Function, mechanism and evolution of the moubatin-clade of soft tick lipocalins

Ben J. Mans[*] and **José M.C. Ribeiro**
Laboratory for Malaria and Vector Research, National Institutes of Health, Rockville, MD, 20852, USA.

Abstract

The "moubatin-clade" of soft tick lipocalins, although monophyletic show clear signs of paralogy as indicated by the various functions associated with this protein family. This includes anti-platelet (moubatin), anti-complement (OMCI) and toxic (TSGP2) activities in the vertebrate host. In order to understand the evolution of function and how it relates to the various paralogs in this clade, we characterized a number of different proteins in regard to undefined function and mechanism. By utilizing gain-of-function for TSGP2 and loss-of-function for TSGP3, we show that inhibition of collagen-induced platelet aggregation by moubatin and TSGP3 are due to scavenging of thromboxane A_2. Moubatin, TSGP2 and TSGP3 are also able to bind leukotriene B_4 with high affinity. TSGP2 and TSGP3, but not moubatin binds complement C5, with kinetics that indicates that conformation change occurs during interaction. A conserved loop and histidine residue in the sequences of OMCI, TSGP2 and TSGP3 are implicated in the interaction with complement C5. The data presented suggests that the ancestral function evolved in this clade was aimed at inhibition of vasoconstriction, platelet aggregation and neutrophil aggregation, primarily by scavenging of thromboxane A_2 and leukotriene B_4. C5 complement targeting activity evolved within this clade, probably within the Old World Ornithodorinae. The moubatin-clade itself most probably derived from the related histamine and serotonin-binding lipocalin sub-family that is conserved within the Argasidae.

Keywords

1. Introduction

Ticks (Acari) are obligate, blood-feeding ectoparasites that can be divided into two main families, the soft (Argasidae) and hard (Ixodidae) ticks. During a blood meal the vertebrate host's immune and hemostatic systems are modulated by secretion of salivary gland derived components into the feeding site (Ribeiro, 1995; Mans and Neitz, 2004a). The complexities of this secreted salivary gland cocktail are being revealed by the high-throughput descriptions of sialomes (salivary gland transcriptomes and proteomes of secreted proteins). In the case of soft ticks there might be up to 200 secreted proteins, while the numbers for hard ticks reaches well

*Address correspondence to: B.J. Mans. Current address: Onderstepoort Veterinary Institute, Agricultural Research Council, Onderstepoort, 0110, South Africa. E-mail: mansb@arc.agric.za.

beyond 500 (Mans et al. 2008a; Francischetti et al. 2008; Ribeiro et al. 2006). Functional characterization of these proteins is important for the understanding of tick evolution in regard to parasite-host interactions.

While each tick family and even species described thus far have unique proteins and protein families, a core set of protein families are conserved between ixodids and argasids (Mans et al. 2008a). Of these the most prominent in terms of family members, as well as protein abundance, includes the lipocalin, basic tail secretory family (BTSP), kunitz/BPTI-family and metalloprotease families (Mans et al. 2008a). In soft ticks, the lipocalin family seems to be by far the most abundant in terms of sheer numbers of family members and levels of protein expression (Mans et al. 2001; Mans et al. 2003; Francischetti et al. 2008; Mans et al. 2008a; Oleaga et al. 2007).

Lipocalins are ubiquitous in nature and is characterized by an eight-stranded anti-parallel β-barrel closed off at one end by an N-terminal helix and stabilized by a C-terminal α-helix that packs against the side of the barrel (Skerra, 2000). The open end of the barrel is enclosed by 4 variable loops that make lipocalins particularly suited in regard to antigenic diversity, as well as contributing towards ligand binding specificity (Skerra, 2000). A core set of lipocalins (kernel lipocalins) are relatively easily distinguised based on conserved sequence motifs. These lipocalins are found throughout all kingdoms of life and is probably the ancestors to those lipocalins that are highly divergent in sequence and are classified as outliers (Flower et al. 2000).

In blood-feeding arthropods, notably triatomine bugs and ticks, lipocalins show extreme sequence diversity and is classified as outliers, even though their structures clearly relate them to the rest of the lipocalin family (Paesen et al. 1999; Andersen et al. 2005; Mans et al. 2008b). In ticks, their exon-intron patterns also show conservation with the rest of the lipocalin family indicating a definite evolutionary relationship between outlier and kernel lipocalins (Mans and Neitz, 2004b). Evolutionary pressure on ticks to adapt to a novel blood-feeding environment as well as immune challenge from vertebrate hosts might account for the large sequence divergence observed between tick and kernel lipocalins. The ability of the lipocalin scaffold to accommodate extensive differences in primary structure could also have played an important role in sequence divergence.

Lipocalins have been implicated in modulation of inflammation by scavenging of histamine and serotonin in both hard and soft ticks (Paesen et al. 1999; Sangamnatdej et al. 2002; Mans et al. 2008b). Additional functions for lipocalins from soft ticks include targeting of platelet aggregation and the complement system, as well as being involved in toxicoses and as allergens (Waxman and Connolly, 1993; Mans et al. 2002a; Hilger et al. 2005; Nunn et al. 2005). The majority of these functions are performed by lipocalins that fall within the moubatin-clade of soft tick lipocalins. The moubatin-clade is composed of at least six lipocalins from three soft tick species. These include for *Ornithodoros moubata*: moubatin, an inhibitor of collagen-induced platelet aggregation and OMCI, an inhibitor of the complement cascade that targets C5 (Waxman and Connolly, 1993; Nunn et al. 2005). TSGP2 and TSGP3 have been described in the soft tick *O. savignyi* and have been implicated in sand tampan toxicoses (Mans et al. 2002a). OP-15 and OP-16 have been identified in *O. parkeri* from a sialome project (Francischetti et al. 2008). High levels of sequence identity (45–88%) exist between moubatin, OMCI, TSGP2 and TSGP3. As such several questions regarding this clade still exist.

1. What is the mechanism by which moubatin inhibits platelet aggregation?

2. By what mechanism do OMCI target C5?

3. How conserved is biological function within this seemingly related group of proteins?

4. What is the history of this clade in regard to functional evolution?

We addressed these issues and showed that the moubatin-clade primarily evolved to target vasoconstriction in the vertebrate host, primarily via scavenging of thromboxane A_2 and leukotriene B_4. This also allowed soft ticks to target platelet aggregation and inflammatory responses in the host. Anti-complement activity evolved in a subset of these proteins, notably OMCI, TSGP2 and TSGP3 and its mechanism of action could be traced to a conserved loop linking the beta-barrel and C-terminal helix of the lipocalin fold. This study allows us for the first time to approximate the evolution of function within this orthologous clade.

2. Materials and Methods

2.1 Phylogenetic analysis

Sequences were extracted from the Genbank non-redundant protein sequence database using the sequence of moubatin (Genbank accession: 462613) and BLASTP analysis (Altschul et al. 1990). Sequences with E-values less than one were aligned using ClustalX and the edges trimmed manually to provide a conserved core of sequence alignment (127 informative positions) (Jeanmougin et al. 1998). Neighbor-joining analysis using this alignment was performed using the Mega2 software (Kumar et al. 2001). Gaps were treated as pairwise deletion, amino acid distance were calculated using Poisson correction and branch support were estimated using bootstrap analysis (10 000 bootstraps).

2.2 Recombinant protein expression

Genes were synthesized for moubatin (Genbank accession number: 159944), TSGP2 (Genbank accession number: 25991388) and TSGP3 (Genbank accession number: 25991390). The gene for "moubatin-like 3" (Genbank accession number: 149287030) was obtained from the *O. parkeri* cDNA library previously constructed (Francischetti et al. 2008). Genes contained a start codon for methionine on the 5'-end as well as 5'-NdeI and 3'-XhoI restriction sites that allowed for directional cloning into the pET17b vector (Novagen). Mutants were constructed using conventional PCR amplification methods. In all cases constructs were confirmed by DNA sequencing. For routine manipulation of plasmids, TOP10 *E. coli* (Invitrogen) cells were used as transformation line. For expression, plasmids were transformed into BL21(DE)Lys S *E. coli* cells (Invitrogen). Protein were expressed and refolded from inclusion bodies as previously described (Calvo et al. 2006). Active, refolded protein was purified using size exclusion and ion exchange chromatography. Protein quality and identity was assessed using electrospray mass spectrometry and N-terminal Edman sequencing. In all cases, the recombinant protein's molecular mass differed from that of the calculated mass with values less than 1 Da (results not shown).

2.3 Isothermal titration calorimetry

Isothermal titration calorimetry was performed using a VP-ITC calorimeter (Microcal, Northhampton, MA) as described (Calvo et al. 2006). Briefly, proteins were equilibrated in ITC buffer (20 mM Tris-HCl, pH 7.4, 0.15 M NaCl), that was also used to prepare all ligands tested. All lipid-derived ligands were freshly prepared before use and were dried under nitrogen and redissolved in ITC buffer to a final concentration of 20μM before sonication in a water bath for 10 minutes to ensure homogeneous suspensions. Lipid-derived ligands included arachidonic acid (AA), leukotriene B_4, C_4, D_4 and E_4 (LTB$_4$, LTC$_4$, LTD$_4$, LTE$_4$), U46619, carbocyclic thromboxane A2 (cTXA$_2$), prostaglandin E2 and platelet activating factor (PAF) (Cayman Chemical, MI, USA). Other ligands included serotonin and histamine (Sigma-Aldridge). All solutions were degassed under vacuum for 4 min before use. Aliquots (10 µl) were injected every 120 sec, and the syringe was stirred at 290 rpm while measuring heat of binding at 30°C. After subtraction of the heats of dilution, the net enthalphy data were analyzed

with a single binding site model using the Microcal Origin software package. All proteins were tested at 2 µM protein and 20 µM ligand unless otherwise indicated.

2.4 Platelet aggregation

Platelet aggregation was measured as described previously using an aggregometer (Mans et al. 2008c). Briefly, platelet rich plasma was prepared from medication-free donors by platelet-pheresis (DTM/NIH Blood Bank). Platelets (300 ul) in the presence or absence of sample (10 ul) were pre-stirred in the aggregometer for 10 minutes to monitor pre-aggregation effects. Aggregation was then induced using collagen (0.1 mg/ml total).

2.5 Contraction of rat aorta

Contraction of rat aortic ring preparations by U46619, a TXA_2 mimetic was measured isometrically. A modified Tyrode solution (with 10 mM HEPES buffer) that was oxygenated by continuous bubbling of air was used in the assays (Webster and Prado, 1970). In the first assay aortic rings were pre-constricted by 100 nM U46619 before addition of proteins to give final concentrations of 1 µM. In the second assay, aortic ring preparations were pre-incubated with 100 nM of either moubatin or TSGP3 and increments of 100 nM U46619 were added until maximum contraction was reached.

2.6 Plasmon resonance analysis

Surface plasmon resonance was measured using a Biacore T-100 (Uppsala, Sweden). All proteins (TSGP2, TSGP3, moubatin and variants) were immobilized using standard amine-coupling chemistry and was pre-concentrated using sodium acetate, pH 4.5. Final immobilization levels are indicated in the results section. C5 complement (Sigma-Aldridge) was used as analyte in standard HBS buffer (10 mM Hepes, pH 7.4, 150 mM NaCl) and were injected over 120 second intervals at 30 ul/min. Dissociation was generally allowed to proceed for 600 seconds before regenerating with 50 mM sodium hydroxide at 30 ul/min for 20 seconds. Data were fitted to the 1:1 and two-state conformational change binding models included in the Biacore T-100 software.

3. Results

3.1 Phylogenetic analysis of moubatin-related lipocalins

Using moubatin as seed sequence, BLASTP analysis of the non-redundant database retrieved a number of sequences with E-values less than one. All sequences belonged to lipocalins from the soft tick family (Fig. 1). Neighbor-joining analysis resolved two clades with good support, one that included serotonin and histamine binding proteins previously described for soft ticks (Mans et al. 2008b).The other clade included moubatin and the closely related sequences, OMCI, TSGP2 and TSGP3 (tick salivary gland proteins), as well as two basal sequences from the soft tick *O. parkeri*. Several interpretations and questions arise from the structure of the phylogenetic tree obtained.

Firstly, the moubatin-clade forms a monophyletic group that recapitulates the expected species relationships for *O. moubata, O. parkeri* and *O. savignyi* (Klompen and Oliver, 1993). Even so, the clade shows evidence of paralogy. In this regard, moubatin functions as an inhibitor of collagen-induced platelet aggregation, while OMCI acts as a complement C5 inhibitor (Waxman and Connolly, 1993; Nunn et al. 2005). Both proteins derive from *O. moubata* and are as such paralogs, suggesting that functional divergence and specialization occurred within the moubatin clade. This raises questions about the functional evolution of the other proteins within this clade, in regard to being either platelet aggregation or complement C5 inhibitors,

especially TSGP2 and TSGP3 which are closely related to both moubatin and OMCI (Mans, 2005).

Secondly, a number of lipocalins from *O. parkeri* was previously annotated as moubatin-like proteins (Francischetti et al. 2008), but groups within the serotonin and histamine binding clade, raising questions to their functionality (Fig. 1). As such, knowledge on the molecular basis for the inhibition of collagen-induced platelet aggregation by moubatin and the interaction of OMCI with complement C5 should assist in the understanding of the evolution of this protein family. We thus investigated the functional mechanisms of action for moubatin, TSGP2 and TSGP3.

3.2 Inhibition of collagen-induced platelet aggregation by recombinant lipocalins

Heterologous expressed moubatin and TSGP3 inhibited collagen-induced platelet aggregation in platelet-rich plasma in a similar fashion, giving IC_{50} values of ~300–600 nM (Fig. 2). This corresponded well with values previously obtained for wild-type and recombinant moubatin (Keller et al. 1993). In contrast, TSGP2 did not inhibit platelet aggregation at concentrations as high as 20 µM (Fig. 2). It should be noted that previously, no platelet inhibitory activity was detected for TSGP2 or TSGP3 when collagen-specific inhibitory activity were purified from crude salivary gland extracts of *O. savignyi* (Mans et al. 2003). This could have been due to the low quantities purified from crude salivary gland extracts or loss of function after reversed phase chromatography of the TSGP's.

3.3 Binding of thromboxane A$_2$ analogs by moubatin and TSGP3

Previous results suggested that the inhibition of collagen-induced platelet aggregation by moubatin might be due to competition of moubatin with the TXA_2 receptor (Keller et al. 1993). Alternatively, the related complement inhibitor, OMCI, have been co-crystallized with ricinuleic acid and based on this, it was suggested that moubatin might scavenge TXA_2 (Roversi et al. 2007). Binding of the stable TXA_2 analogs, U46619 and $cTXA_2$ by the tick lipocalins was tested using microcalorimetry. Moubatin and TSGP3 bound both analogs with low nanomolar affinities. TSGP2 did not bind any of the analogs at the concentrations tested, although it bound arachidonic acid at affinities comparable to that of moubatin and TSGP3 (Fig. 3; Table 1).

TSGP2 and TSGP3 show 88% sequence identity (Mans et al. 2003), with an R85G substitution difference. This residue position is located within the binding cavity of the lipocalin barrel and was suggested previously as being a determinant for ligand binding in moubatin (Roversi et al. 2007). In order to confirm that this difference is responsible for the functional difference observed between TSGP2 and TSGP3, mutants were constructed to give TSGP2_R85G and TSGP3_G85R.

For TSGP3_G85R, U46619 binding activity was abolished, while a gain of function in regard to U46619 binding was observed for TSGP2_R85G, with an affinity comparable to wild-type TSGP3 (Fig. 4; Table 1). In addition, TSGP2_R85G inhibited collagen-induced platelet aggregation at concentrations similar to that observed for moubatin and TSGP3, while TSGP3_G85R failed to inhibit platelet aggregation even at concentrations as high as 35 µM (Fig. 4). This gives evidence that binding of TXA_2 analogs and inhibition of platelet aggregation is causally linked, most probably due to scavenging of TXA_2. OMCI also possess the G85R substitution seen for TSGP2, which suggests that it would also be unable to bind TXA_2 (Roversi et al. 2007).

3.4 Biological significance of TXA$_2$ scavenging

TXA$_2$ is a potent platelet aggregation agonist as well as a vasoconstrictor (Coleman et al. 1981). In addition to anti-platelet activity, moubatin and TSGP3 could thus also act as vasodilators. The ability of moubatin and TSGP3 to act as inhibitors of vasoconstriction was tested using contraction of rat aorta (Fig. 5). Both relaxed rat aorta pre-constricted by U46619 as well as inhibiting contraction induced by U46619 in a concentration dependent manner. TSGP3 was previously shown to exist as ~5% of the total protein present in the salivary gland extracts of *O. savignyi* (Mans et al. 2001; Mans et al. 2004c). If it is assumed that only 50% of the salivary gland protein is secreted during feeding and that the feeding site is between 10–50µl in volume, then TSGP3 will be present at ~3–18 µM. This is well above the concentrations needed by TXA$_2$ to induce vasoconstriction or platelet aggregation and given the high affinities for U46619, suggests that moubatin and TSGP3 will be able to act as modulators of vasoconstriction and platelet aggregation at the site of feeding.

3.5 Binding of leukotrienes by tick lipocalins

TXA$_2$ is derived from arachidonic acid, which is the precursor to other bio-active eicosanoids, such as prostaglandins and leukotrienes (Khanapure et al. 2007). The structure of OMCI was solved in complex with the fatty acid ricinuleic acid, suggesting that the moubatin-clade might be able to bind other eicosanoids as well, such as leukotrienes and prostaglandins (Roversi et al. 2007). Binding results indicated that moubatin, TSGP2 and TSGP3 bind LTB$_4$ with affinities in the nanomolar range (Fig. 6; Table 1). Prostaglandin E$_2$, LTC$_4$, LTD$_4$, LTE$_4$ as well as platelet activating factor was not significantly bound by moubatin, TSGP2 or TSGP3 under similar conditions tested, suggesting specificity for TXA$_2$ and LTB$_4$ (results not shown). Leukotriene B$_4$ act as a stimulus for neutrophil migration, aggregation and degranulation during inflammation. It was also shown that LTB$_4$ can act as an indirect vasoconstrictor by causing release of histamine and TXA$_2$ through BLT1 receptor activation (Bäck et al. 2004; Bäck et al. 2007). As such, members from the moubatin clade could also function as regulators of neutrophil function by scavenging of LTB$_4$. It is of interest that TSGP2 can bind arachidonic acid and LTB$_4$, but not TXA$_2$ analogs. LTB$_4$ and arachidonic acid lacks the hydroxyl group located at carbon 15 of TXA$_2$ (Fig. 3 and Fig. 6), which most probably sterically clashes with Arg85 in the structures of TSGP2 and OMCI.

3.6 Binding to C5 complement

In addition to anti-platelet and anti-inflammatory lipocalins, the moubatin clade also contains OMCI, an important inhibitor of the complement cascade via targeting of C5 (Nunn et al. 2005). The high sequence similarities shared between OMCI, moubatin, TSGP2 and TSGP3 was previously indicated and the question was raised as to the conservation of C5-binding function (Mans, 2005). Plasmon resonance analysis indicated that TSGP2 and TSGP3 bind to C5 with affinities similar to that observed for OMCI (Fig. 7A; Table 2) (Hepburn et al. 2007). Compared to TSGP2, moubatin did not interact with C5, even though similar levels of immobilization were obtained for both proteins (Fig. 7C).

3.7 Conformation change during binding of C5 to lipocalins

Kinetic data fitted best to a two-state conformational change model (Fig. 7A; Table 2). A linked-reaction assay, that compares the dissociation rate of C5 when injected for different time periods, showed that the rates of the dissociation curves differ (Fig. 7C). This is the expected behavior for time-dependent conformation change and suggests that the two-state conformation model is valid, with the implication that a conformation change occurs during the binding of TSGP2 and TSGP3 to C5.

3.8 Mapping the active site regions of C5 interaction

The fact that TSGP2 and TSGP3, but not moubatin binds to C5, suggests that OMCI, TSGP2 and TSGP3 share determinants important for C5 binding. As such, positions conserved between OMCI, TSGP2 and TSGP3, but not moubatin were mapped onto a multiple sequence alignment. Of these, two main regions of potential interest were identified (Fig. 8A). The first was four residues (TSGP2: Thr66, Lys73, Thr75 and Thr82) that mapped to a common interface formed by beta-strands D and E. This interface corresponds roughly with a similar region previously proposed as a model for the interaction of OMCI and C5 (Roversi et al. 2007). A construct (T2M4) with all four positions mutated to amino acids found in moubatin (TM4: T66M, K73N, T75V, T82K) did not have any effect on binding of C5 to TSGP2 (Fig. 9A), indicating that this region is irrelevant to interaction with C5.

3.9 The βH-α2 loop of TSGP2 is important for C5 interaction

The other region of potential interest is a conserved loop located between beta-strand H of the β-barrel and the C-terminal α-helix (Fig. 8). This loop is highly conserved between OMCI, TSGP2 and TSGP3 but not moubatin. Binding to C5 was abolished when the loop for TSGP2 was replaced with that of moubatin to yield construct T2ML (TSGP2_MWMLEAGGLEV-QWQSNGSADDK) (Fig. 9B). This recombinant still bound LTB_4 with a similar affinity as the wild-type indicating that it was still correctly folded (Table 1).

3.10 Orientation of the βH-α2 loop by His95 is important for high affinity binding to C5

The βH-α2 loop in the structure of OMCI have the interesting feature that a conserved histidine (TSGP2: His95) located on the loop between beta-strands F and G seems to orient the loop via structured waters and steric hindrance (Fig. 8C). The nitrogen atoms of histidine form hydrogen bonds to the water molecules, which in turn interact with the backbone carbonyl atoms of Ala137-Gly138-Gly139 (Fig. 8). In moubatin this histidine is replaced by an aspartic acid (Fig. 8). A mutant of TSGP2 (TSGP2_H95D) that simulates this position in moubatin led to a 10-fold decrease in binding levels of C5 (Fig. 9C). The TSGP2_H95D mutant showed increased rates of association as well as dissociation compared to the wild type protein. Kinetic analysis indicated that affinity decreased ~10-fold and that the data could be fitted to a single site binding model (Table 4). This suggests that in this mutant the second-order phase associated with conformation change was lost and that this conformation change is important for increased affinity. It also supports a function for the βH-α2 loop in the interaction with C5 and suggests that the specific conformation of the loop is important and necessary to facilitate conformation change. It should be noted that this loop area as region of interaction is novel and differ from the proposed model for OMCI-C5 interaction previously proposed (Roversi et al. 2007). It is beyond the scope of the current study to predict how the soft tick lipocalins will interact with C5 complement, but the results thus far suggest that it will not mimic interaction of the C8γ lipocalin with complement C5 as proposed by Roversi et al. (2007).

3.11 "Moubatin-like" lipocalins in the biogenic amine-binding clade

Thus far, the present study focused on members from the moubatin clade. The question has been raised whether those lipocalins annotated as "moubatin-like" (Francischetti et al. 2008), but groups into the biogenic amine-binding clade, will have moubatin-like functions. Previously, the residues involved in the binding of serotonin and histamine in soft tick lipocalins has been elucidated (Mans et al. 2008b). Mapping of these residues and those identified as active site residues in the structure of OMCI indicates that most residues involved in ligand interaction are conserved within but not between the different clades (Fig. 10).

"Moubatin-like 3" lies well within the histamine and serotonin-binding clade (Fig. 1). It stoichiometrically binds one serotonin or two histamine molecules (Fig. 11; Table 1), as was

previously indicated for its ortholog TSGP1 (Mans et al. 2008b). It did not bind LTB_4, U46619 or cTXA2 (results not shown). Taken with the phylogenetic analysis and conservation of ligand-binding residue data, this strongly suggests that other members from this clade will also be serotonin and histamine binders, rather than binders of eicosanoids.

4. Discussion

Blood-feeding lifestyles evolved independently almost 20 times within arthropods. This occurred in triatomine bugs (Triatoma and Rhodnius genera), bed bugs (Cimex sp.), fleas (Siphonaptera), lice (Phthiraptera), nematocerous flies (Culicidae, Ceratopogonidae, Psychodidae, Simuliidae), brachycerous flies (Glossinidae, Hippoboscidae, Tabanidae), moths (Calyptra sp.), parasitiform mites (Ixodidae, Argasidae, Laelapidae, Macronyssidae) and acariform mites (Hirstiella sp.). Independent evolution is reflected in the diversity and redundancy of the biochemical activities found within the saliva's of these arthropods (Ribeiro, 1995). The complexities of salivary gland derived transcriptomes underscore this and have revealed the presence of multi-member protein families (Ribeiro and Francischetti, 2003). Multimember protein families emphasize the important role that gene duplication has played in the evolution of novel function during the adaptation to a blood-feeding lifestyle (Mans and Neitz, 2004). It also poses significant challenges for the functional annotation of salivary gland transcriptomes. Gene duplication and the subsequent evolution of novel function is the focus of the current study. We studied a specific group of closely related lipocalins (the moubatin-clade) found within soft ticks and mapped out functions for paralogous proteins. We show how closely related proteins can have different or mixed functions, an important consideration in the functional annotation of paralogous gene families.

The current study characterized functions for moubatin, TSGP2 and TSGP3 as well as a histamine and serotonin binding lipocalin from the soft tick *O. parkeri*. We showed that moubatin and TSGP3 inhibits platelet aggregation by scavenging of TXA_2. This allows them to act as potent inhibitors of TXA_2 mediated vasoconstriction. TSGP2 is unable to bind TXA_2 due to a G85R substitution in the lipocalin ligand-binding cavity. All proteins can, however, bind LTB_4, which implicate them in modulation of neutrophil function. TSGP2 and TSGP3, but not moubatin, can bind to C5 complement, in a similar fashion as the complement pathway inhibitor OMCI. The differences in C5 targeting ability was linked to a conserved loop found in OMCI, TSGP2 and TSGP3, but not moubatin.

TSGP2 has been implicated in toxicoses caused by *O. savignyi* that affects the cardiac system of the host (Mans et al. 2002a; Mans et al. 2003). As yet, the mechanism for this cardio-pathology has not been resolved. It should be noted that tick salivary glands are rich in blood-meal derived arachidonic acid (Shipley et al. 1993). As such, it is possible that TSGP2 will bind eicosanoid-related compounds before secretion into the host. To this end, the adverse effects of arachidonic acid metabolites on the cardiac system are well known and it is possible that toxicoses induced by *O. savignyi* could be associated with eicosanoid metabolism (Feuerstein and Hallenbeck, 1987). How this relates to the non-toxicity of TSGP3, is not yet apparent. It should be noted that TSGP4, another lipocalin implicated in sand tampan toxicoses has been shown to bind arachidonic acid and cysteinyl leukotrienes as well (Mans and Ribeiro, 2008).

The data thus far suggests several lines of thought on the evolution and conservation of functions associated with the moubatin clade.

1. Most of the residues involved in the interaction of OMCI with ricinuleic acid are conserved in the lipocalins that group within the moubatin-clade (Fig.1 and Fig. 9). This would suggest that this whole clade, including the two homologs from *O. parkeri*, should be able to bind arachidonic acid, LTB_4 and TXA_2. The exceptions

are TSGP2 and OMCI that possess the TSGP2_G85R substitution that will preclude their binding of TXA_2. As such, binding of TXA_2 and LTB_4 are probably the ancestral functions associated with this clade.

2. The importance of the βH-α2 loop and His95 in the binding of TSGP2 to complement C5 and their conservation patterns suggests that C5-binding function is limited to OMCI, TSGP2 and TSGP3. The absence of conserved sequences in the *O. parkeri* lipocalins and the position of OMCI, TSGP2 and TSGP3 in the phylogenetic tree, suggests that moubatin is the ancestral gene and that C5 binding function evolved from a gene duplication that occurred within ticks of the Old World. This would also imply that a moubatin-paralog to TSGP2/TSGP3 exists in *O. savignyi* as previously suggested (Mans et al. 2003).

3. No homologs for the moubatin clade have been found in an *Argas monolakensis* salivary gland cDNA library (Mans et al. 2008a). Interestingly, most *Argas* species feeds on birds, which lack platelets, the main producer of TXA_2 in the mammalian hemostatic system. This could suggest that the "moubatin-clade" of lipocalins evolved after the divergence of the major soft tick genera, possibly after irradiation of mammals. The implication of this would be that TXA_2 and LTB_4 scavenging capability, as well as complement inhibitory activity evolved after the divergence of the main tick families, as has been suggested for anti-hemostatic components as well (Mans et al. 2002b; Mans and Neitz, 2004a). In this regard, complement inhibitors that targets C3 of the complement cascade are found in prostriate ticks and belong to the unique "Isac-family" (Valenzuela et al. 2000; Daix et al. 2007). This would support independent evolution of complement and inflammation targeting function in the different tick families. Host-specificity and host-specific immunity could have been major determinants in the evolution of these functions, which might thus be either genus or even species specific adaptations. In contrast, it was recently shown that anti-thrombin and GPIIbIIIa antagonists are conserved within the soft tick family, presumably due to the conservation of basic hemostatic functions in vertebrates and evolution of anti-hemostatic functions in the ancestral soft tick lineage (Mans et al. 2008c).

4. Conserved disulphide bond patterns, BLAST as well as phylogenetic analysis suggests that the moubatin and the serotonin and histamine binding clades are evolutionary related (Mans et al. 2003; Mans, 2005; Mans et al. 2008a). It was also shown that serotonin and histamine binding proteins are conserved between the major soft tick genera (Mans et al. 2008b). This would suggest that the moubatin clade originated from a gene duplication of a serotonin and histamine binding ancestral lipocalin, most probably after divergence of the *Argas* and *Ornithodoros* genera. This has the interesting implication that LTB_4 and TXA_2 binding activity evolved from a lipocalin binding cavity that at first bound serotonin (Mans et al. 2008b). Comparison of ligand binding residues responsible for either binding of serotonin or LTB_4/TXA_2 shows that several are indeed conserved, although important differences exists that could explain ligand binding specificity as well as functional shift that occurred during evolution (Fig. 10).

Acknowledgements

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J. Mol. Biol 1990;215:403–410. [PubMed: 2231712]

Andersen JF, Gudderra NP, Francischetti IM, Ribeiro JM. The role of salivary lipocalins in blood feeding by *Rhodnius prolixus*. Arch. Insect Biochem. Physiol 2005;58:97–105. [PubMed: 15660358]

Bäck M, Qiu H, Haeggström JZ, Sakata K. Leukotriene B4 is an indirectly acting vasoconstrictor in guinea pig aorta via an inducible type of BLT receptor. Am. J. Physiol. Heart. Circ. Physiol 2004;287:H419–H424. [PubMed: 15016629]

Bäck M, Sakata K, Qiu H, Haeggström JZ, Dahlén SE. Endothelium-dependent vascular responses induced by leukotriene B4. Prostaglandins Other Lipid. Mediat 2007;83:209–212. [PubMed: 17481557]

Calvo E, Mans BJ, Andersen JF, Ribeiro JM. Function and evolution of a mosquito salivary protein family. J. Biol. Chem 2006;281:1935–1942. [PubMed: 16301315]

Coleman RA, Humphrey PP, Kennedy I, Levy GP, Lumley P. Comparison of the actions of U-46619, a prostaglandin H2-analogue, with those of prostaglandin H2 and thromboxane A2 on some isolated smooth muscle preparations. Br. J. Pharmacol 1981;73:773–778. [PubMed: 7248665]

Daix V, Schroeder H, Praet N, Georgin JP, Chiappino I, Gillet L, de Fays K, Decrem Y, Leboulle G, Godfroid E, Bollen A, Pastoret PP, Gern L, Sharp PM, Vanderplasschen A. *Ixodes* ticks belonging to the *Ixodes ricinus* complex encode a family of anticomplement proteins. Insect Mol. Biol 2007;16:155–166. [PubMed: 17298559]

Feuerstein G, Hallenbeck JM. Prostaglandins, leukotrienes and platelet-activating factor in shock. Ann. Rev. Pharmacol. Toxicol 1987;27:301–313. [PubMed: 3034139]

Flower DR, North AC, Sansom CE. The lipocalin protein family: structural and sequence overview. Biochim. Biophys. Acta 2000;1482:9–24. [PubMed: 11058743]

Francischetti IM, Mans BJ, Meng Z, Gudderra N, Veenstra TD, Pham VM, Ribeiro JM. An insight into the sialome of the soft tick, *Ornithodorus parkeri*. Insect Biochem. Mol. Biol 2008;38:1–21. [PubMed: 18070662]

Hepburn NJ, Williams AS, Nunn MA, Chamberlain-Banoub JC, Hamer J, Morgan BP, Harris CL. *In vivo* characterization and therapeutic efficacy of a C5-specific inhibitor from the soft tick *Ornithodoros moubata*. J. Biol. Chem 2007;282:8292–8299. [PubMed: 17215252]

Hilger C, Bessot JC, Hutt N, Grigioni F, De Blay F, Pauli G, Hentges F. IgE-mediated anaphylaxis caused by bites of the pigeon tick *Argas reflexus*: cloning and expression of the major allergen Arg r 1. J. Allergy Clin. Immunol 2005;115:617–622. [PubMed: 15753913]

Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ. Multiple sequence alignment with Clustal X. Trends Biochem. Sci 1998;23:403–405. [PubMed: 9810230]

Keller PM, Waxman L, Arnold BA, Schultz LD, Condra C, Connolly TM. Cloning of the cDNA and expression of moubatin, an inhibitor of platelet aggregation. J. Biol. Chem 1993;268:5450–5456. [PubMed: 8449907]

Khanapure SP, Garvey DS, Janero DR, Letts LG. Eicosanoids in inflammation: biosynthesis, pharmacology, and therapeutic frontiers. Curr. Top. Med. Chem 2007;7:311–340. [PubMed: 17305573]

Klompen JSH, Oliver JH. Systematic relationships in the soft ticks (Acari: ixodida: Argasidae). Sys. Entomol 1993;18:313–331.

Kumar S, Tamura K, Jakobsen IB, Nei M. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 2001;17:1244–1245. [PubMed: 11751241]

Mans BJ, Venter JD, Very PJ, Louw AI, Neitz AW. Identification of putative proteins involved in granule biogenesis of tick salivary glands. Electrophoresis 2001;22:1739–1746. [PubMed: 11425229]

Mans BJ, Steinmann CM, Venter JD, Louw AI, Neitz AW. Pathogenic mechanisms of sand tampan toxicoses induced by the tick, *Ornithodoros savignyi*. Toxicon 2002a;40:1007–1016. [PubMed: 12076655]

Mans BJ, Louw AI, Neitz AW. Evolution of hematophagy in ticks: common origins for blood coagulation and platelet aggregation inhibitors from soft ticks of the genus *Ornithodoros*. Mol. Biol. Evol 2002b; 19:1695–1705. [PubMed: 12270896]

Mans BJ, Louw AI, Neitz AW. The major tick salivary gland proteins and toxins from the soft tick, *Ornithodoros savignyi*, are part of the tick Lipocalin family: implications for the origins of tick toxicoses. Mol. Biol. Evol 2003;20:1158–1167. [PubMed: 12777525]

Mans BJ, Neitz AW. Adaptation of ticks to a blood-feeding environment: evolution from a functional perspective. Insect Biochem. Mol. Biol 2004a;34:1–17. [PubMed: 14723893]

Mans BJ, Neitz AW. Exon-intron structure of outlier tick lipocalins indicate a monophyletic origin within the larger lipocalin family. Insect Biochem. Mol. Biol 2004b;34:585–594. [PubMed: 15147759]

Mans BJ, Neitz AW. Molecular crowding as a mechanism for tick secretory granule biogenesis. Insect Biochem. Mol. Biol 2004c;34:1187–1193. [PubMed: 15522614]

Mans BJ. Tick histamine-binding proteins and related lipocalins: potential as therapeutic agents. Curr. Opin. Investig. Drugs 2005;6:1131–1135.

Mans BJ, Andersen JF, Francischetti IM, Valenzuela JG, Schwan TG, Pham VM, Garfield MK, Hammer CH, Ribeiro JM. Comparative sialomics between hard and soft ticks: Implications for the evolution of blood-feeding behavior. Insect Biochem. Mol. Biol 2008a;38:42–58. [PubMed: 18070664]

Mans BJ, Ribeiro JM, Andersen JA. Structure, function and evolution of biogenic amine-binding proteins in soft ticks. J. Biol. Chem. 2008b(In press).

Mans BJ, Andersen JF, Schwan TG, Ribeiro JM. Characterization of anti-hemostatic factors in the argasid, *Argas monolakensis*: implications for the evolution of blood-feeding in the soft tick family. Insect Biochem. Mol. Biol 2008c;38:22–41. [PubMed: 18070663]

Mans BJ, Ribeiro JM. A novel group of peptide leukotriene scavengers in soft ticks. Insect Biochem. Mol. Biol. 2008(Submitted).

Nunn MA, Sharma A, Paesen GC, Adamson S, Lissina O, Willis AC, Nuttall PA. Complement inhibitor of C5 activation from the soft tick *Ornithodoros moubata*. J. Immunol 2005;174:2084–2091. [PubMed: 15699138]

Oleaga A, Escudero-Población A, Camafeita E, Pérez-Sánchez R. A proteomic approach to the identification of salivary proteins from the argasid ticks *Ornithodoros moubata* and *Ornithodoros erraticus*. Insect Biochem. Mol. Biol 2007;37:1149–1159. [PubMed: 17916501]

Paesen GC, Adams PL, Harlos K, Nuttall PA, Stuart DI. Tick histamine-binding proteins: isolation, cloning, and three-dimensional structure. Mol. Cell 1999;3:661–671. [PubMed: 10360182]

Ribeiro JM. Blood-feeding arthropods: live syringes or invertebrate pharmacologists? Infect. Agents Dis 1995;4:143–152. [PubMed: 8548192]

Ribeiro JM, Francischetti IM. Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. Ann. Rev. Entomol 2003;48:73–88. [PubMed: 12194906]

Ribeiro JM, Alarcon-Chaidez F, Francischetti IM, Mans BJ, Mather TN, Valenzuela JG, Wikel SK. An annotated catalog of salivary gland transcripts from *Ixodes scapularis* ticks. Insect Biochem. Mol. Biol 2006;36:111–129. [PubMed: 16431279]

Roversi P, Lissina O, Johnson S, Ahmat N, Paesen GC, Ploss K, Boland W, Nunn MA, Lea SM. The structure of OMCI, a novel lipocalin inhibitor of the complement system. J. Mol. Biol 2007;369:784–793. [PubMed: 17445829]

Sangamnatdej S, Paesen GC, Slovak M, Nuttall PA. A high affinity serotonin- and histamine-binding lipocalin from tick saliva. Insect Mol. Biol 2002;11:79–86. [PubMed: 11841505]

Skerra A. Lipocalins as a scaffold. Biochim. Biophys. Acta 2000;1482:337–350. [PubMed: 11058774]

Shipley MM, Dillwith JW, Bowman AS, Essenberg RC, Sauer JR. Changes in lipids of the salivary glands of the lone star tick, *Amblyomma americanum*, during feeding. J. Parasitol 1993;79:834–842. [PubMed: 8277375]

Valenzuela JG, Charlab R, Mather TN, Ribeiro JM. Purification, cloning, and expression of a novel salivary anticomplement protein from the tick, *Ixodes scapularis*. J. Biol. Chem 2000;275:18717–18723. [PubMed: 10749868]

Waxman L, Connolly TM. Isolation of an inhibitor selective for collagen-stimulated platelet aggregation from the soft tick, *Ornithodoros moubata*. J. Biol. Chem 1993;268:5445–5449. [PubMed: 8449906]

Webster M, Prado ES. Glandular kallikreins from horse and human urine and from hog pancreas. Meth. Enzymol 1970;19:681–699.
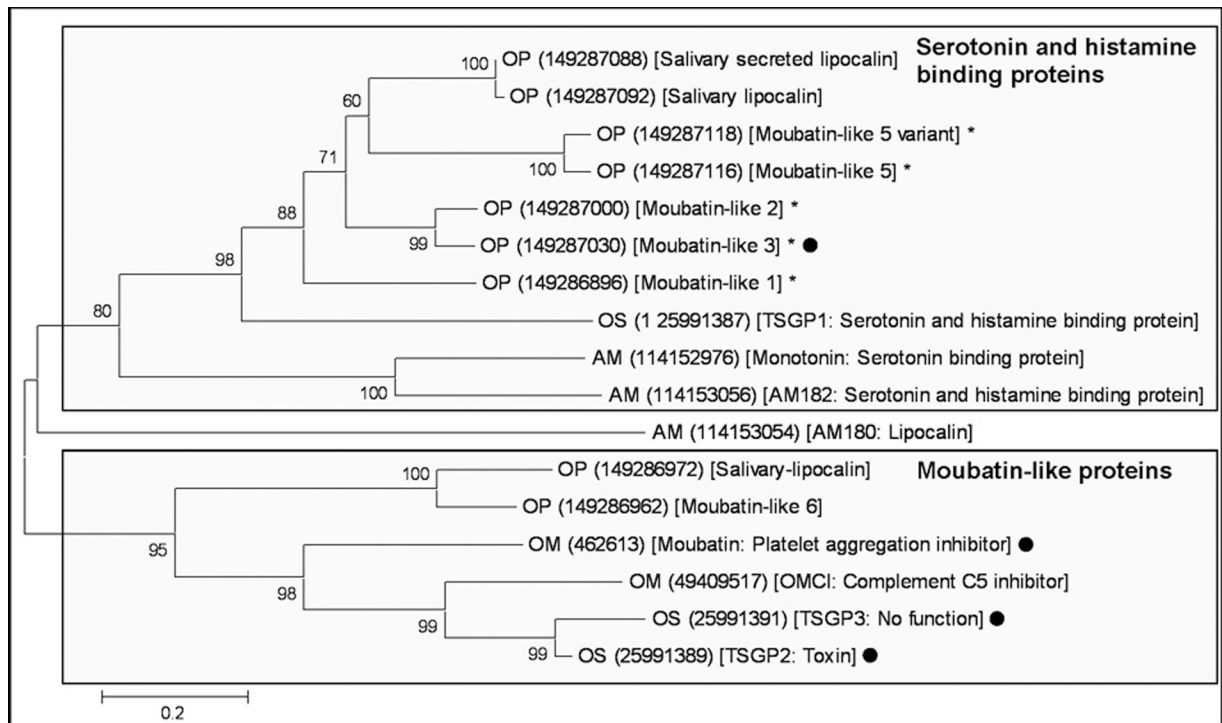
Fig. 1.
Phylogenetic analysis of moubatin-related lipocalins. Lipocalins were retrieved from the non-redundant protein database using the sequence for moubatin. All hits with E-values below one were extracted and used to construct a neighbor-joining tree. Support values (10 000 bootstraps) for nodes is indicated. The two major supported clades are boxed. Sequences are identified by a species designation (AM: *Argas monolakensis*; OM: *Ornithodoros moubata*; OP: *O. parkeri*, OS: *O. savignyi*) and a Genbank accession number within round brackets. The Genbank description line and functional information extracted from the literature is included within square brackets. Sequences marked with dots were characterized in the present study. Those marked with asterisks indicate proteins whose functional and evolutionary relationship with moubatin is uncertain.
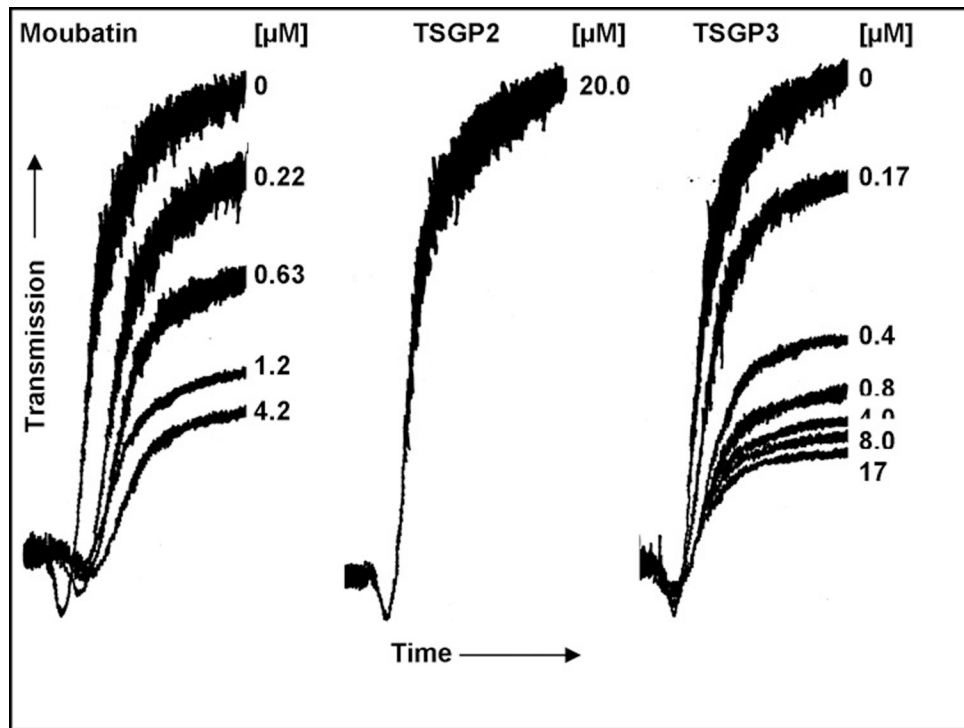
Fig. 2.
Inhibition of collagen-induced platelet aggregation. Indicated are curves representing platelet aggregation of platelet-rich plasma induced by 0.1 mg/ml collagen in the presence of various concentrations of moubatin, TSGP2 and TSGP3.
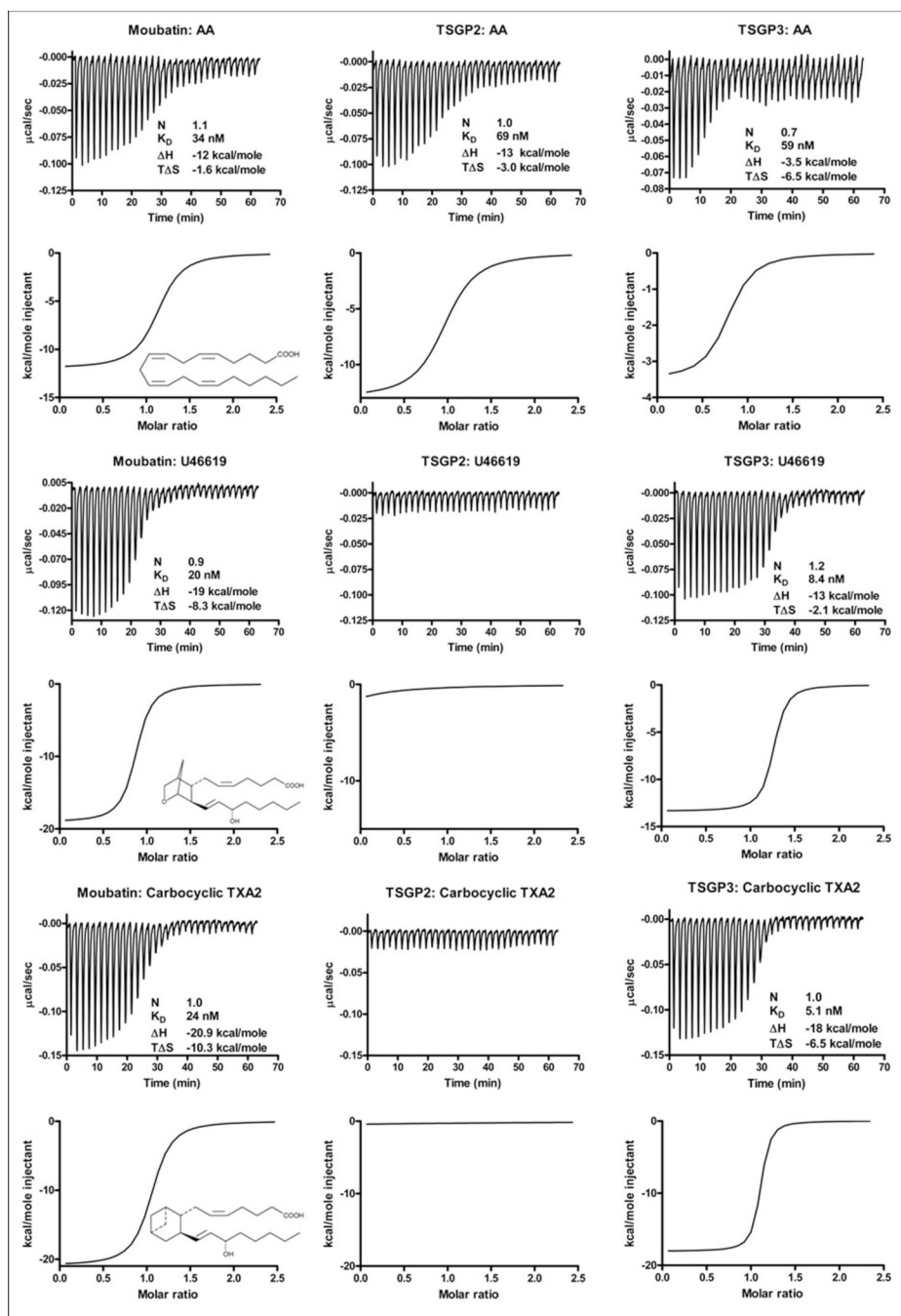
Fig. 3.
Binding of arachidonic acid (AA), thromboxane A_2 (TXA_2) and carbocyclic TXA_2 ($cTXA_2$) by moubatin, TSGP2 and TSGP3 as measured by microcalorimetry. In each case, protein (2μM) was titrated against 20μM ligand. Indicated are thermodynamic parameters derived for each titration that includes the stoichiometry of binding (N), dissociation constant (K_D), change in enthalpy (ΔH) and change in entropy (TΔS) upon binding. Also shown are the structures of arachidonic acid, U46619 and $cTXA_2$.
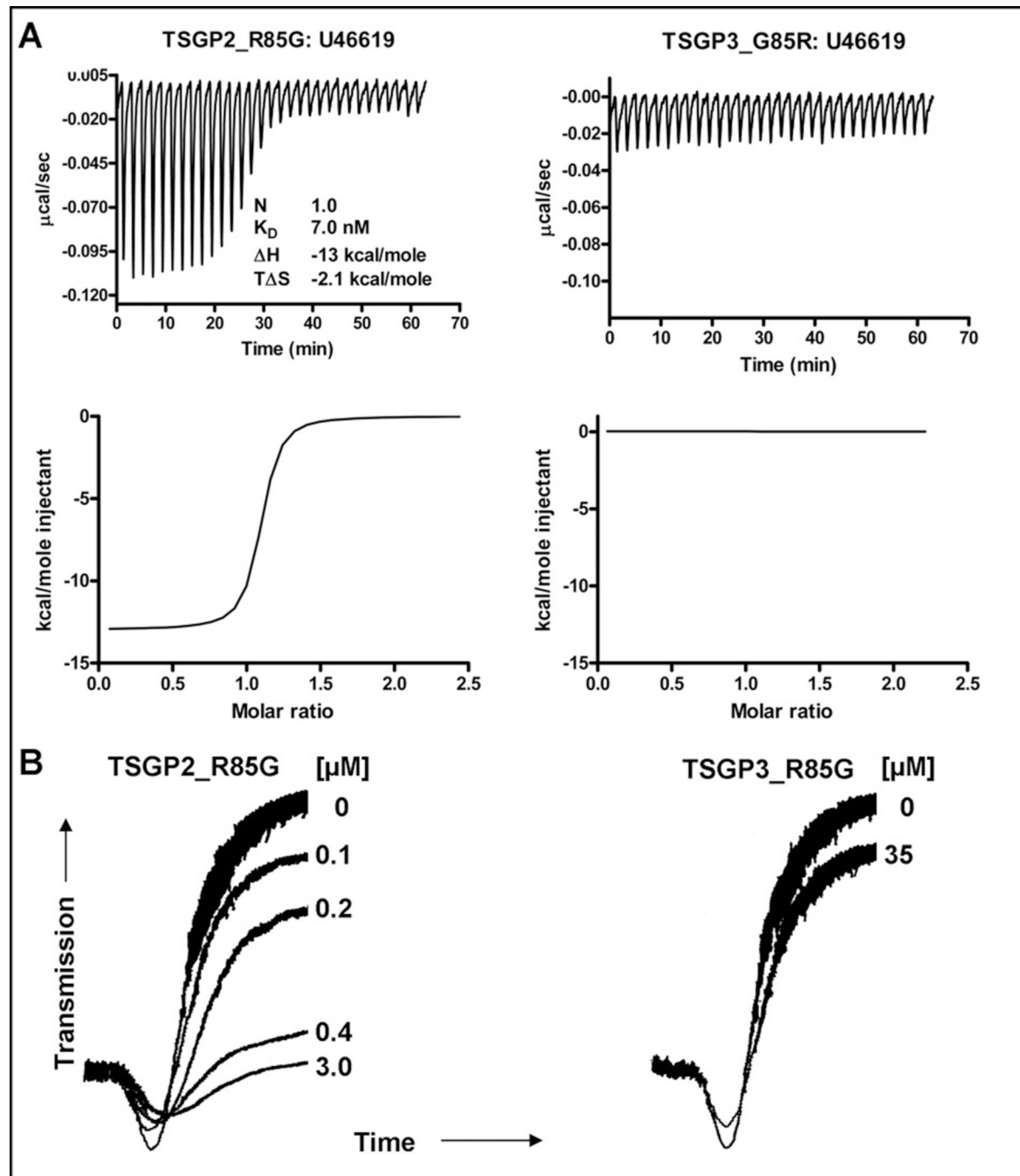
Fig. 4.
Functional shift of TSGP2_R85G and TSGP3_G85R mutants. A) Binding of U44619 by
TSGP2_R85G and TSGP3_G85R mutants as measured by microcalorimetry. Protein (2 µM)
was titrated using 20 µM ligand. B) Inhibition of collagen-induced platelet aggregation by
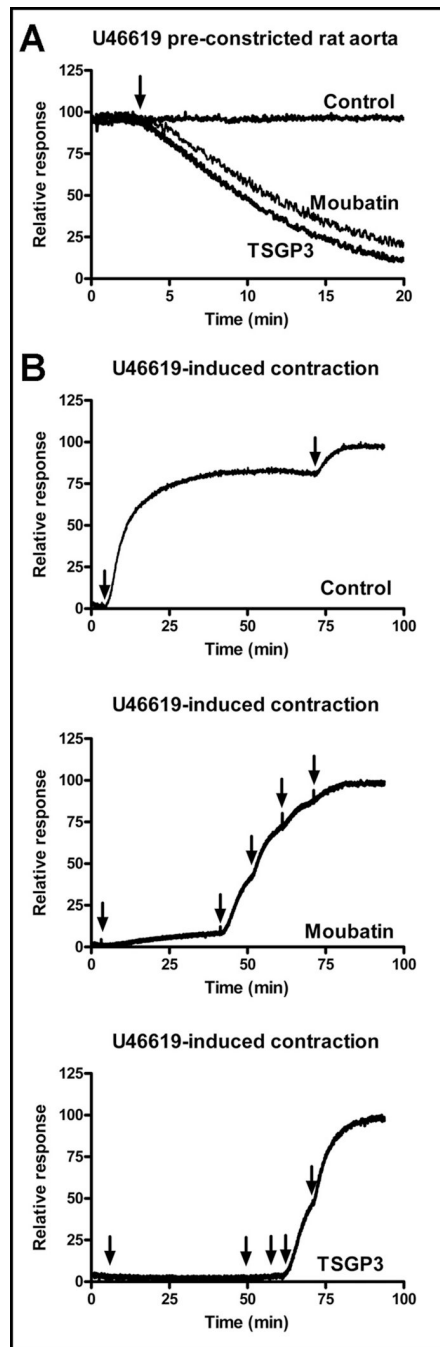TSGP2_R85G and TSGP3_G85R mutants.

Fig. 5.
Contraction of U46619-induced rat aorta. A) Rat aorta was contracted using 100 nM U46619, before addition of moubatin or TSGP3 (arrow) to give 1μM final concentration. B) Rat aorta was pre-incubated with saline (control), 100 nM moubatin or TSGP3 before addition of U46619 in 100 nM increments as indicated by arrows.
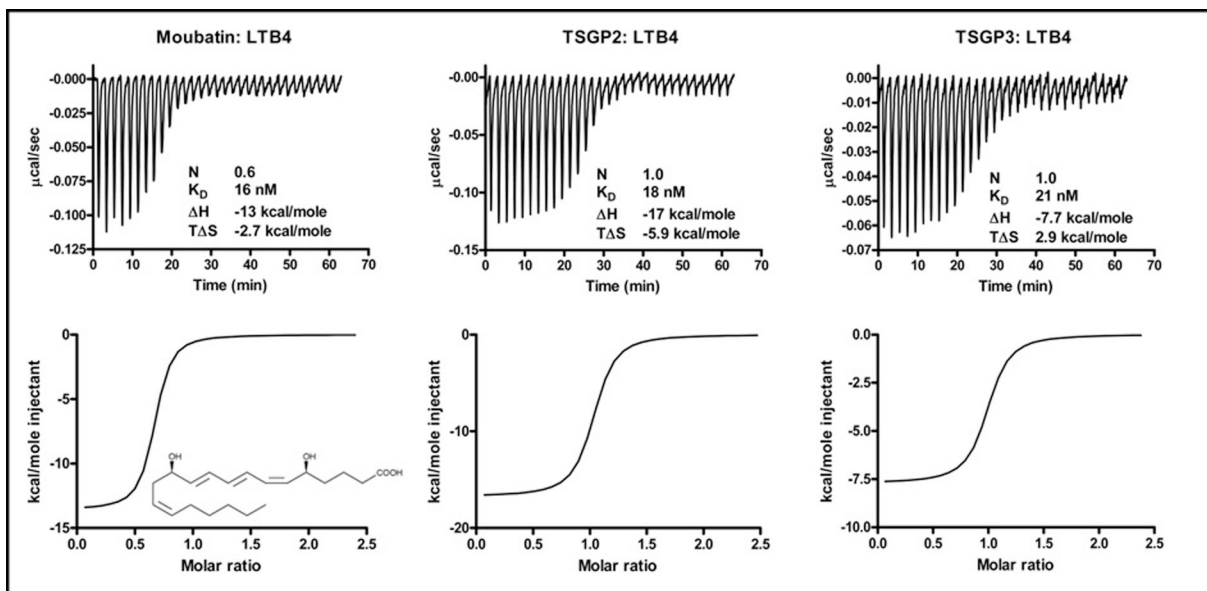
Fig. 6.
Binding of LTB$_4$ by moubatin, TSGP2 and TSGP3 measured by microcalorimetry. Protein (2µM) was titrated against 20 µM ligand. Indicated are thermodynamic parameters derived for each titration that includes the stoichiometry of binding (N), dissociation constant (K$_D$), change in enthalpy (ΔH) and change in entropy (TΔS) upon binding. Also shown is the structure of LTB$_4$.
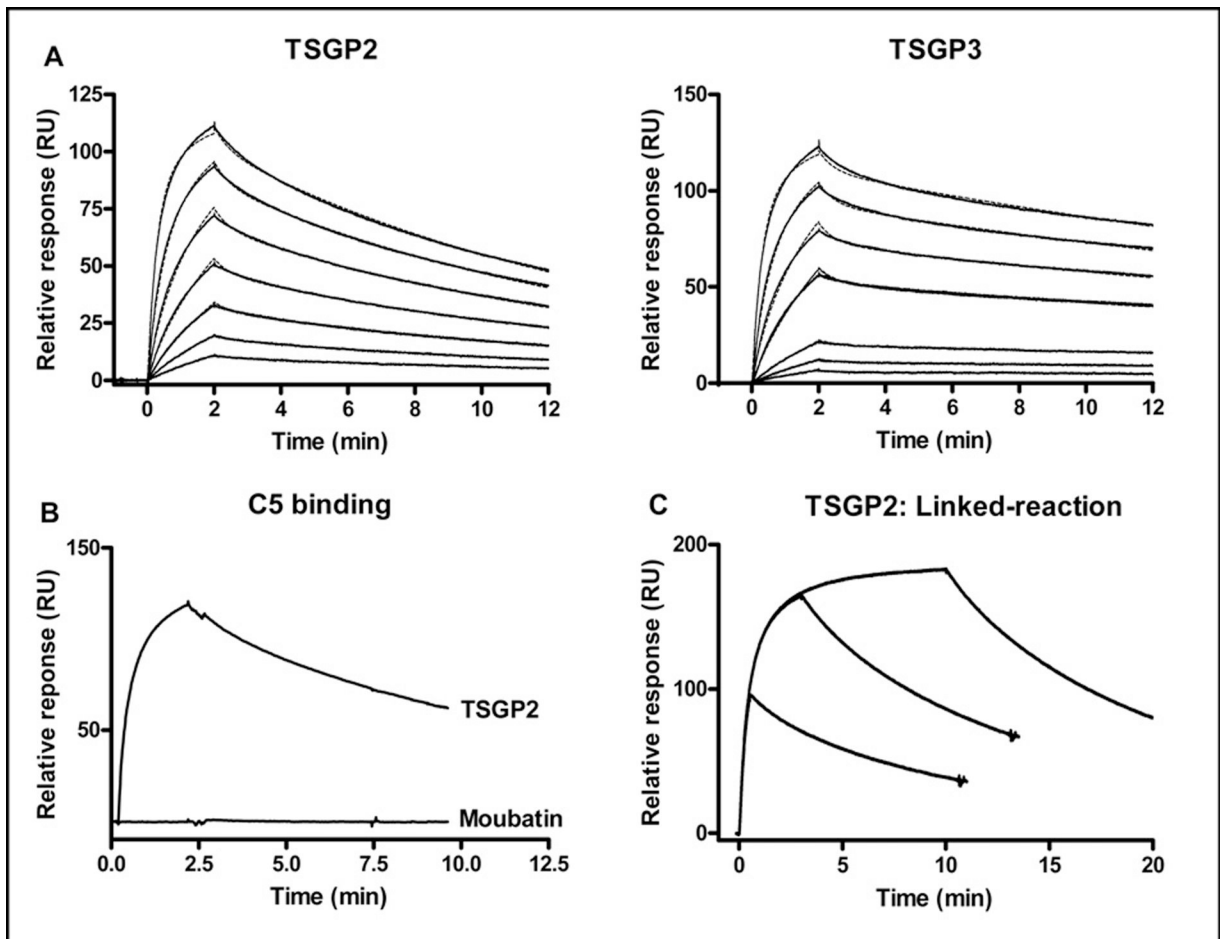
Fig. 7.
Interaction of lipocalins with complement C5. A) Binding of human C5 to immobilized TSGP2 and TSGP3. The fit to a two-state conformational change model is shown by the dashed line. Curves from bottom to top, indicate 8, 16, 33, 66, 132, 263, 526 nM, respectively. B) Binding of human C5 to immobilized TSGP2 (channel 2: 219 RU final immobilized level) or moubatin (channel 4: 253 RU final immobilized level). C) C5 (300 nM) injected for different time periods over an immobilized TSGP2 surface.
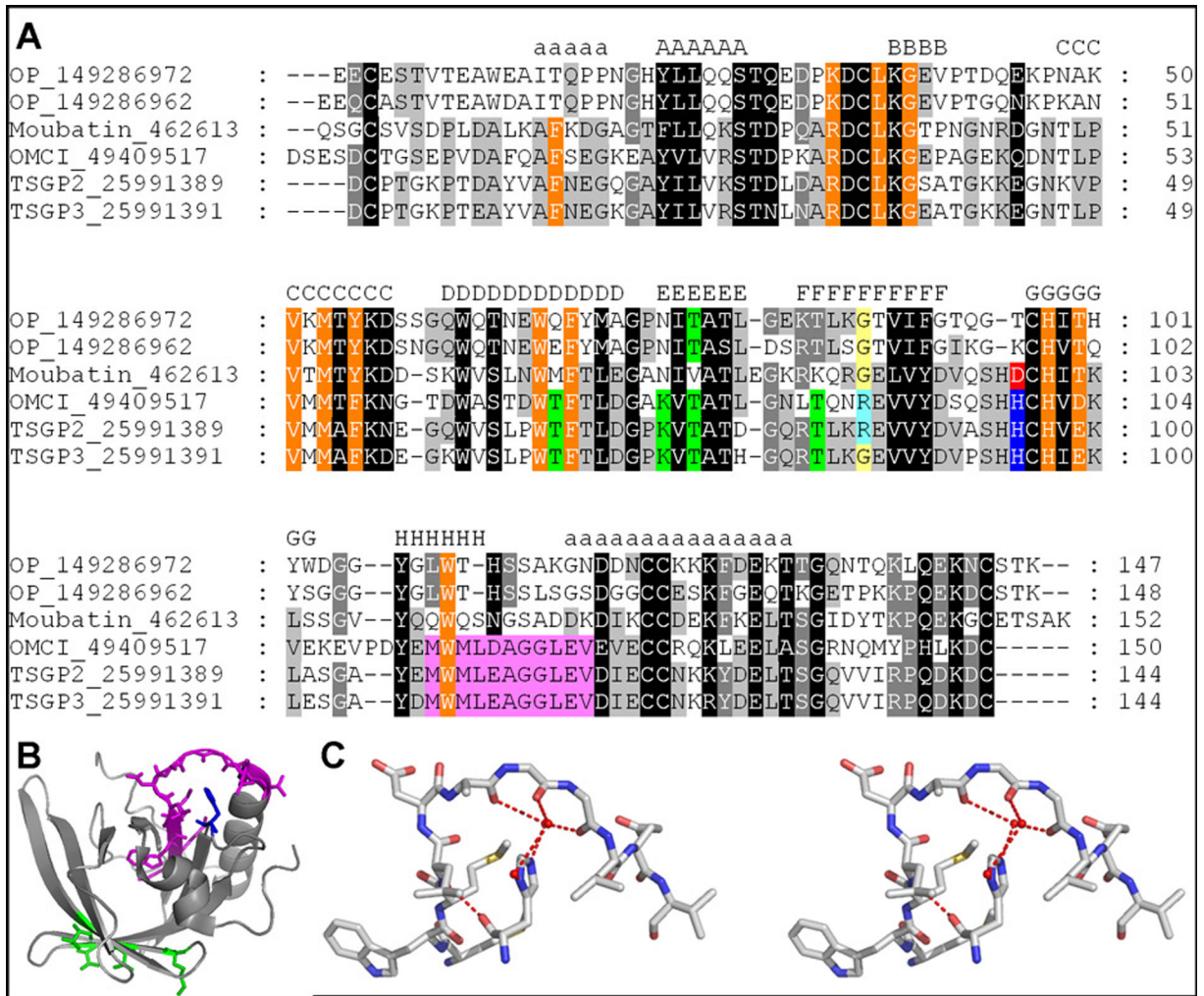
Fig. 8.
Active site residues in the moubatin clade of lipocalins. A) Alignment of lipocalins from the moubatin clade. Residues involved in ricinuleic acid binding by OMCI are indicated in orange (Roversi et al. 2007). The four residues mutated in construct T2M4, that had no effect on C5 binding are colored in green. The loop, important for C5 binding that was mutated in construct T2ML is colored in purple. The H95D substitution in construct T2_H95D that led to loss of second order association of TSGP2 with C5 is colored in blue and red, respectively. The R85G substitution in construct TSGP2_R85G that led to gain of TXA_2 binding function is colored in yellow and light blue respectively. B) The structure of OMCI, with corresponding residues colored that was mutated in TSGP2 to test for C5 interaction. These include those for constructs T2M4 (green), T2ML (purple) and TSGP2_H95D (blue). C) Stereo view of the orientation of histidine and the βH-α2 loop. Waters are indicated as van der Waals radii.
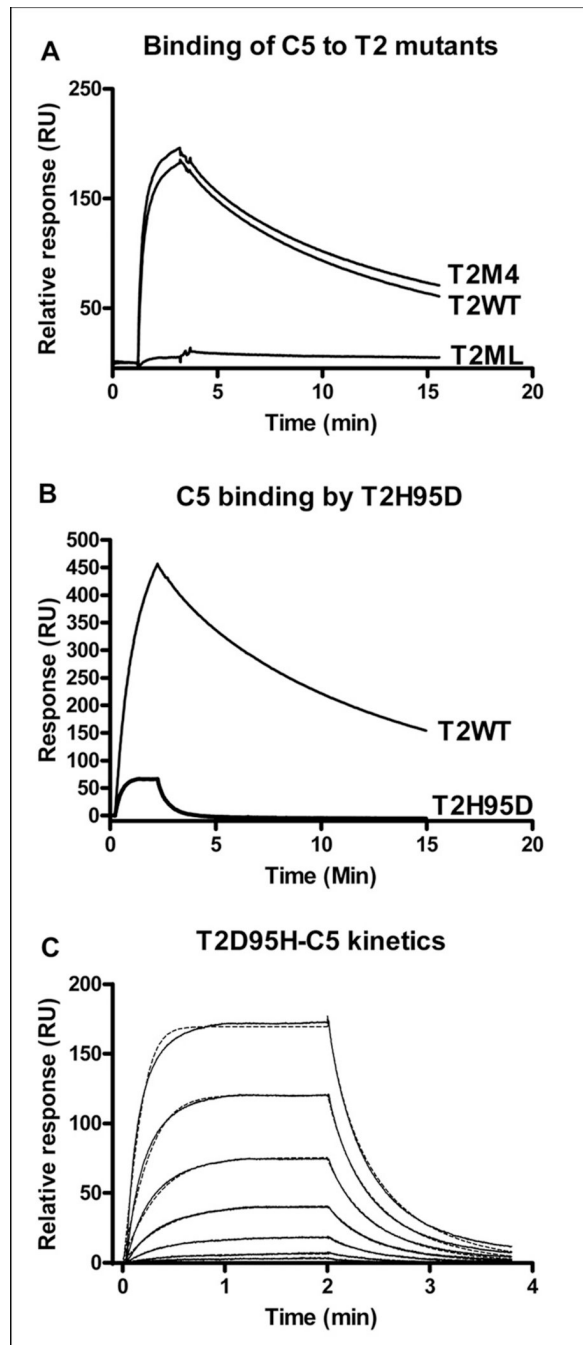
Fig. 9.
C5 binding to TSGP2 mutants. A) C5 binding to immobilized T2WT (310 RU final
immobilized level), T2M4 (174 RU final immobilized level) and T2ML (251 RU final
immobilized level). B) C5 binding to immobilized T2WT (channel 2: 533 final immobilized
level) and T2H95D (channel 4: 572 final immobilized level). C) Binding kinetics for C5 to
immobilized T2H95D. A single binding site model fit is indicated by the dashed line. Curves
from bottom to top, indicate 8, 16, 33, 66, 132, 263, 526 nM, respectively.
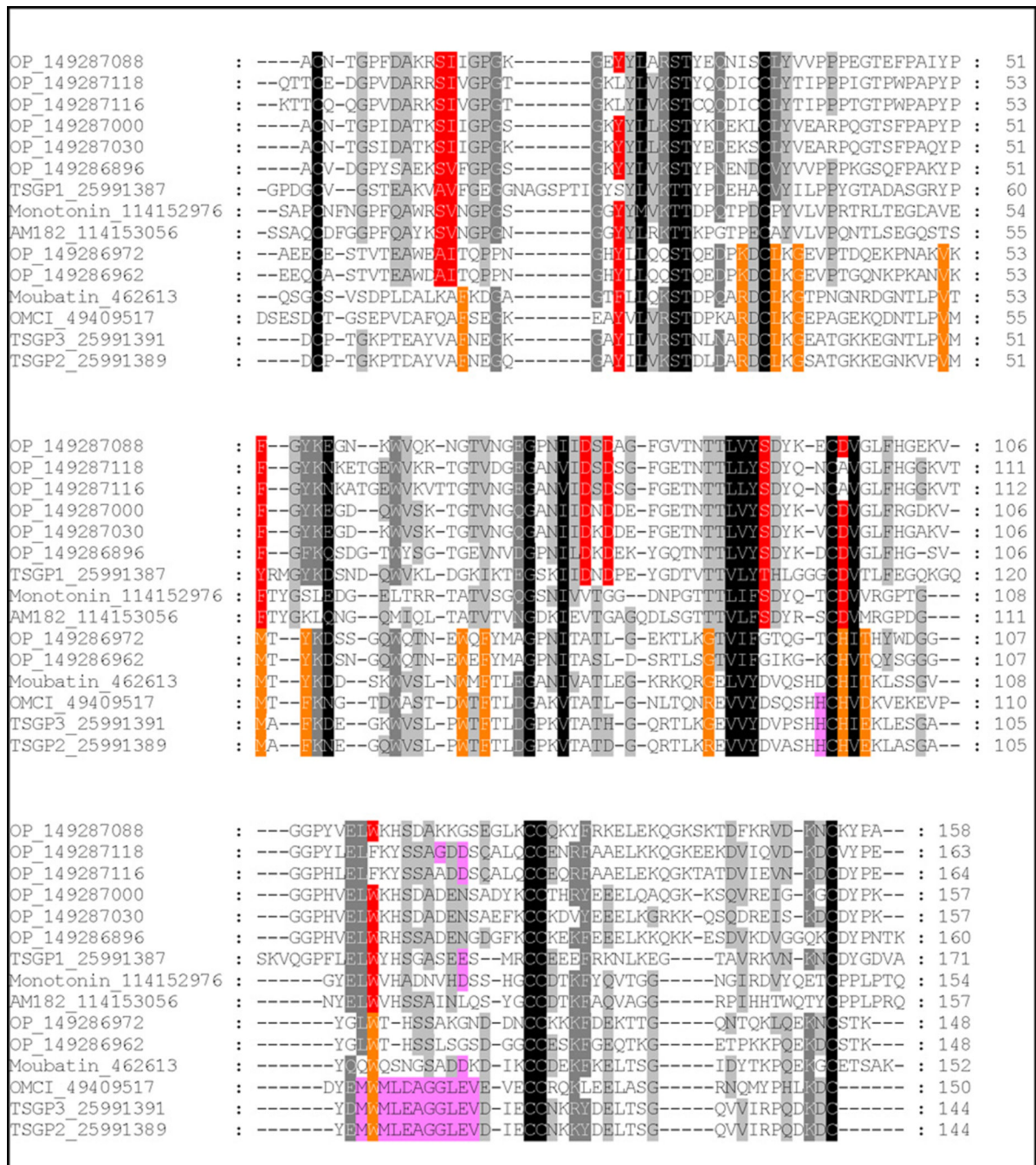
Fig. 10.
Multiple alignment of tick lipocalins from the moubatin and serotonin and histamine-binding
clades. Conserved residues that line the ligand binding pocket of the moubatin-clade lipocalins
are colored in green. Binding pocket data were inferred from those residues that interact with
ricinuleic acid in the structure of OMCI (Roversi et al. 2007). Conserved residues that interact
with serotonin and histamine are colored in yellow. Data were obtained from the structures of
monotonin (Mans et al. 2008). Residues implicated in complement C5 interaction are colored
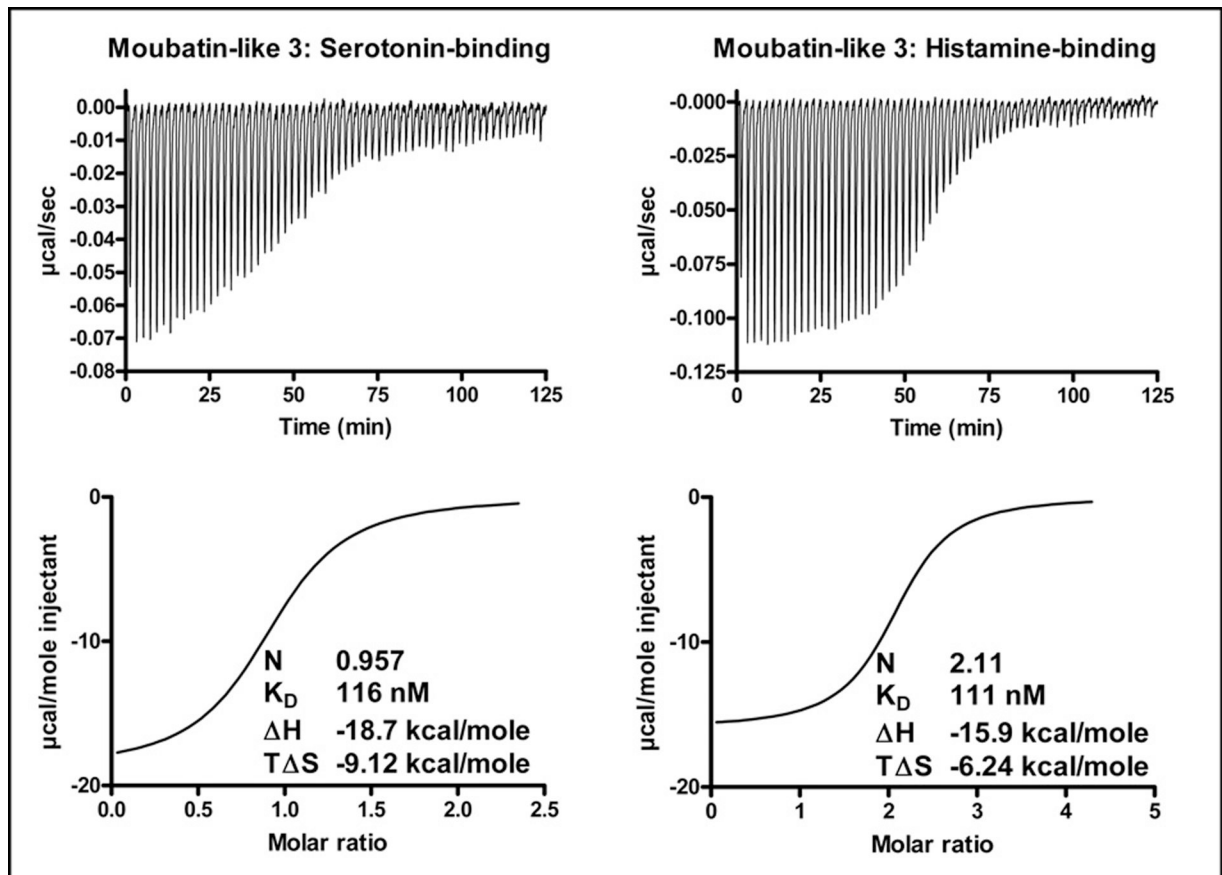purple.

Fig. 11.
Serotonin and histamine binding by "Moubatin-like 3" measured by microcalorimetry. Protein (2μM) was titrated against 20 μM ligand. Indicated are thermodynamic parameters derived for each titration that includes the stoichiometry of binding (N), dissociation constant (K_D), change in enthalpy (ΔH) and change in entropy (TΔS) upon binding.

Table 1

Thermodynamic parameters for binding of arachidonic acid, U44619, $cTXA_2$ and LTB_4 by moubatin, TSGP2 and TSGP3. Also indicated is binding of 5-HT and histamine by "moubatin-like 3".

Protein	n[a]	ΔH[b]	TΔS[b]	Kd[c]	ΔG[b]
Arachidonic acid					
Moubatin	2.2	−12.0 ± 0.2	−1.6	34 ± 6	−10.4
TSGP2	1.0	−13.0 ± 0.2	−3.0	69 ± 7	−10.0
TSGP3	0.7	−3.5 ± 0.2	6.5	59 ± 24	−10.0
U46619					
Moubatin	0.9	−19.0 ± 0.3	−8.3	20 ± 3	−10.7
TSGP2	NB[d]	NB[d]	NB[d]	NB[d]	NB[d]
TSGP3	1.2	−13.3 ± 0.1	−2.1	9 ± 1	−11.2
TSGP2_R85G	1.0	−13.0 ± 0.1	−1.7	7 ± 1	−11.3
TSGP3_G85R	NB[d]	NB[d]	NB[d]	NB[d]	NB[d]
cTXA$_2$					
Moubatin	1.0	−20.9 ± 0.3	−10.3	24 ± 4	−10.6
TSGP2	NB[d]	NB[d]	NB[d]	NB[d]	NB[d]
TSGP3	1.0	−18.0 ± 0.1	−6.5	5 ± 0.7	−11.5
LTB$_4$					
Moubatin	0.6	−13.5 ± 0.2	−2.8	16 ± 3	−10.7
TSGP2	1.0	−16.7 ± 0.2	−6.0	18 ± 3	−10.7
TSGP3	1.0	−7.7 ± 0.1	2.9	21 ± 5	−10.6
TSGP2_R85G	1.0	−10.7 ± 0.1	0.3	10 ± 1.6	−11.0
TSGP3_G85R	1.0	−9.1 ± 0.3	0.1	228 ± 36	−9.2
T2_ML	0.9	−14.0 ± 0.2	−3.0	12 ± 2	−11.0
T2_M4	1.0	−16.3 ± 0.1	−4.9	6 ± 0.8	−11.4
T2_H95D	0.9	−14.4 ± 0.2	−3.3	12 ± 2	−11.1
Serotonin					
Moubatin-like 3	1.0	−18.7 ± 0.3	−9.1	116 ± 13	−9.6
Histamine					
Moubatin-like 3	2.1	−15.9 ± 0.1	−6.2	106 ± 3	−9.7

[a]Values indicate the stoichiometry of ligand binding.

[b]Values are given as kcal/mol. S.E.M. represents the deviation of the experimental data from the fitted data.

[c]Values are given as nanomolar.

[d]NB No binding detected at 2μM protein and 20μM ligand. S.E.M. represents the deviation of the experimental data from the fitted data.

Table 2

Kinetic parameters for C5 interaction with TSGP2, TSGP3 and T2H95D, using plasmon resonance analysis.

Protein	Ka1 (1/Ms)	Kd1 (1/s)	Ka2 (1/s)	Kd2 (1/s)	KD (nM)	Chi2 (RU2)	Model
TSGP2	1.0E5	0.017	0.019	0.003	25.8	0.371	Two-state
TSGP3	6.8E4	0.018	0.019	0.001	13.9	0.652	Two-state
T2H95D	5.5E5	0.13	-	-	236	1.71	1:1 binding