, Armstrong and coworkers (Armstrong,

1997; Parsons and Armstrong, 1996) reported that the GS thiolate-tyrosine-type Michaelis complex is prevalent in the rat mu class GST (whose catalytic Tyr has a pK_a of ~10), although the enzyme may switch to a GSH-tyrosinate-type complex if the catalytic Tyr is fluorinated (a modification that lowers the pK_a to 7.5); thus, the physiological mechanism for GSH deprotonation would be the former. Our data do not impinge directly on this point, which critically depends on the pK_a of the Tyr^{in} conformer. However, the reactions given in Figure 3 imply that this pK_a may be significantly higher than previously expected, given that the low overall pK_a of the catalytic Tyr is assigned to the weighted average of the two conformers. Thus, if anything, our scheme would suggest that GSH preferentially binds to the deprotonated Tyr^{out} conformer, but the only proton of the ES complex is completely transferred to the Tyr^{in}, consistent with isotope effects experiments. The main reaction path would be:

$$E\_Tyr^{out}\text{-}O^- + GSH \rightarrow E\_Tyr^{out}\text{-}O^- \ldots HSG \rightarrow E\_Tyr^{in}\text{-}O^- — HSG \rightarrow E\_Tyr^{in}\text{-}OH — {}^-SG.$$

The mechanism proposed in Figure 3 qualitatively accounts for the experimental data on alpha and pi class GSTs, and is consistent with the thermodynamics of the process, assuming that $Tyr^{in}10$ has the usual pK_a of 9–10, whereas $Tyr^{out}10$ has a pK_a close to 7. According to our scheme, the two conformers $Tyr^{in}10$ and $Tyr^{out}10$ are both populated at equilibrium, even in the absence of GSH; at physiological pH Tyr^{in} is protonated, whereas Tyr^{out} is significantly unprotonated. As far as our hypothesis can be generalized, the different pK_a values reported for the catalytic Tyr, ranging between pK_a = 10 for Mu1-1 GST from rat (Xiao et al., 1996) and pK_a = 6.7 for A4-4 GST from human (Hubatsch and Mannervik, 2001), would merely reflect the equilibrium population of the two conformers that strongly favors Tyr^{in} in the former case and Tyr^{out} in the latter.

In our hypothesis, GSH may combine with both conformers and favor the translocation of the catalytic Tyr to the active site. If GSH binds to an enzyme presenting the Tyr^{in} conformation, this residue is likely to be protonated and the complex will behave like a weak diprotic acid. In this case Tyr donates an H bond, favoring the release of the hydrogen ion from GSH. If, on the other hand, GSH binds to an enzyme presenting the Tyr^{out} conformer, it promotes the switch of the tyrosinate toward the catalytic site and extraction of the GSH proton. Because experimental evidence is consistent with both mechanisms, we conclude that they are both possible, and either one may appear to prevail depending on the fraction of Tyr occupying either of the two alternative positions. Our hypothesis does not contradict any of the relevant findings, and actually reconciles many of them; in particular: (1) it explains the low apparent pK_a of the catalytic Tyr, as well as that of bound GSH; (2) it allows both protonated and unprotonated Tyr to be catalytically competent; and (3) it is not inconsistent with the decrease of the catalytic activity with increasing pH because this would stabilize the Tyr^{out} conformer and would preclude one of the two possible reaction pathways.

Evolutionary Conservation of Arg and Its Interaction Network

Prompted by the hypothesis that the π-cation bond between $Tyr^{out}10$ and Arg21 is crucial, we further investigated the evolutionary conservation of the latter residue and its interaction network. Two hundred forty-three sequences of GSTs from different organisms are listed in the SwissProt database (Bairoch et al., 2005); because the sequence and the 3D structure of all GSTs are conserved, especially in the amino-terminal region (Sheehan et al., 2001), a comparison is justified. Of the 243 sequences, 95 are assigned to the classes alpha, mu, pi, and sigma, having Tyr as the catalytic residue (we call this group set #1); 83 to the classes beta, kappa, omega, phi, tau, theta, and zeta, having either Cys or Ser as the catalytic residue (set #2); and 65 are not assigned to any class (set #3). The analysis shows that 94% of the GSTs of set #1 (89 out of 95) present a characteristic structural motif constituted by the catalytic Tyr at position *n* followed by Glu, Gln, or His at *n + 8*, by Arg at *n + 11*, and by Asp or Glu at *n + 22* or *n + 23*, that is, ([Y] 7X [E,Q,H] 2X [R] 10-11X [E,D]). In the alpha class human GST, the functional relevance of the interaction between Arg19 and Glu31 (topologically homologous to Arg21 and Asp33 in Sh28GST) has been substantiated by mutagenesis (Lee et al., 1995), but has not been explained. The conservation of Arg at *n + 11* and its interaction network is specific for Tyr-containing GSTs, as demonstrated by the fact that only 6% of GSTs of set #2 (5 out of 83) present the extended structural motif ([Y,C,S] 7X [E,Q,H] 2X [R] 10-11X [E,D]).

The residues of the motif are identified by their putative functional and structural roles: that is, the catalytic Tyr, the Arg residue required to establish the π-cation bond, and the two residues that can interact with the Arg and stabilize its correct orientation. Their evolutionary conservation is an independent, though indirect, confirmation of their putative role. Because these four residues are located on two β strands and an α helix, their stereochemical proximity depends on the correct folding of the amino-terminal region of the protein; the short coil regions between these secondary structure elements might allow insertions or deletions between the Arg and the Glu or Asp. The structural organization of these residues is strongly conserved, as evident from comparison of human alpha class GST A1-1 (Cameron et al., 1995; Protein Data Bank (PDB) code 1GSD), sigma class GST from squid (Ji et al., 1996; PDB code 2GSQ), and human pi class GSTP1-1 (Oakley et al., 1998; PDB codes 16GSA and 16GSB), all reported in Figure 4.

As an unbiased empirical control of the statistical frequency of the ([Y] 7X [E,Q,H] 2X [R] 10-11X [E,D]) motif identified in GSTs containing a Tyr in the active site, we searched the whole SwissProt database which lists 178,171 protein sequences. The motif was found in 3750 (2.1%) sequences, a value that does not seem unreasonably high, in view of the relative frequency of these amino acids (Klapper, 1977).

It is interesting to analyze the six Tyr-containing GSTs of set #1 that do not contain the ([Y] 7X [E,Q,H] 2X [R] 10-11X [E,D]) motif. Four of the exceptions have a substitution at *n + 8*: the mu class GSTs from rabbit (*Oryctolagus cuniculus*; SwissProt code P46409) and

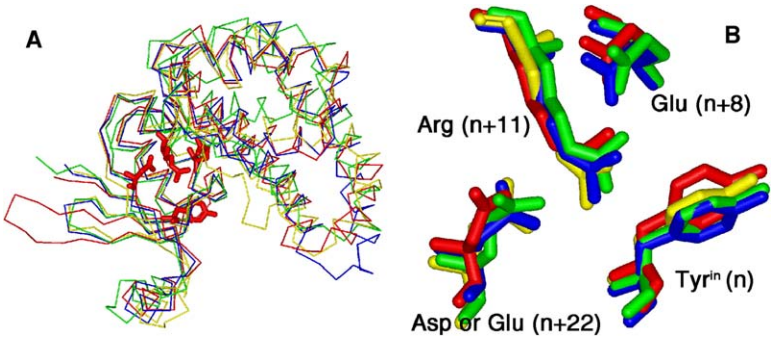


Figure 4. Comparison of the Structure of Representative GSTs

(A) Superimposition of the C-α traces of Sh28GST (red), α class GST A1-1 (green), sigma class GST (blue), and pi class GST P1-1 (yellow); PDB codes are 1OE7, 1GSD, 2GSQ, and 16GS, respectively. Tyr10 ($Tyr^{in}10$), Glu18, Arg21, and Asp33 in Sh28GST are highlighted in red.

(B) Details of the superimposition shown in (A). Arg is kept in place by interaction with two residues of the same polypeptide chain, ready to stabilize the possible Tyr^{out} conformer (see text).

Y1 from hamster (*Cricetulus longicaudatus*; SwissProt code Q00285) have Leu and Asn at *n + 8*, respectively, whereas human GSTP1-1 (SwissProt code P09211) and GSTP1 from macaque (*Macaca mulatta*; SwissProt code Q28114) both have Ala. All other residues of the motif are conserved. The last two GSTs have a full, but probably functionally silent, ([Y] 7X [E,Q,H] 2X [R] 10-11X [E,D]) motif at significant distance from the amino terminus.

The amino-terminal residues of the sigma class GST2 from *Manduca sexta* (SwissProt code P46429) could not be aligned satisfactorily with the other GSTs. Our best guess is that the catalytic residue is Tyr8 and, if this enzyme has the external pocket at all, it is organized around Arg16 (*n + 8*) and Glu27 (*n + 19*), respectively; this could be the consequence of a three-residue deletion between positions 8 and 16, also involving one of the residues that were expected to interact with Arg16. The sigma class GST2 from *Caenorhabditis elegans* (SwissProt code O16115) presents the extended ([Y,C,S] 7X [E,Q,H] 2X [R] 10-11X [E,D]) motif and aligns nicely with the other sequences, but has a Cys in the putative catalytic position instead of a Tyr; perhaps it is misclassified.

The sequence conservation of the amino-terminal region of Tyr-containing GSTs is more extended than described above, and several other residues appear to be conserved in the region between the amino terminus and the first 25 residues or more downstream of the catalytic Tyr; for example, the second residue before the catalytic Tyr (***n – 2***) is always large and nonpolar, and the Glu or Asp at *n + 22* or *n + 23* is very often followed by a second Glu or Asp. A search for the motif ([Y] 7X [E,Q,H] 2X [R] 9-12X [E,D][E,D]) in set #1 found 85 GSTs out of 95 and only 875 proteins in the whole SwissProt database (i.e., 0.5%). However, we do not have a clear-cut functional explanation for these other residues in the catalytic mechanism of Figure 3.

Conclusions

As far as we know, there is no other example of a protein in which a Tyr in a π-cation interaction with an sp^2 nitrogen of an amino acid changes its acidic properties with functional consequences. Only in artificial "receptors" has it been demonstrated that electron-rich phenolate could bind quaternary ammonium ions more tightly than neutral ones (Ma and Dougherty, 1997). Sh28GST may be the first example of a new role for the π-cation interactions, although Nurizzo et al. (2001) noticed that a similar motif may be of significance in other proteins.

Although we cannot extend our mechanism to all Tyr-containing GSTs, we suggest that it may at least be considered when the three following conditions apply: (1) the pK_a of the catalytic Tyr is lower than the usual value of Tyr residues by more than 1 pH unit; (2) a water-filled external pocket is present; and (3) the pocket hosts Arg ***n*** **+ 11** and its interaction network. Conditions (2) and (3) can be easily verified by computer modeling if the 3D structure of the enzyme is available.

The in and out movement of Tyr10 may be relevant to the search for GST inhibitors. This is intriguing because the GSTs are detoxifying enzymes characterized by promiscuous substrate specificity, which makes it difficult to design class-specific drugs. If the swing off of the active site Tyr is involved in the detoxifying activity, as we propose, it may be profitable to orient drug search toward the pocket in which the Tyr is held by the interaction with the conserved Arg. On the other hand, if this mechanism were unique to Sh28GST, it may be a specific clue to inhibit the schistosomial enzyme with possible therapeutic implications.

Acknowledgments

We are very grateful to Ken A. Johnson and Adriana E. Miele for helpful suggestions, and to François Trottein, Maxime Hervé, and Josette Fontaine of the Institut Pasteur of Lille, Lille, France for providing the Y10F GST mutant. P.B. is supported by Elettra (Trieste, Italy). This work was partially supported by Progetto di Ateneo (University "La Sapienza," Rome, Italy) and Consorzio Interuniversitario Biotecnologie (Urbino, Italy).

Received: April 23, 2005
Revised: June 19, 2005
Accepted: June 20, 2005
Published: September 13, 2005

References

Armstrong, R.N. (1997). Structure, catalytic mechanism, and evolution of the glutathione S-transferase. Chem. Res. Toxicol. *10*, 2–18.

Bairoch, A., Apweiler, R., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., et al. (2005). The Universal Protein Resource (UniProt). Nucleic Acids Res. *33*, D154–D159.

Bjornestedt, R., Stenberg, G., Widersten, M., Board, P.G., Sinning, I., Jones, T.A., and Mannervik, B. (1995). Functional significance of

arginine 15 in the active site of human class α glutathione transferase A1-1. J. Mol. Biol. *247*, 765–773.

Caccuri, A.M., Antonini, G., Board, P.G., Parker, M.W., Nicotra, M., Lo Bello, M., Federici, G., and Ricci, G. (1999). Proton release on binding of glutathione to α, mu and delta class glutathione transferases. Biochem. J. *344*, 419–425.

Cameron, A.D., Sinning, I., L'Hermite, G., Olin, B., Board, P.G., Mannervik, B., and Jones, T.A. (1995). Structural analysis of human α-class glutathione transferase A1-1 in the apo-form and in complexes with ethacrynic acid and its glutathione conjugate. Structure *3*, 717–727.

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

Dietze, E.C., Ibarra,, C., Dabrowski,, M.J., Bird,, A., and Atkins, W.M. (1996b). Rational modulation of the catalytic activity of A1-1 glutathione S-transferase: evidence for incorporation of an on-face (pi…HO-Ar) hydrogen bond at tyrosine 9. Biochemistry *35*, 11938–11944.

Gallivan, J.P., and Dougherty, D.A. (1999). Cation-π interactions in structural biology. Proc. Natl. Acad. Sci. USA *96*, 9459–9464.

Graminsky, G.F., Kubo, Y., and Armstrong, R.N. (1989). Spectroscopic and kinetic evidence for the thiolate anion of glutathione at the active site of glutathione S-transferase. Biochemistry *28*, 3562–3568.

Hervé, M., Angeli, V., Pinzar, E., Wintjens, R., Faveeuw, C., Narumiya, S., Capron, A., Urade, Y., Capron, M., Riveau, G., and Trottein, F. (2003). Pivotal roles of the parasite PGD_2 synthase and of the host D prostanoid receptor 1 in schistosome immune evasion. Eur. J. Immunol. *33*, 2764–2772.

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 pK(a) of tyrosine-9 and the C-terminal dynamics of glutathione S-transferase A1-1. Biochemistry *40*, 10614–10624.

Ibarra, C.A., Chowdhury, P., Petrich, J.W., and Atkins, W.M. (2003). The anomalous pKa of Tyr-9 in glutathione S-transferase A1-1 catalyzes product release. J. Biol. Chem. *278*, 19257–19265.

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. USA *93*, 8208–8213.

Johnson, K.A., Angelucci, F., Bellelli, A., Hervé, M., Fontaine, J., Tsernoglou, D., Capron, A., Trottein, F., and Brunori, M. (2003). Crystal structure of the 28 kDa glutathione S-transferase from *Schistosoma haematobium*. Biochemistry *42*, 10084–10094.

Klapper, M.H. (1977). The independent distribution of amino acid near neighbor pairs into polypeptides. Biochem. Biophys. Res. Commun. *78*, 1018–1024.

Kolm, R.H., Sroga, G.E., and Mannervik, B. (1992). Participation of the phenolic hydroxyl group of Tyr-8 in the catalytic mechanism of human glutathione transferase P1-1. Biochem. J. *285*, 537–540.

Lee, H.C., Toung, Y.P., Tu, Y.S., and Tu, C.P. (1995). A molecular genetic approach for the identification of essential residues in human glutathione S-transferase function in *Escherichia coli*. J. Biol. Chem. *270*, 99–109.

Liang, J., Herbert, E., and Woodward, C. (1998). Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. Protein Sci. *7*, 1884–1897.

Ma, J.C., and Dougherty, D.A. (1997). The cation-p interaction. Chem. Rev. *97*, 1303–1324.

Nurizzo, D., Baker, H.M., He, Q.Y., MacGillivray, R.T., Mason, A.B., Woodworth, R.C., and Baker, E.N. (2001). Crystal structures and iron release properties of mutants (K206A and K296A) that abolish the dilysine interaction in the N-lobe of human transferrin. Biochemistry *40*, 1616–1623.

Oakley, A.J., Lo Bello, M., Ricci, G., Federici, G., and Parker, M.W. (1998). Evidence for an induced-fit mechanism operating in pi class glutathione transferases. Biochemistry *37*, 9912–9917.

Parsons, J.F., and Armstrong, R.N. (1996). Proton configuration in the ground state and transition state of a glutathione transferase-catalyzed reaction inferred from the properties of the tetradeca(3-fluorotyrosil)glutathione transferase. J. Am. Chem. Soc. *118*, 2295–2296.

Raghunathan, S., Chandross, R.J., Kretsinger, R.H., Allison, T.J., Penington, C.J., and Rule, G.S. (1994). Crystal structure of human class mu glutathione transferase GSTM2-2. Effects of lattice packing on conformational heterogeneity. J. Mol. Biol. *238*, 815–832.

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

Sinning, I., Kleywegt, G.J., Cowan, S.W., Reinemer, P., Dirr, H.W., Huber, R., Gilliland, G.L., Armstrong, R.N., Ji, X., Board, P.G., et al. (1993). Structure determination and refinement of human α class glutathione transferase A1-1, and a comparison with the mu and pi class enzymes. J. Mol. Biol. *232*, 192–212.

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

Xiao, G., Liu, S., Ji, X., Johnson, W.W., Chen, J., Parsons, J.F., Stevens, W.J., Gilliland, G.L., and Armstrong, R.N. (1996). First-sphere and second-sphere electrostatic effects in the active site of a class mu glutathione transferase. Biochemistry *35*, 4753–4765.

Zacharias, N., and Dougherty, D.A. (2002). Cation-π interactions in ligand recognition and catalysis. Trends Pharmacol. Sci. *23*, 281–287.