DNA sequence analysis of the rat $RT1.B_\alpha$ gene

Paul A. Barran and W. Robert McMaster

Department of Medical Genetics, University of British Columbia, Vancouver, B.C., Canada V6T 1W5

Abstract. The major histocompatibility complex (MHC) of the rat ($RT1$) complex encodes two sets of class II molecules referred to as RT1.B and RT1.D. The $RT1.B_\alpha$ gene was isolated for a Sprague-Dawley (RT1b) rat genomic library using a rat RT1.B chain cDNA as a hybridization probe. The coding and the majority of the intron DNA sequence was determined. The structure of the $RT1.B_\alpha$ gene is equivalent to that of $H-2$ and $HLA \alpha$ chain genes. Comparison of the nucleotide and predicted amino acid sequences of the $RT1.B_\alpha$ gene with those of the $H-2$ and $HLA$ genes revealed a high degree of overall sequence conservation. However, two regions of the first external domain (c1), residues 19-23 and 45-78, exhibit marked sequence variation. Two blocks of conserved nucleotide sequence were identified in the 5' promoter region of the $RT1.B_\alpha$ gene that have been described in all MHC class II genes sequenced to date. These conserved sequences may be involved in the coordinate regulation of expression of class II genes. The cloned $RT1.B_\alpha$ gene was efficiently transcribed when transfected to mouse L cells.

Introduction

The major histocompatibility complex (MHC) encodes cell-surface glycoproteins which are involved in immune regulation in all mammalian species. The MHC class I and class II molecules function in the associative recognition of foreign antigen by T lymphocytes (Klein et al. 1983). The rat MHC ($RT1$ complex) encodes two sets of cell surface class II molecules referred to as RT1.B and RT1.D, which are homologous to the H-2 class II A and E molecules, respectively (Blankenhorn et al. 1983). RT1.B and D class II molecules are composed of two noncovalently bonded polypeptide chains referred to as $B_\alpha B_\beta$ and $D_\alpha D_\beta$. The RT1 class II $\alpha$ and $\beta$ chains have apparent molecular masses of 30,000 and 24,000, respectively (Fukomoto et al. 1982).

The MHC shows the highest degree of polymorphism of any known mammalian genetic system. This polymorphism appears to have been selected for during evolution and is related to the function of MHC molecules. Comparison of the genes encoding these molecules among closely related species, such as rat and mouse, which diverged about 8 million years ago, with more distantly related species, such as rodent and man, which diverged about 80 million years ago, and comparisons of alleles within a species may reveal mechanisms for the generation and maintenance of polymorphism. The structure of MHC class II genes has been shown to be highly conserved during evolution (Wallis and McMaster 1984, Eccles and McMaster 1985, Robertson and McMaster 1985, Figueroa and Klein 1986). This article describes the isolation and nucleotide sequence of the $RT1.B_\alpha$ chain gene and compares the RT1 sequence with that of an allelic $RT1.B_\beta$ cDNA as well as with the corresponding $H-2$ and $HLA \alpha$ chain sequences.

Materials and methods

Isolation and characterization of genomic clones. A Sprague-Dawley rat ($RT1^b$) genomic DNA library, constructed with Hae III-digested liver DNA and the $\lambda$ vector Charon 4A (L. Jagodzinski, unpublished data), was screened with the cDNA insert of pRIa.2 which encodes a $RT1.B_\alpha$ chain (Wallis and McMaster 1984). Screening was carried out by plaque hybridization (Woo 1979) using the 782 bp insert of pRIa.2, radioactively labeled by nick translation (Rigby et al. 1977). Hybridizations were carried out in 6 x SSC at 68 °C and filters washed at 68 °C in 1 x SSC (1 x SSC: 0.015 M NaCl, 0.0015 M sodium citrate, pH 7.6). Recombinant plagues that hybridized to the cDNA probe were analyzed by Southern blot procedures (Southern 1975) using the labeled cDNAs pRIa.2 and pACD3 (which encodes the entire H-2 $A_\alpha$ chain (Benoiest et al. 1983)).

DNA sequence determination. Random DNA fragments 200-500 bp in length were generated from pRT1B.4 and pRT1B.5 by restriction endonuclease digestion with Alu I, Rsa I, Hae III, and Sau 3A, or by sonication essentially as described by Deininger (1983). Fragments were size fractionated, made blunt ended with mung bean nuclease (Kroeker and Kowalski 1978), and ligated to Smal I-digested M13 mp8 (Messing 1983). Subclones containing RT1.B chain coding sequence were identified by
Southern hybridization with the cDNAs pRIa.2 and pACD3 as described above and sequenced by the deoxy chain termination procedure (Sanger et al. 1977). Specific DNA fragments were isolated from pRT1B.4 and subcloned into M13 mp9, mp18, and mp19 to complete the sequence. The computer programs of Staden (1980) and Delaney (1982) were used to compile the sequence data.

**DNA transfection and selection.** DNA transfections were carried out essentially as described previously (Graham and Van der Eb 1973, Eccles et al. 1986). Thymidine kinase (TK)-deficient mouse L cells were maintained in Dubecco’s modified Eagle’s medium (DMEM) plus 10% fetal calf serum (FCS) (Gibco, Burlington, Ontario, Canada). One day prior to transfection, 8 x 10⁵ cells were plated per 10 cm petri dish. An aliquot (0.5 ml) of a calcium phosphate precipitate, prepared as described (Graham and Van der Eb 1973) using 10 µg phage or cosmid DNA and 100 ng of plasmid pOPF DNA which contains the herpes simplex (HSV) TK gene (Grosfeld et al. 1982), was added to each dish. The cells were incubated for 6 h and then the DNA precipitate was removed and 5 ml 15% glycerol in phosphate-buffered saline added for 1 min. The cells were washed with DMEM and 10% FCS and cultured for 2 days. The medium was then replaced with DMEM and 10% FCS containing hypoxanthine, aminopterin, and thymidine (HAT). TK-positive colonies appeared after 2–3 weeks growth and individual colonies were removed and grown as clones.

**Results and Discussion**

**Nucleotide sequence analysis.** The RTIα genomic library was screened with the cDNA insert of pRIa.2 and from a screen of 1 x 10⁶ recombinant phage, three strongly hybridizing clones were identified. One clone, L7, was shown by hybridization to contain the entire RT1.Bα gene in two Eco RI fragments. These two fragments were subcloned into the plasmid vector pUC9 and were designated pRT1B.4 and pRT1B.5. Restriction maps of L7, pRT1B.4, and pRT1B.5 are shown in Figure 1A. The nucleotide sequences of the coding regions of the RT1.Bα gene were determined according to the strategy outlined in Figures 1B and C with the resulting nucleotide and predict-

**Fig. 2.** The nucleotide and predicted amino sequence of the RT1.Bα gene. The numbers correspond to the amino acid positions in the mature protein. Sequences underlined are (5' to 3'): putative CAAT box, conserved regulatory elements, putative TATA box, and the polyadenylation signal. † † † indicates 5' conserved sequences found in other MHC class II genes (Okada et al. 1985)