CONCLUDING DISCUSSION

"Nothing in biology makes sense except in the light of evolution"

Theodosius Dobzhansky, 1973

The current ideas on tick evolution and their adaptation to a blood-feeding environment were outlined in the introduction to this thesis. Ticks adapted to an efficient host hemostatic system by evolving numerous new protein functions capable of regulating this complex system. While many tick salivary gland proteins have been described, knowledge on the evolution of tick protein families is lacking. Evolution of novel protein function entails gene duplication and subsequent gain or loss of protein function. This implies that proteins with different functions but common ancestors must still share fundamental properties. In this study tick evolution was approached from the perspective of gene duplication and subsequent gain/loss of protein function. This thesis presents the first comprehensive analysis of tick protein families in an evolutionary context.

Two main protein families have been studied, a family of BPTI-like anti-hemostatic factors (Chapter 2-4) and a family of proteins that are part of the lipocalin protein family (Chapter 6-7). Protein family is being defined here as proteins that share a common structural protein fold, although their functions might differ, i.e. paralogous proteins.

The BPTI-like family consists of at least three proteins with different functions. Phylogenetic analysis clearly showed that these proteins are monophyletic, which imply common ancestry. Common ancestry in this sense can only mean that there were multiple gene duplication events and subsequent mutation that led to a gain or loss of protein function.

Several isoforms of the platelet aggregation inhibitor, savignygrin were identified (Chapter 2). As savignygrin was never described before, extensive characterization on protein and functional level was performed. These showed that savignygrin targets the integrin, $\alpha_{IIb}\beta_3$, possesses an RGD motif and is the ortholog of disagregin (Mans, Louw

and Neitz, 2002b). The first evidence of gene duplication in argasid ticks is also presented. It appears to be a very recent gene duplication event, probably after the divergence of *O. moubata* and *O. savignyi*. An interesting question is why this second copy of savignygrin has not yet been inactivated as expected for duplicated genes (Lynch and Conery, 2000). A recent study has shown that duplicate genes might persist in a lineage if they confer an advantage to the organism (Kondrashov *et al.* 2002). Advantage conferred at this early stage of gene duplication is considered to be due to the presence of higher concentrations of the translated protein from a gene. In the case of the savignygrins, this could be a relevant observation as the (+)/(-) forms each constitute ~6% of the total salivary gland proteins. Compared to the yields obtained for disagregin (0.225% of the total SGE) this seems to be exceptionally high (Karczewski, Endris and Connolly, 1994).

O. moubata has a limited number of hosts and predominantly feeds either on man or warthogs, and inhabits the burrow of its host. *O. savignyi* probably has a wider host range and will feed on any animal that shelters in the shade of trees where it resides (Hoogstraal, 1956). This behaviour of *O. savignyi* may explain the higher concentrations of platelet aggregation inhibitor. A wide host range may expose a tick to a diverse repertoire and concentration range of platelet receptors. It should thus be interesting to determine the number of $\alpha_{IIb}\beta_3$ receptors expressed on the platelet surfaces of different hosts. Expression of high concentrations of platelet aggregation inhibitor would obviously be advantageous for the tick to cover as wide a range of platelet receptor concentrations as might be encountered in different hosts. To test the hypothesis that purifying selection could be responsible for the maintenance of two gene copies for the savignygrins, the number of integrin $\alpha_{IIb}\beta_3$ per platelet should be investigated for different domestic and wild animals. It is predicted that animals considered as hosts for *O. savignyi* should have higher numbers of this integrin than *O. moubata* hosts. If this proves to be correct, it can be speculated that the control of platelet aggregation during feeding of *O. savignyi* is still important. While the use of anti-coagulants has been dismissed as a suitable target for vaccine development, targeting of platelet aggregation inhibitors have never been attempted and can still be viable for vaccine development.

An alternative explanation for the two gene copies of the savignygrins, is that this is a very recent gene duplication event, so that one copy still needs to be inactivated, or change function. A date for this gene duplication event can probably be estimated if the divergence of *O. savignyi* and *O. moubata* is considered. If we assume as upper limit that ticks originated 120 MYA (Klompen *et al.* 1996), then a divergence rate for the platelet aggregation inhibitors can be calculated:

$$r = d/2T$$

Where r = rate of mutation expressed as changes per site per million years, d = total number of amino acid substitutions calculated from: -ln (1-p), where p = d/n (d is the number of amino acid differences between two sequences and n is the total number of amino acid residues in sequences of similar length). T = time of divergence in millions of years (Nei and Kumar, 2000).

This gives a substitution rate of 3.32×10^{-9} per site per year for the platelet aggregation inhibitors. Assuming that this rate stays the same for all members of the platelet aggregation family, then the calculated time of divergence for the +/- isoforms of savignygrin is 4.9 MYA. *O. savignyi* and *O. moubata* have most probably diverged after the origination of ticks. If 92 MYA is taken as the time of divergence (Klompen and Grimaldi, 2001), a substitution rate of 4.43×10^{-9} per site per year can be calculated for the platelet aggregation inhibitors. Using this rate gives a divergence time for the +/- isoforms at 3.7 MYA. If very fast rates of substitution are assumed (as exemplified by the positive selection observed for the platelet aggregation inhibitors) then the fastest realistic rate would probably be 1×10^{-8} per site per year (Grauer and Li, 2000). Using this gives a time of divergence between *O. savignyi* and *O. moubata* at ~40 MYA and can probably be taken as the lower limit for time of divergence between these two species. It also gives a time of divergence for the +/- forms at 1.6 MYA.

The lower limit is still close to what is considered as ample time for a duplicated gene to be inactivated (Lynch and Conery, 2000). It could thus be concluded that purifying selection and concerted evolution is probably playing a role in the maintenance of two savignygrin gene copies.

In Chapter 3 the thrombin inhibitor, savignin was characterized on molecular level in order to obtain data for a comprehensive phylogenetic analysis of the tick BPTI-family. Characterization of savignin in terms of its sequence and predicted structure confirmed that it is the ortholog of ornithodorin and that they should have similar mechanisms of inhibition (Mans, Louw and Neitz, 2002a). The results from this study provide a link between the kinetic characterization of savignin (Nienaber, Gaspar and Neitz, 1999) and the structural characterization of ornithodorin, for which no kinetic data were provided (van de Locht *et al.* 1996). The predictions of the intra-domain interactions of savignin provided insights into the evolutionary significance of domain interaction in terms of the conserved nature of the thrombin inhibitors (Chapter 4).

Chapter 4 described the relationship between the platelet aggregation inhibitors, savignygrin and disagregin and other members of the tick BPTI-family. The platelet aggregation inhibitors are identified as novel members of the BPTI-family and it is shown that the RGD motif of savignygrin is presented on what is known as the substrate-binding presenting loop of the canonical inhibitors (Mans, Louw and Neitz, 2002b). Common ancestry between different tick BPTI-like proteins was also investigated. These studies presented the first in-depth analysis of how tick proteins evolved new functions (Mans, Louw and Neitz, 2002c). The conclusions derived in this chapter have far reaching implications for the evolution of ticks and how they adapted to a blood-feeding environment (see below). Investigations into the validity of these conclusions should provide the tick research community with ample study opportunities that will stimulate a new synthesis of tick evolution.

Chapter 5 provides important information on salivary gland morphology, not previously described for the tick, *O. savignyi*. For the first time data is provided that shows the

presence of more than three granular cell types in argasid ticks. This is of cardinal importance for the classification of the granular cell types in tick salivary glands. These results clearly show that morphological and histochemical approaches need to be supplemented by biochemical and immuno-cytochemical techniques in order to describe granular cell types on a more analytical level. The localization of the same anti-hemostatics to different granular types (typed on the basis of morphology), however, complicates the assignment of definite granular types.

Chapter 6 describes proteins that may be involved in granule biogenesis (Mans *et al.* 2001). Tick salivary gland granule biogenesis has not been described to date and this study is the first step in this direction. The nature of sand tampan toxicoses has also been revisited after a number of years since the first descriptions of this form of toxicoses (Neitz, Howell and Potgieter, 1969; Howell, Neitz and Potgieter, 1975). This culminated in the identification and characterization of a novel basic toxin and the confirmation of the properties of the acidic toxin previously described (Mans *et al.* 2001; Mans *et al.* 2002). Pathogenesis of the cardiac system caused by these toxins is a novel form of tick toxicosis and distinguishes sand tampan toxicoses from paralysis toxins (Mans *et al.* 2002).

The identification of the TSGPs led to the discovery of a whole family of argasid tick lipocalins (Mans, Louw and Neitz, 2002d). There are at least 6 different lipocalin proteins in the *Ornithodoros* species (these include toxins and anti-platelet agents) and three more in hard tick species (histamine binding proteins). Except for moubatin, the TSGPs are the first lipocalins described for argasid tick species. It appears as if these proteins have different functions in the host although a common function in granule biogenesis have been proposed. Phylogenetic analysis also showed that these proteins are monophyletic. This again implies gene duplication events and subsequent gain or loss of protein function through mutation. The conclusions derived for the evolution of the tick toxins show that toxins might be recently acquired traits of the sand tampan. This has important implications for the origin and evolution of other tick toxins. The extensive lipocalin

family further supports the concept that gene duplication played an important role during the evolution of novel protein functions in argasid ticks.

A consideration of the results obtained during this study leads us back to the introduction and the questions that were asked concerning the adaptation of ticks to a blood-feeding environment. We can now consider these questions in the light of the findings presented in this study.

Question: How did ticks acquire novel anti-hemostatic strategies?

This study has confirmed evidence of gene duplication of two protein families. A general mechanism for the acquisition of novel protein functions through the use of a few relatively simple protein folds are proposed. Ticks generated functional diversity, by gene duplication and gain or loss of protein function. This allowed the generation of different anti-hemostatic factors by which ticks could regulate the hemostatic system of their host's. Other unknown functions were also generated, which are exemplified by peripheral toxic activities of some of these proteins such as TSGP2 and TSGP4.

The use of a few common protein folds to generate diversity is a simple but elegant way for an organism to become more complex (complexity is here defined as having more potential protein functions and hence being able to adapt to a wide variety of possible environments). It can be foreseen that new functions could still be generated in time from these same protein folds in response to a changing environment. This same principle probably also holds for other, as yet undiscovered proteins present in the salivary glands.

Question: What was the nature of the ancestral tick before adaptation to a blood-feeding environment?

The question can be posed why a few common protein folds are used over and over again, while ticks should at least have the potential of having many more protein folds in its repertoire. It can be predicted that the tick genome should be between 100-200 Mb and ~15000 genes (Grauer and Li, 2000). Perhaps the answer could be found in the ancestor to the holothyrida and Ixodida sub-orders. Such a mite may be presumed to be a free-

living scavenger and would have needed its salivary glands to make the food source more accessible. It can be foreseen that matrix-degrading proteins, such as hyaluronidase and proteases, such as matrix metalloproteases may have been present, as this would have allowed the mite to enlarge its feeding site (Neitz and Vermeulen, 1987). A range of protease activities has been described in the SGE of *O. savignyi* (Mahlaku, Gaspar and Neitz, 2002). Serine protease inhibitors may have played a role to inhibit any protease activity present in the body fluids of the dead organisms it fed on. The ancestral lipocalins probably functioned in granule biogenesis, as this would have been an intrinsic property of the salivary gland as a secretory organ. Depending on the complexities of its food sources, it could be speculated that the salivary glands at the time of divergence between the holothyrida and ixodida, had already been differentiated organs. This means, that a certain subset of proteins were expressed in these glands and that the number targeted to the secretory pathway would have been limited to those used during their feeding process. Ticks may have originated with a limited set of proteins, from which new diversity had to be generated. This may account for the re-use of an existing protein fold. The state of this primitive salivary gland could also have influenced the protein repertoire of the ancestral tick. Proteins are not only expressed in the glands, but also need the signals necessary for secretion.

Question: Did ticks evolve anti-hemostatic strategies before or after divergence into the main tick families?

It has been suggested that hard and soft ticks adapted independently to a blood-feeding environment (Chapter 4). This was based on anti-hemostatic factors unique to hard and soft ticks. If the main tick families adapted independently to their new hemostatic environments, other features in their biology, that is also dependent on a blood-feeding lifestyle, should corroborate this. The feeding and reproductive strategies of hard and soft ticks differ quite remarkably. Adult female hard ticks feed once, secrete a salivary gland degenerating factor that causes the salivary gland to atrophy and proceeds to lay several thousands of eggs in one stage after which the tick dies. Adult soft ticks on the other hand, feed several times and lay several hundreds of eggs after each feeding period (Sonenshine, 1991). As these strategies are general throughout the Ixodidae and

Argasidae it can be assumed that these traits only evolved after the divergence of hard and soft ticks, probably at a stage when ticks were adapting to a blood-feeding environment.

Question: What was the driving force behind the evolution of hematophagy in ticks?

It was proposed that the divergence of ticks into their different families might have been triggered by the divergence of mammals and birds (Chapter 4). The divergence of mammals and birds could also have been a possible reason for the evolution of hematophagy in ticks. The ancestral holothyrid related tick probably fed on dead arthropods. The emergence of numerous new potential mammalian and avian hosts, with relatively thin epidermises that allows for the penetration and location of a fluid meal could have been a sufficient trigger that allowed a rapid adaptation to this new niche. This is probable when it is considered that ticks belong to the superorder Parasitiformes, which consists of the Holothyrida, Mesostigmata and Ixodida. Many Mesostigmata live as parasites on mammalian and avian hosts from where they feed on secretions obtained by penetrating the skin, causing blood and lymph fluid flow (Radovsky, 1985). An ancestral holothyrid-like mite could have made this transition from dead arthropods, to dead mammals and finally to live mammals, in the process starting to evolve components that can control its host's defense mechanisms.

Implications derived from this study for future research into the "tick genome"

It seems in retrospect that more questions were raised during this study than answered. While gene duplication events have been indicated for two protein families, the question of how many undiscovered members of both protein families exist in different tick species is now more pressing than before. More problematic even is the high level of divergence of these sequences and the indications that many functional proteins might only have acquired novel functions after the divergence of hard and soft ticks. This implies that even if the genome of a hard tick species is sequenced to serve as a model for tick genome studies, we might still miss some of the most important molecules involved in tick feeding from other species. The choice of a model tick or ticks for genome sequencing will remain a central problem. Choosing a tick closest to the ancestral tick (if

we can predict the more ancient lineages correctly) will allow us to study early evolutionary events although we will miss those proteins recently diverged. The most fruitful endeavor would be to sequence the genomes of representatives from both hard and soft ticks, as well as the Nuttallillidae. The latter is probably one of the most important targets, as it will provide us with an external reference point for the other tick families. As the Nuttallillidae is also the only monotypic family and occur only in Gondwanaland (the suggested cradle of tick origins), this tick might be the closest living fossil that we have to the ancestral tick. The scarcity of living specimens of this tick makes it a priority of study for tick biodiversity. If this tick species should be lost or become extinct we might lose one of the most valuable tick species still living for the study of tick evolution.

Future studies on the gene duplication mechanisms in argasid ticks

Gene duplication events for both the BPTI-like inhibitors and lipocalins can now be studied at chromosome level. Questions remaining to be answered are whether the duplicated genes are localized to the same chromosomes, in what order they are arranged on the chromosomes and what patterns of introns/exons they exhibit. Such knowledge has been fruitfully used to study mammalian lipocalin evolution (Salier, 2000). The known gene sequences can be utilized using *in situ* hybridization techniques and genomic libraries to answer these questions. If the genes from individual families are localized on the same chromosome, novel members of the different families could be identified, by limited chromosome sequencing. Studies into this field can allow the preliminary mapping of a tick genome, which could be important for future genome sequencing projects.

Future studies on the TSGPs

The expression of the TSGPs will allow the next step of characterization of these molecules. These include studies into the mechanisms of toxicity and the feasibility of their use as vaccine targets. While the TSGPs might not be a universal vaccine target that will provide cross-protection for all tick species, neutralization of the toxic activities alone would be a viable option to relieve the burden that these ticks place on their hosts.

Future studies on the BPTI-like hemostatic inhibitors

Expression of the different members of the BPTI-family should provide the means for further structural characterization in terms of mechanisms. An intriguing possibility that emerges from this study is the design of a chimeric protein with fXa, thrombin and platelet aggregation inhibitory capacities. Such a protein might be useful as a multi-functional agent to control thrombosis in a regulated manner and could also be used as a possible vaccine agent, to generate immune responses that could knock out more than one function necessary for tick feeding. The different evolutionary pathways proposed for the BPTI-like family can also be tested by a phylogenetic reconstruction of ancestral proteins and subsequent engineering of chimeric proteins (Chang and Donoghue, 2000). This could also allow the prediction of as yet undiscovered ancestral proteins from the same family.

Conclusion

Evolution as a biological phenomenon is multi-disciplinary. It can be considered on the level of the gene, the genome, protein function, cellular metabolic processes, individual cells, organs or whole organisms and populations. As such no individual study into any of the above categories can provide an accurate description of evolution if the other parts are not also considered. This study considered gene duplication and subsequent gain/loss of protein function, and its role during the adaptation of ticks to a blood-feeding environment. However, gene duplication implies multi-gene families and gain/loss of function and the existence of various distinct functions. To this end, studies into possible functions of several proteins have been conducted, which include inhibitors of platelet aggregation, blood clotting, granule biogenesis and toxins. It can be certain that not all functions of the proteins studied have yet been elucidated as some may be multi-functional and/or be in the process of evolving new functions. Preliminary studies into salivary gland biology has also been conducted as it was made clear that a study of the evolution of bio-active components secreted during feeding cannot fail to take into account the organ of origin. This study has shown that a holistic approach to evolution is necessary if we are to understand tick adaptation to a blood-feeding environment. The

results generated in this manner have left many avenues to be explored that will hopefully lead to a deeper understanding of ticks as highly adapted parasites. While a control method of ticks is envisaged at the end of the road, we hope that this study also shows ticks not only as villains but also as a part of a wonderful world of interaction at molecular level. Tick control after all is a necessity imposed by the drive of man to conquer this world. As always we should consider the possibility that in changing this world we might also lose the keys to its mysteries.

"The theory of evolution is quite rightly called the greatest unifying theory in biology. The diversity of organisms, similarities and differences between kinds of organisms, patterns of distribution and behavior, adaptation and interaction, all this was merely a bewildering chaos of facts until given meaning by the evolutionary theory"

Ernst Mayer, 1970