

## Zusammenfassung

Cobra Venom Factor (CVF), das komplement-aktivierende Protein aus dem Gift der Monokel-Kobra *Naja kaouthia*, ist strukturell und funktionell homolog zur Komplement-Komponente C3. Ähnlich wie C3b, die aktive Form von C3, bildet CVF mit Faktor B eine C3-Konvertase. In vorangegangenen Arbeiten wurde beim Screening von cDNA-Bibliotheken aus Giftdrüsen-mRNA die Sequenz eines weiteren CVF-Moleküls, genannt CVF2, identifiziert. Die cDNA-Sequenzen von CVF1, CVF2 und C3 sind ca. 92% identisch.

In der vorliegenden Arbeit wurde die Struktur des CVF1-Gens aufgeklärt und mit der des humanen C3-Gens verglichen. Für die Aufklärung der Genstruktur wurden eine genomische Cosmid-Bibliothek, genomische PCR und Genome Walking eingesetzt.

Die sich aus den erhaltenen Daten ergebende Struktur des vollständigen CVF1-Gens umfaßt 40 Exons und 39 Introns. Die Exongrößen variieren von 46 bp bis 1140 bp, und Introngrößen variieren von 100 bp bis ca. 9000 bp. Es konnte gezeigt werden, dass das CVF1-Gen mit einer Mindestgröße von 89 kb mehr als doppelt so groß wie das humane C3-Gen (42 kb) ist. Dennoch sind die Strukturen beider Gene in hohem Ausmaß konserviert. Mehr als die Hälfte der einander entsprechenden Exons zeigen identische Größen. Unterschiede in den Exongrößen variieren meist von 3 bp bis 15 bp und repräsentieren stets das Vielfache eines Codons. Dabei sind die Intronphasen nahezu identisch. Sequenzhomologien bei den Introns wurden nur am äußersten Randbereich festgestellt. Ein bedeutsamer Unterschied jedoch besteht darin, dass die konservierten Exons 31 und 32 im CVF1-Gen kontinuierlich und nicht durch ein Intron unterbrochen sind, was im Gegensatz zu den Strukturen der humanen C3, C4 und C5-Gene steht.

In den darüber hinaus untersuchten Strukturen der Gene von CVF2 und coC3 sind die Exons 31 und 32 ebenfalls kontinuierlich und nicht durch ein Intron unterbrochen. Es zeigte sich, dass die Exongrößen der CVF/C3-Genfamilie stark konserviert sind. Als Ursache für einen Größenunterschied von 9 bp in Exon 33 des CVF1-Gens im Vergleich zu den entsprechenden Exons in den CVF2- und coC3-Genen konnte eine singuläre Mutation in der Intron-Spleißstelle identifiziert werden. Obwohl die Introngrößen der drei Gene teilweise stark variieren, sind ausgeprägte Sequenzhomologien auch an internen Positionen vorhanden.

Zusätzlich wurden die Promotorregionen der drei homologen Gene näher analysiert, um mögliche Ursachen für die Expression in unterschiedlichen Geweben zu finden. Putative TATA-Boxen wurden in allen drei Genen an konservierten Positionen gefunden. Es konnte

gezeigt werden, dass sich die stromaufwärts gelegenen Regionen allerdings in den möglichen Bindungsstellen für Transkriptionsfaktoren unterscheiden.

Einige Ergebnisse wiesen darauf hin, dass das Kobra-Genom zusätzlich zu den bereits bekannten Genen CVF1, CVF2 und coC3 möglicherweise noch zwei weitere homologe Gene bzw. Pseudogene enthält. Diese Ergebnisse sind im Einklang mit multiplen C3-Genen in Teleost-Fischen und multiplen Toxin-Genen in verschiedenen *Naja*-Spezies.

## Summary

Cobra Venom Factor (CVF), the complement-activating protein in the venom of the monocle cobra (*Naja kaouthia*), is a structural and functional analog of complement component C3. Like C3b, the active form of C3, CVF can form a C3 convertase with factor B. The screening of cDNA libraries prepared from cobra venom gland mRNA revealed the presence of a second cobra venom factor, named CVF2. The cDNA sequences of CVF (now named CVF1), CVF2, and cobra C3 are approximately 92% identical.

As an approach to understand the molecular basis for the structural and functional homology and difference between CVF1 and C3, the CVF1 gene structure was elucidated using a genomic cosmid library, genomic PCR, and genome walking. The CVF1 gene consists of 40 exons and 39 introns. Exon sizes vary from 46 bp to 1,140 bp, and introns vary from 100 bp to 9,000 bp. Whereas the CVF1 gene is with greater than 89 kb more than twice as large as the human C3 gene (42 kb), the structures of both genes exhibit a very high degree of conservation. More than half of the corresponding exons show identical sizes. Minor variations in exon sizes of 3 bp to 15 bp represent multiples of a codon. Intron phases are almost identical. In contrast, intron sequences and sizes show no homology between the CVF1 and human C3 genes. The most important difference is that the conserved exons 31 and 32 are contiguous in CVF1 and not separated by an intron, which is in contrast to the human C3, C4, and C5 gene structures.

As is the case for the CVF1 gene, exons 31 and 32 in the CVF2 and cobra C3 genes are also not separated by an intron. Intron sizes vary between the three genes. Exon 33 of the CVF1 gene is 9 bp smaller than the corresponding exons of the CVF2 and cobra C3 genes. The exon size difference is caused by a single nucleotide substitution in the intron splice site rather than by deletion.

Moreover, the promoter regions of the three homologous genes were analyzed in more detail to find the molecular basis for expression in different tissues. Putative TATA boxes were found in all three genes at conserved positions. However, the upstream regions of the three genes differ in putative binding sites for transcription factors. In addition to the genes for CVF1, CVF2, and cobra C3, some results indicate that the cobra genome may contain two more homologous genes or pseudogenes. These findings are consistent with multiple C3 genes in teleost fish and multiple toxin genes in various *Naja* species.