Available online at www.sciencedirect.com
SCIENCE DIRECT
BioSystems 82 (2005) 266–272
BioSystems
www.elsevier.com/locate/biosystems

Evolutionary history of the retinoblastoma gene from archaea to eukarya

Masaharu Takemura *

Department of Human Functional Genomics, Life Science Research Center, Mie University, Tsu, Mie 514-8507, Japan

Received 9 June 2005; received in revised form 19 August 2005; accepted 19 August 2005

Abstract

The retinoblastoma gene product (Rb protein) has a role in progression through the cell cycle, regulating the activities of several transcription factors such as E2F. Since its functional loss results in impaired differentiation in the nervous, hematopoietic, and muscular systems, the Rb protein is very important for cell regulation in multicellular eukaryotes. To gain an insight into the evolutionary history of the Rb gene, I have compared the amino acid sequences of Rb proteins in multicellular eukaryotes and unicellular organisms including yeast, archaeotes, and viruses. Two short amino acid sequences, in the N-terminal and pocket A regions of human Rb protein, found to be well conserved, also in a single protein of *Saccharomyces cerevisiae*. These sequences were also found in proteins of two archaeotes, *Archaeoglobus fulgidus* and *Methanococcus jannaschii*. Further, the most conserved sequence in the pocket B region among multicellular eukaryotic Rb proteins was also conserved in several poxviruses. From these data, I conclude that the pocket A and B regions, backbones of the Rb protein, are derived from different organisms, respectively, the ancestors of archaeote and poxvirus, and that the ancestral pocket B region has been lost during evolutionary history of unicellular eukaryotes.

Keywords: Retinoblastoma gene; Rb; Poxvirus; Symbiosis; Molecular evolution

1. Introduction

Functional loss by deletion mutation of the retinoblastoma susceptibility (Rb) gene was first found in cases with retinoblastoma, a familial childhood intraocular tumor, and subsequently in association with other human tumors such as osteosarcoma (Riley et al., 1994; Weinberg, 1995). The Rb gene product, Rb protein, forms inactive complexes with E2F, a family of transcription factors comprising E2F-1 through -6, presumably inhibiting the formation of the basal transcription complex, resulting in repression of cell growth (Dyson, 1998). Therefore it has been termed a cell growth suppressor or tumor suppressor.

Phosphorylation and dephosphorylation of cellular proteins are recognized as important mechanisms regulating many eukaryotic cell events (Hunter, 1987). In response to growth signals from the extra cellular environment, cyclin-dependent kinases (cdk's) are induced to phosphorylate Rb protein at multiple sites. The resulting hyper-phosphorylated Rb protein does not bind to E2F, so that the Rb-E2F complex dissociates to generate free and active E2F (Nevins et al., 1997).

The Rb gene product has a role not only in the regulation of cell cycle, but also in cellular differentiation in multicellular eukaryotes. Functional inactivation of both copies of the Rb gene results in incorrect differentiation in the nervous and hematopoietic systems of

* Tel.: +81 59 231 5390; fax: +81 59 231 5388.
E-mail address: takemura@gene.mie-u.ac.jp.

0303-2647/$ – see front matter © 2005 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.biosystems.2005.08.005

Rb-deficient mice (Clarke et al., 1992; Jacks et al., 1992; Lee et al., 1992). The Rb protein directly binds to MyoD, a muscle-specific basic-helix-loop-helix (bHLH) regulator protein, and Rb protein inactivation inhibits myogenesis (Gu et al., 1993). Other than in mammals such as *Homo sapiens* (human) and *Mus musculus* (mouse), Rb homologues have been reported in *Gallus gallus* (chicken), *Notophthalmus viridescens* (newt), *Xenopus laevis* (frog), *Oncorhynchus mykiss* (trout), *Drosophila melanogaster* (fruit fly), *Arabidopsis thaliana* (cress), *Zea mays* (maize), *Pisum sativum* (pea), *Caenorhabditis elegans* (nematode) (Bernards et al., 1989; Boehmelt et al., 1994; Brunelli and Thorgaard, 1999; Destree et al., 1992; Du et al., 1996; Kong et al., 2000; Lee et al., 1987; Lu and Horvitz, 1998; Shimizu and Mori, unpublished data; Tanaka et al., 1997; Xie et al., 1996). However, Rb-like genes have not been hitherto identified in unicellular eukaryotes such as yeast and protozoans. The results suggest that the Rb gene is indispensable for cellular differentiation, a phenomenon specific to multicellular organisms, and that it is not necessary for the life of unicellular organisms.

To obtain insights into the evolution of the Rb gene, I have compared amino acid sequences of the products between animalia and plantae, and focused on highly conserved sequences, in the pocket A and N-terminus regions. Interestingly, I have found these sequences also in a protein of *Saccharomyces cerevisiae*, whose function is not known. Similar sequences were also detected in archaea species, *Archaeoglobus fulgidus* and *Methanococcus jannaschii*. Furthermore, a sequence homologous with a well-conserved sequence in the pocket B region of eukaryotic Rb proteins was also found in several poxviruses. In the present study, I performed a comparative study of Rb, Rb-like, and ancient Rb genes using sequences from GenBank, to approach the evolutionary branching of multicellular organisms from unicellular organisms on a molecular basis.

2. Materials and methods

The nucleotide sequences of multicellular eukaryotic Rb genes were obtained from GenBank: *H. sapiens*, M15400; *Mus musculus*, M26391; *Gallus gallus*, X72218; *Notophthalmus viridescens*, Y09226; *Xenopus laevis*, AAB23173 (peptide sequence); *Oncorhynchus mykiss*, AF102861; *Drosophila melanogaster*, Q24472 (peptide sequence); *Arabidopsis thaliana*, AF245395; *Zea mays*, X98923; *Pisum sativum*, AB012024; *Caenorhabditis elegans*, AF116529.

For these multicellular eukaryotes, homology and maximum matching searches of amino acid sequences of Rb or Rb-like proteins were performed using the DNASIS (Hitachi Software Engineering Co. Ltd., Japan) program.

Identification of a protein carrying a sequence homologous for the Rb protein in *Saccharomyces cerevisiae* was performed at the Saccharomyces Genome Database (http://www.yeastgenome.org/) using the BLASTP program (Karlin and Altschul, 1990). Identification of proteins possessing this sequence in *Archaeoglobus fulgidus*, *Methanococcus jannaschii*, *Escherichia coli*, and *Bacillus subtilis* was performed at the Institute for Genomic Research, Comprehensive Microbial Resource (http://tigrblast.tigr.org/cmr-blast/), again with the BLASTP program (Karlin and Altschul, 1990).

The complete amino acid sequences of virus proteins were obtained from GenBank: fowlpox virus, NC_002188; melanoplus sanguinipes entomopoxvirus, NC_001993; vaccinia virus, NC_001559; variola virus, X69198; Yaba-like disease virus, AJ293568; beet curly top virus, NC_001412; ebola virus, NC_002549; adenovirus type 2, J01917; human herpesvirus type 1, NC_001806.

3. Results

Among vertebrates, amino acid sequences of Rb proteins are highly conserved. Rates of maximum matching of amino acid sequences compared with *H. sapiens* are: *Mus musculus*, 90%; *Gallus gallus*, 72%; *Notophthalmus viridescens*, 59%; *Xenopus laevis*, 59%; *Oncorhynchus mykiss*, 56%. On the other hand, those of insects, plants, and worms are lower: *Drosophila melanogaster*, 29%; *Arabidopsis thaliana*, 30%; *Zea mays*, 27%; *Pisum sativum*, 30%; *Caenorhabditis elegans*, 28%. Highly conserved amino acid sequences between vertebrates and the latter species presumably are very important for Rb function and thus also for estimation of the molecular evolution of the Rb gene.

Fig. 1 shows amino acid alignments of Rb pocket A and B regions for four species, *H. sapiens* (vertebrate), *Arabidopsis thaliana* (plant), *Drosophila melanogaster* (insect), and *Caenorhabditis elegans* (worm). An especial focus was on a sequence whose position is a.a. 530–539 in *H. sapiens*, indicated by box 1 in Fig. 1A. This sequence in *H. sapiens* (KVIESFIKAE) is well conserved in the other three species, especially in *Arabidopsis thaliana* (KVIESFIRHE) (Figs. 1A and 2). According to a previous report, this sequence is part of the α8 helix in the pocket A region, which is a component of the A–B interface (Lee et al., 1998). Interestingly, it was found that a homologous sequence exists in a protein of *Saccharomyces cerevisiae* (KLIESFISKAE) (Fig. 2, left alignment). The function of this protein, encoded by a gene named YLR419W, is not yet fully understood, but it was predicted to be a member of DEAH-box family of RNA helicases, which are responsible for RNA processing such as mRNA splicing (De la Cruz et al., 1999).