D. Zhang · M.S. Cupp · E.W. Cupp

Polymorphism of the thrombostasin gene in the horn fly (Haematobia irritans) revealed in a cDNA library and in genomic DNA

Received: 3 May 2001 / Accepted: 29 June 2001 / Published online: 15 August 2001
© Springer-Verlag 2001

Abstract Thrombostasin (TS) is a newly described thrombin-inhibiting protein isolated from the saliva of the horn fly (Haematobia irritans), a blood-sucking ectoparasite of cattle. This report provides a detailed characterization of the TS gene and the first analysis of the allelic complexity of a gene for an anti-hemostatic protein from a blood-feeding insect. Multiple point mutations at fixed positions in the TS gene were identified in a cDNA library prepared from mRNA isolated from horn fly salivary glands. When translated, the variant mRNAs would specify five biochemically active peptides that differ in molecular weight, isoelectric point and predicted secondary structure. Allelic variation with the same mutation pattern was revealed in the genomes of individual flies collected in the field and sampled from a long-standing laboratory colony. Approximately 60% of flies examined carried heterozygous alleles, including five additional alleles not found in the cDNA library. Comparative analysis of the allelic mutations and the predicted effects on secondary structures of the active proteins produced suggest that the TS gene may be undergoing evolutionary selection.

Keywords Horn fly · Anti-thrombin · Thrombostasin · Heterozygosity · Allele-specific PCR

Introduction

Thrombostasin (TS) is a newly described anti-thrombin protein found in the salivary glands of horn flies (Haematobia irritans) (Cupp et al. 2000; Zhang et al. 2001). TS is postulated to serve a critical function in blood-feeding by this fly, which is a major pest of cattle and is of economic importance worldwide. The horn fly, a migrant species, was introduced into North America from Europe in the 1880s (McLintock and Depner 1954), and was recorded in South America in the 1980s (Guglielmone et al. 1999). Survival of this insect is solely dependent upon the availability of mammalian bloodmeals and cattle blood is the preferred choice (Kuramochi 2000). Analysis of the components of horn fly saliva indicated that TS is the major factor responsible for delaying blood clotting (Cupp et al. 2000) and, hence, may be critical for the fly’s ability to counter its host’s hemostatic mechanism and maintain its blood feeding lifestyle.

HPLC-purified TS showed a single protein band in SDS-PAGE analysis, two groups of protein species that differ in pI in two-dimensional gels, and more than four ion species in electrospray ionization mass spectroscopy (ESI-MS) (Zhang et al. 2001). These results indicate that TS is comprised of a family of protein molecules with similar amino acid sequences. The cloned TS gene revealed a coding sequence of 471 bp which could be conceptually translated into a secretory peptide of 157 amino acids. N-terminal sequence analysis of the purified TS, isolated from horn fly saliva, showed, however, that the mature form of TS was composed of 81 amino acids, starting at the 77th residue (serine) from the N-terminus of the full-length peptide. This purified 81-amino acid peptide was active against thrombin in kinetic assays and delayed blood clotting in vitro.

Since the control of horn flies by reliance on pesticides has been challenged by the rapid development of resistance, immunization to neutralize the anti-clotting activity of TS is under investigation as an alternative measure. Success in this endeavor may hinge on a clear understanding of the structural variations of TS in the horn fly populations as related to potency and immunogenicity.

In the present study expressed transcripts of the TS gene were isolated from a cDNA library made from RNA prepared from horn fly salivary glands, obtained

Communicated by E. K. F. Bautz

D. Zhang · M.S. Cupp · E.W. Cupp
Department of Entomology and Plant Pathology, 301 Funchess Hall, Auburn University, Auburn, AL36849-5413, USA
E-mail: ecupp@acesag.auburn.edu
Fax: +1-334-8445005
from a laboratory colony, and the structure of the TS gene was analyzed using genomic DNA from individual flies from both laboratory and field-collected populations.

**Materials and methods**

cDNA library construction and screen

Total RNA was prepared from horn fly salivary glands using the micro RNA isolation kit (Stratagene). Messenger RNA was subsequently isolated by affinity chromatography on a column of oligo(dT) cellulose (Stratagene). The reagents and methods supplied in the ZAP expression cDNA synthesis kit (Stratagene) were used to synthesize and construct the cDNA library according to the manufacturer’s instructions. The resulting cDNAs, flanked by EcoRI sites at the 5′-end and XhoI sites at the 3′-end, were cloned into the phagemid vector pBK-CMV.

Protocols for plaque hybridization and colorimetric detection with NBT and X-phosphate from the Genius system (Boehringer Mannheim) were followed to screen the cDNA library using a DIG (digoxigenin)-labeled probe. The probe was made by incorporating DIG-dUTP via PCR amplification of the known sequence with the primers Hits1 (5′-CGGCTCCCATCACACTGCA-3′, forward) and Hits6 (5′-CTGGAAACCTGTCACAAAC-3′, reverse). Individual positive phagemids were excised in vivo using ExAssist helper phage with the XLOLR bacterial strain (Stratagene). XLOLR clones with putative inserts were subjected to plasmid purification for sequencing.

Genomic DNA extraction and PCR amplification

Field populations of horn flies were collected from cattle kept at the Auburn University College of Veterinary Medicine in Auburn, Ala. Colonies were raised in the laboratory from pupae provided by the U.S. Livestock Insects Research Laboratory in Kerrville, Tex. Individual flies (approximately 3 mg) were homogenized with 40 μl of buffer (10 mM TRIS-HCl, pH 8.0, 2 mM EDTA, and 0.4 M NaCl) and genomic DNA was extracted from the homogenate (Aljanabi and Martínez 1997). Ten nanograms of each genomic DNA were used for PCR with Taq DNA polymerase using the following primer pairs: Hits8 (5′-ATCATGAGCATTCTCTAG-3′, forward, containing the initiation codon ATG) and Hits9 (5′-GCTTATGAGGCTTGTGAACA-3′, reverse, containing the stop codon TAA). The PCR was carried out by mixing the following components in a final volume of 50 μl: 20 mM TRIS-HCl (pH 8.4), 50 mM KCl, 1.5 mM MgCl2, 200 μM dNTPs, 0.5 μM of each primer, and genomic DNA. The mixture was incubated at 94°C for 3 min and then maintained at 80°C until 2.5 U of polymerase was added to each tube. A 0.3-cycle amplification followed: 94°C for 45 s, 54°C for 45 s, and 72°C for 45 s. A final extension step was carried out at 72°C for 7 min. PCR products were directly sequenced after separation by agarose gel and purification with Sephaglas BP (Pharmacia Biotech).

**Heterozygous allele discrimination by allele-specific PCR**

Heterozygous TS alleles, revealed in sequences of individual horn fly genomes, typically showed two equally strong signals for two different bases in the same location on sequencing gels (see below). Thus, an allele-specific PCR (AS-PCR) method (Sommer et al. 1992; Mitterer et al. 1999) was used to resolve the specific composition of each allele pair in the region of the gene coding for the mature, active protein. Four pairs of allele-specific primers (all reverse, Table 1) were used to discriminate between four heterozygous codons wherever two or more codons were heterozygous in a single PCR product: codon 29 G/A, codon 52 G/A, codon 74 A/G, and codon 76 A/G. Using the purified PCR product as template (the same as that used for sequencing described above), a common primer Hits1 (forward) was first paired with each one of the allele-specific primers at the extreme 5′-end, such as TSR226A and TSR226G (codon 76) in AS-PCR, together with an AAT-positive template and a GAT-positive template as controls. The PCR products amplified with TSR226A and TSR226G then served as templates for the next round of PCR with the next pair of allele-specific primers toward the 5′-end, and so on. The AS-PCR conditions were optimized as follows: in a final volume of 50 μl containing 60 mM TRIS-HCl (pH 8.8), 150 mM (NH4)2SO4, 3.5 mM MgCl2, 50 μM dNTPs, 0.3 U Taq polymerase, 0.1 μM Hits11, 0.05 μM allele-specific primer, and approximately 1.0 ng of template. A 30-cycle reaction was carried out after an initial de-naturation step at 94°C for 3 min, each cycle comprising 30 s at 94°C, 30 s at 56-64°C (individual annealing temperatures are shown in Table 1) and 30 s at 72°C, with a final extension at 72°C for 7 min.

**Table 1 Sequences of primers used for allele-specific PCR**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5′→3′)</th>
<th>Coordinates (bp)</th>
<th>Annealing temperature (°C)</th>
<th>Specific for codon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hits11</td>
<td>AGTGCGGGTCCCCATCACACTG</td>
<td>1–21</td>
<td>56</td>
<td>29</td>
</tr>
<tr>
<td>TSR86A</td>
<td>GATTGTCATTAGAGCTTC</td>
<td>86–103</td>
<td>56</td>
<td>29</td>
</tr>
<tr>
<td>TSR86T</td>
<td>GATTGTCATTAGAGCTCA</td>
<td>86–103</td>
<td>56</td>
<td>29</td>
</tr>
<tr>
<td>TSR115A</td>
<td>TCTAGTCAAGCTT</td>
<td>155–171</td>
<td>58</td>
<td>52</td>
</tr>
<tr>
<td>TSR115T</td>
<td>TCTAGTCAAGCTT</td>
<td>155–171</td>
<td>58</td>
<td>52</td>
</tr>
<tr>
<td>TSR220A</td>
<td>CATGGGGACATCTACGCT</td>
<td>220–238</td>
<td>64</td>
<td>74</td>
</tr>
<tr>
<td>TSR220G</td>
<td>CATGGGGACATCTACGCT</td>
<td>220–238</td>
<td>64</td>
<td>74</td>
</tr>
<tr>
<td>TSR226A</td>
<td>TGCAGCATTGGAACATT</td>
<td>226–243</td>
<td>60</td>
<td>76</td>
</tr>
<tr>
<td>TSR226G</td>
<td>TGCAGCATTGGAACATT</td>
<td>226–243</td>
<td>60</td>
<td>76</td>
</tr>
</tbody>
</table>

*Refer to the coding sequence for mature thrombostasin in Fig. 1A*