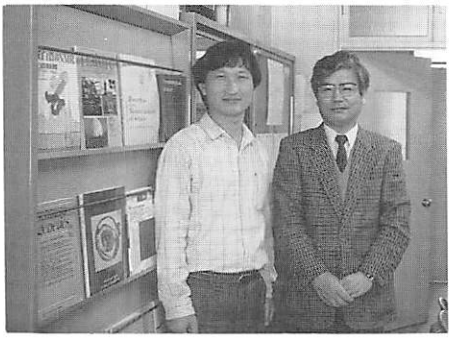
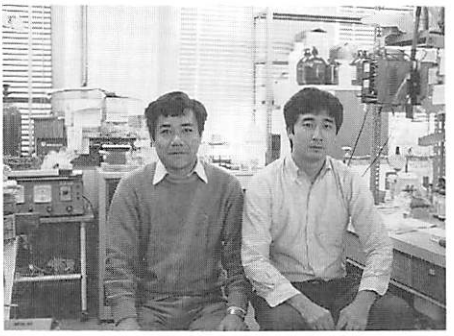
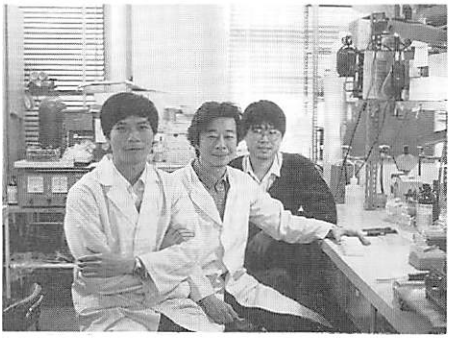
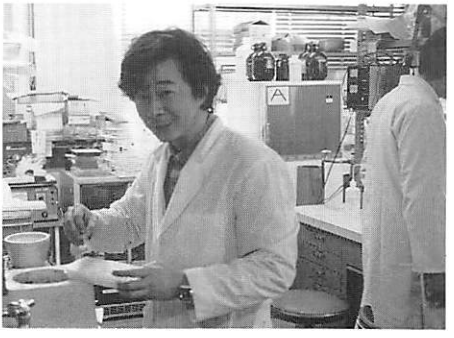
SCIENTIFIC REPORTS

Unpublished data should not be cited without permission of the authors.

Immunobiology

DEPARTMENT OF IMMUNOBIOLOGY

GENERAL SUMMARY

Since a few years ago, our research has been centered on elucidation of the molecular genetic mechanisms controlling the biosynthesis of complement proteins. We are particularly concerned with the regulatory mechanisms for the fourth complement component(C4) of the mouse and its isotype (sex-limited protein: Slp), since C4 plays a central role in the classical activation and the regulation of these proteins offers many fascinating genetic aspects in terms of the serum variation, hormone-dependence and tissue specificity. In man, many hereditary disorders are associated with the genetic defects of the expression of C4 genes. It is interesting to note that almost half of the HLA haplotypes associated with SLE-like disorders due to C4 deficiency show no major defect or rearrangements of the C4 genes. Clearly a considerable portion of C4 deficiency arises from the genetic defects in the regulation of the complement genes. We are therefore interested in identification and characterization of the genetic machinery controlling the expression of mouse C4 genes, because many expression variants are available among inbred mouse strains in terms of the expression of the C4 and Slp genes which provide us with an ideal model system to study gene regulation. When we isolated cDNA clones for mouse C4 and Slp in 1983 and published the experimental results in 1984 (Proc. Natl. Acad. Sci. USA, *81*, 6822-6826) we were far behind the two leading laboratories in this research field. Since then we have been successful in isolating the genomic DNA for C4 and Slp genes from several mouse strains and in characterizing the regulatory regions of these genes, now probably leading the field. We are still pursuing this very interesting as well as very important research.

In 1986, two principal researchers of this laboratory moved to the United States for furthering their study. In April, Dr. Masaru Nonaka joined Dr. H. Colten's lab (Department of Pediatrics, Washington University School of Medicine), while Dr. Shunnosuke Sakai started to work in November with Dr. C. H. Shreffler in the Department of Genetics of the same Medical School. Both of them are apparently making great progress in their study in St. Louis and are expected to come back to this lab in the spring of 1987 with the fruits of their study. During their absence, we have welcomed three students as graduate students or research fellows from Asian countries. Mr. Sa-nga Pattanakitsakul arrived from Thailand in October of 1985; Mr. Lai Jing Erh from Taiwan started to work in this lab in April, 1986; Mr. Huang Zhu Ming from China arrived here in October of 1987. In turn, Dr. Toshiaki Asahi left this lab for Australia to study in Melbourne University for one year in September, 1987.

(1) Complete nucleotide and derived amino acid sequences of the fourth component of mouse complement (C4)

M Nonaka, K. Nakayama, D. Y. Yu and M. Takahashi

The fourth component of complement (C4) is a serum glycoprotein composed of three disulfide-linked polypeptide chains, but it is translated as a single chain precursor before it is secreted. In addition to the hemolytically active C4, a testosterone-regulated isotype, sex-limited protein (Slp), has been identified in the mouse. Mouse C4 and Slp display extensive structural homology and both appear to undergo essentially identical intracellular processing. Neverthless, Slp has no known function probably because it is not susceptible to enzymatic cleavage by activated Cls. In the human, two isotypes of C4 (C4A and C4B) have been identified. Two mouse C4 genes (*C4* and *Slp*) and two human C4 genes (C4A and C4B) have evolved from single common ancester genes by the process fo gene duplication. Thus, structural studies on these genes serve as an appropriate model for the evolution of homologous genes that result in the multigenic family. Previously, we isolated cDNA and genomic DNA clones encoding for mouse C4 and Slp (1,2). As part of our studies for clarifying the structural differences between mouse C4 and Slp genes, we undertook the determination of the complete nucleotide sequence for mouse C4 and Slp cDNA.

The nucleotide sequence coding for the fourth component of mouse complement (C4) was determined from a cloned genomic DNA fragment and a cloned cDNA fragment. The amino acid sequence of the protein was deduced. The single chain precursor protein (pro-C4) consists of 1719 amino acid residues. The mature β, α, and γ subunits contain 654, 766, and 291 amino acids, respectively. One potential carbohydrate attachment site is predicted for the β chain, three for the α chain, and none for the γ chain. From a comparison with human C4 cDNA sequence an extensive overall sequence homology, 79% in nucleotides and 76% in amino acids, is observed. There is conservation in both the position and number of cysteine residues in human and mouse C4. We compared the mouse C4 amino acid sequences with those of mouse C3 and human α_2-macroglobulin and the evolutionary relationship among these three proteins is descussed.

1) M. Nonaka, M. Takahashi, S. Natsuume-Sakai, M. Nonaka, S. Tanaka, A. Shimizu and T. Honjo: Proc. Natl. Acad. Sci. *81*: 6822-6826, 1984.
2) M. Nonaka, K. Nakayama, D. Y. Yu, A. Shimizu and M. Takahashi: Immunol. Rev. *87*: 81-99, 1985.

(2) Complete nucleotide and derived amino acid sequences of sex-limited protein (slp), nonfunctional isotype of the fourth component of mouse complement (C4)[1]

M. Nonaka, K. Nakayama, D. Y. Yu and M. Takahashi

The nucleotide sequence coding for sex-limited protein (Slp), the testosterone-regulated isotype of the fourth component of mouse complement (C4), was determined from cloned genomic DNA and cDNA fragments. The complete deduced amino acid sequence for the single chain precursor protein of Slp (pro-Slp) consists of 1716 residues. The mature β, α, and γ subunits contain 654, 763, and 291 amino acids, respectively. One potential carbohydrate attachment site is predicted for the β-chain, five for the α-chain, and none for the γ-chain. From a comparison with the mouse C4 sequences, an extensive overall sequence homology, 96.0% in nucleotides and 94.2% in amino acids, is observed (Table I). Only one deletion/inserton event is recognized between C4 and Slp sequences: three residues near the Cls cleavage site are deleted from Slp. The distribution of cysteine residues is completely conserved between pro-Slp and pro-C4.

In addition to the coding sequences, the 3′ untranslated regions of mouse C4 and Slp cDNA show extensive homology. The polyadenylation signal, AATAAA, was found in the identical position, 77bp downstream from the stop codon in both C4 and Slp. Only six nucleotide substitutions were recognized in this 77bp region (92% homology). However, the poly-A tail of the Slp cDNA starts 20 bp further upstream in comparison to the C4 cDNA. This may imply that the C4 and Slp genes each have specific polyadenylation sites. Altenatively, the difference may reflect the clone-dependent variation for the start point of the poly-A tail.

Table 1. *Nucleotide and amino acid sequence homology among mouse C4, Slp and human C4*

Comparison	% homology	
	Nucleotide	Amino acid
Human C4A vs. mouse C4	79	76
Human C4A vs. Slp	78	74
Mouse C4 vs. Slp	96	94

(3) Identification of the 5'-flanking regulatory region responsible for the difference in transcriptional control between mouse complement C4 and Slp genes

M. Nonaka, H. Kimura, D. Y. Yu, S. Yokoyama, K. Nakayama and M. Takahashi

The fourth component of complement (C4) plays a pivotal role in the activation of the classical complement pathway. Two C4 isotypes have been described in humans (C4A amd C4B) and in mice (C4 and Slp), and their structural genes are closely linked to each other and to the C2 and factor B genes in the major histocompatibility complex (MHC). Both human C4A and C4B proteins are active as complement components, although some difference in reactivity has been reported. In mouse, however, only C4 shows complement activity, while Slp is hemolytically inactive and has no known function.

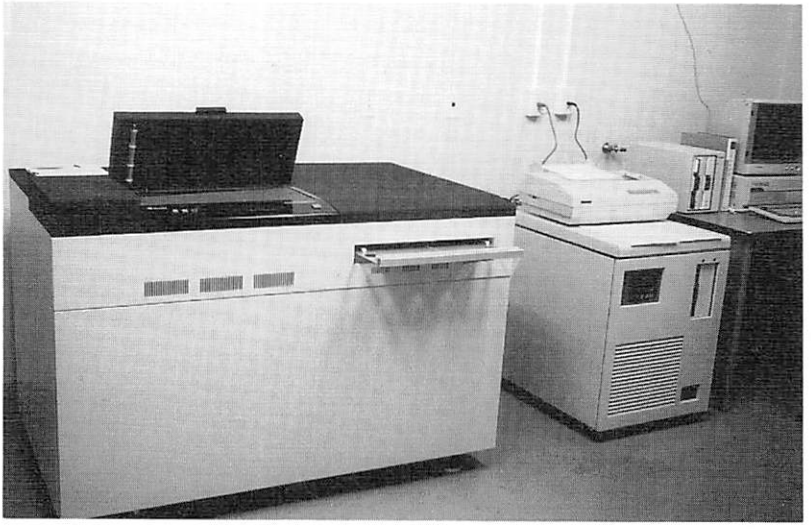
The mode of gene expression of mouse C4 and Slp shows striking contrast. Whereas C4 is constitutively expressed in liver and macrophages in male and female mice of all the strains tested, Slp is produced only in liver of male adult mice of certain limited strains. In most strains Slp is not expressed at all.

To elucidate the molecular basis underlying the difference in the mode of gene expression between mouse complement C4 (constitutive) and sex-limited protein (Slp) (testosterone-regulated), we compared nucleotide sequences and transcriptional regulatory activities of the 5'-flanking regions of these two genes. Although the two sequences showed a high degree of overall homology (95%) up to 1.9kilobases (kb) upstream from the transcription initiation site, the Slp sequence lacked a 31-nucleotide segment containing ACACCC repeats and a 60-nucleotide segment containing ACAC repeats, which are present, respectively, 1.6kb and 200 base pairs (bp) upstream from the transcription initaiation site of the C4 gene. When assayed in human hepatoma-derived HepG2 cells, the 1.8-kb 5'-flanking DNA fragment of the C4 gene demonstrated strong transcriptional acrivity, whereas the corresponding DNA fragment of the Slp gene showed only negligible activity. By progressive deletion experiments, it was shown that the difference in the constitutive transcriptional activity of the C4 and Slp genes was accounted for by the presence or absence of the putative regulatory domain located between 1700 bp and 400 bp upstream from the transcription initiation site.

(4) Recombination of two homologous MHC Class III genes of the mouse (C4 and Slp) that accounts for the loss of testosterone dependence of sex-limited protein expression

K. Nakayama, M. Nonaka, S. Yokoyama, D.Y. Yu, S. Pattanakitsakul and M. Takahashi

In most mouse strains, expression of the gene encoding sex-limited prtein (Slp), an isotype of the fourth component of complement (C4), is induced by testosterone, or the gene is not expressed at all; however, in some wild-derived strains carrying H-2^{W7}, H-2^{W16}, or H-2^{W19} haplotype, Slp is expressed constitutively in the same way as C4. To examine the structural basis for the testosterone-independent expression of Slp, 41 overlapping clones together encoding the S region were isolated from C3H.W7 mouse (H-2^{W7}) cosmid library. Five C4-related genes each spanning approximately 16 kb were identified among the cluster of cosmid clones and were isolated for structural study. One of the genes ($C4^{W7}$) hybridized with the C4-specific oligonucleotide probe but not with the Slp-specific oligonucleotide probe, whereas the other genes (Slp^{W7a}, Slp^{W7b}, Slp^{W7c}, and Slp^{W7d}) hybridized only with the Slp-specific probe. Restriction mapping of these genes and sequencing of the selected 5'-flanking regions of the genes were performed, and the results were compared with the data obtained with the C4 and Slp genes of FM (H-2^{d}) and BIO.BR (H-2^{K}). These studies showed that three of the C4-related genes of C3H.W7 (Slp^{W7b}, Slp^{W7c}, and Slp^{W7d}) are C4-Slp recombinant genes comprising a 5'-region derived from C4 gene and a 3'-region derived from Slp gene. It is suggested that 5'-flanking region derived from C4 in these C4-Slp recombinant genes accounts for testosterone-independent expression of Slp in C3H.W7 mouse. More recent studies have shown that the three C4-Slp recombinant genes of H-2^{W7} have distinctive recombination sites, but all of these sites clusters within about 3 kb of the central region of the genes, suggesting the genetic mechanisms facillitating the unequal crossingover during meiosis (S. Pattanakitsakul et al. manuscript in preparation).



Supersensitive Radio-image Analyzer