NOTE TO USERS

uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Xiuzhong Zheng
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Biochemistry)
GRADE / DEGREE

Department of Biochemistry, Microbiology and Immunology
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Definition of Ku-interaction domains in RAG1 and RAG2 proteins in V(D)J recombination

TITRE DE LA THÈSE / TITLE OF THESIS

Robert Haché
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Majambu Mbikay

Martin Pelchat

Gary W. Slater
LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES /
DEAN OF THE FACULTY OF GRADUATE AND POSTDOCORAL STUDIES

Definition of Ku-interacting domains in RAG1 and RAG2 proteins in V(D)J recombination

By

Xiuzhong Zheng

Thesis submitted to the Department of Biochemistry,

Microbiology and Immunology in partial fulfillment of the

requirements of the degree of **Master of Science**

University of Ottawa

Ottawa, Ontario, Canada

July 2005

Your file Votre référence
ISBN: 0-494-11477-0
Our file Notre référence
ISBN: 0-494-11477-0

ABSTRACT

V(D)J recombination is a process that generates the diversity of the immune repertoire against foreign antigens. During B and T cell development, genes encoding immunoglobulins (Ig) and T cell receptors (TCR) variable region are somatically assembled by selective V (variable), D (diversity) and J (joining) segments pre-existing in the germline. Recombination-activating genes 1 and 2 (RAG1/RAG2), the lymphocyte-specific factors, initiate V(D)J recombination by nicking at the border the heptamer sequence of RSS (recombination signal sequence) and generate four DNA double stranded breaks (DSB) in the cleavage step. After cleavage, RAG1/RAG2 complex are still bound to the DNA ends. In the joining step, DNA breaks are processed and rejoined by non-homologous end joining apparatus, which includes Ku70/Ku80, DNA-PKcs, Artemis, XRCC4 and DNA ligase IV. However, how the cleavage step is linked to the joining step is not yet known.

Minimal or core regions of RAG1 and RAG2 have been delineated to be required for recombination. Here I report that, by using purified recombinant Ku70/Ku80 from insect cells that both core RAG1 and core RAG2 proteins can interact directly with Ku70/Ku80 and that this interaction is DNA independent.

In the present study, the Ku-interacting domain of RAG1 was initially determined by deletion mapping assay. Both N-terminus and C-terminus deletion mutants were in vitro translated as ^{35}S labeled proteins and were tested for binding to immunoprecipitated Ku70/Ku80 from whole cell extract of Jurkat T cells. I have determined that the region

spanning from amino acid residue 499 to 755, which corresponds to the central domain of RAG1, is involved in the Ku-RAG1 interaction. By expressing three peptides (RAG1 $_{499-763}$, RAG1 $_{573-763}$ and RAG1 $_{660-763}$) as maltose binding protein (MBP) fusion proteins and testing for binding to Ku, I have defined a minimal Ku-interaction domain in RAG1 as the region spanning from amino acid residues 660 to 763. In RAG2, results from the deletion mapping assay suggests that the region spanning from amino acid residue 132 to 242 is involved in the interaction with Ku.

My results suggest that a physical interaction between RAG1/2 and Ku antigen may help coordinate the cleavage stage of V(D)J with the non homologous DNA end rejoining of the mature sequences.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Robert J.G. Haché. I truly appreciate the insight, encouragement and guidance I received during my studentship in his laboratory. I would also like to thank Dr. Haché for his critical revision of my thesis. I would also like to express my thanks to my thesis advisory members, Dr. Martin Holcik and Dr. Jonathan Lee for their time and suggestions throughout the course of my research.

I am equally grateful to Dr. Caroline Schild-Poulter for her support, encouragement and guidance to my work, without which my research would not be complete. I also appreciate her time and help in the preparation of my thesis manuscript.

My sincere thanks to all the past and current members of the Haché laboratory. I would like to express my thanks to Amy Shih for her encouragement and technical advice throughout my research. A special thanks to Michelle Leleu-Evans for her truly help, encouragement and smiling face that brighten my life away from home.

I would like to appreciate the financial support and encouragement from my guarantor Mr. Edwin T. Cornelius, Jr. I would also like to extend my thanks to University of Ottawa for their financial support.

Finally, I am truly blessed to have such wonderful parents, whose love and invaluable support give special meaning to every journey in my life. I am sincerely grateful to them.

Abbreviations

DMEM	Dulbecco's Modified Eagle Medium
DNA-PKcs	catalytic subunit of DNA-dependent protein kinase
EMSA	electrophoretic mobility shift assay
FBS	fetal bovine serum
HMG1/2	high mobility group protein 1/2
Ig	immunoglobulin
IPTG	isopropyl-beta-D-thiogalactopyranoside
KARP-1	Ku autoantigen related protein 1
MBP	maltose binding protein
NBR	nomamer binding region
NE	nuclear extract
NHEJ	non-homologous end joining
NLS	nuclear localization signal
NRE1	negative regulatory element 1
OS	Omenn's syndrome
PBS	phosphate-buffered saline
PHD	plant homeodomain
PMSF	phenylmethylsulfonyl fluoride
RAG1/2	recombination activating protein 1/2

RSS	recombination signal sequence
SCID	severe combined immunodeficiency disease
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	signal end complex
TCR	T-cell receptor
TdT	terminal deoxynucleotidyl transferase
vWA domain	von willebrand factor A like domain
WCE	whole cell extract
ZDD	zinc-binding dimerization domain
ZFA/ZFB	zinc finger A/zinc finger B

Table of Contents

List of Figures

1. Introduction

1.1 V(D)J recombination overview

Adaptive immunity is exclusively found in jawed vertebrates and comprises two types of immunities: humoral immunity and cellular immunity. Humoral immunity is mediated by antibodies (immunoglobulins) produced on the surface of B lineage cells, while cellular immunity is directly mediated by T-cell receptors (TCR) on the surface of T cells (Blackwell and Alt, 1989). Each Ig or TCR is composed of a constant region that serves structural, signaling and effector functions, and a variable region that determines the recognition and specificity of antigen. The genes encoding variable region of both Ig and TCR are assembled by selected noncontiguous V (variable), D (diversity) and J (joining) gene segments in a site-directed DNA rearrangement process known as V(D)J recombination (Blackwell and Alt, 1989; Lewis, 1994; Sadofsky, 2001; Gellert, 2002; Jung and Alt, 2004b). A potential functional variable region gene is assembled when a V and J gene segments are joined (for IgL and TCRα and γ loci) and a V, D and J segments are joined (for IgH and TCRβ and δ loci) (Lewis, 1994).

V(D)J recombination is mediated through recombination signal sequence (RSS) flanking V, D and J gene segments. RSSs are composed of a highly conserved palindromic heptamer (CACAGTG) element and an AT-rich nonamer (ATAAAAACC) element, separated by a less conserved 12bp (12RSS) or 23bp (23RSS) spacer. Efficient V(D)J recombination only occurs between two RSSs with different spacer length, process that is termed the 12/23 rule (Tonegawa, 1983). V(D)J recombination can be

divided into two steps: cleavage and joining (Figure 1A and 1B) (Lewis, 1994; Fugmann et al., 2000a; Jung and Alt, 2004).

1.1.1 Cleavage step of V(D)J recombination

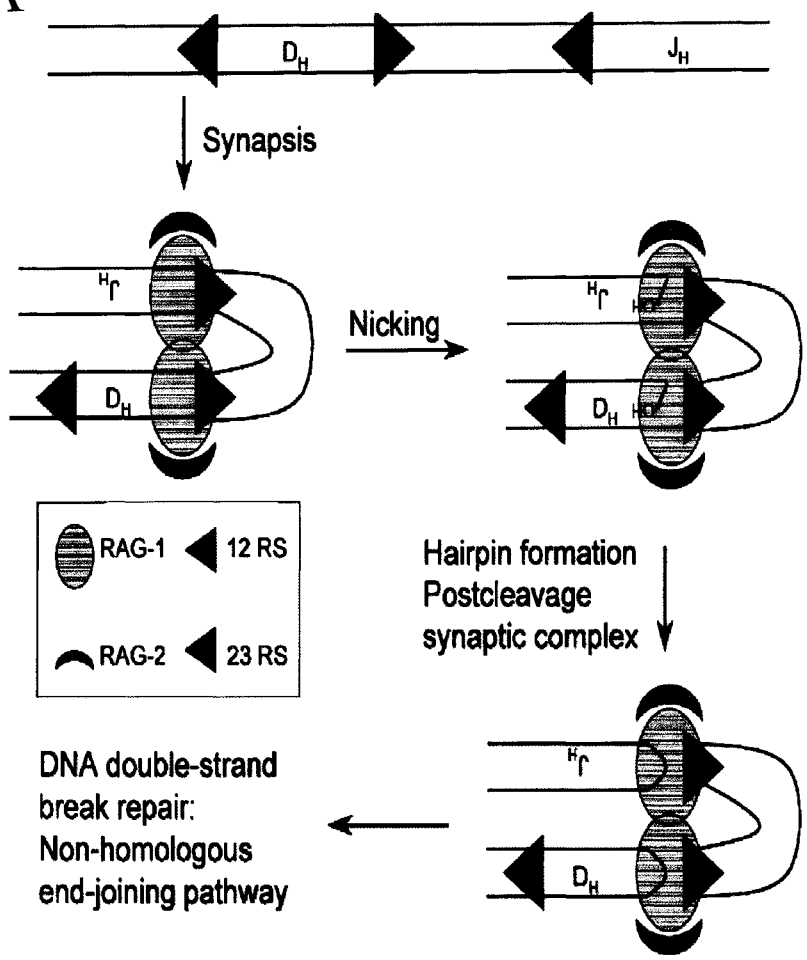
V(D)J recombination is initiated by the RAG1/RAG2 complex (Schatz et al., 1989; Oettinger et al., 1990) (see later in page **8** for details). RAG1 and RAG2 together act as a recombinase and recognize both 12RSS and 23RSS independently or both simultaneously (Hiom and Gellert, 1997; Akamastsu and Oettinger, 1998; Jones and Gellert, 2002; Mundy et al., 2002). Coupled cleavage only occurs when RAG1/RAG2 complex bound to a pair of complementary RSSs in a synaptic complex.

RSS cleavage occurs in two kinetic steps. First, RAG1/RAG2 complex uses water as a nucleophile to introduce a nick at the border of the coding flank and the heptamer of RSS. Nicking liberates a free 3' hydroxyl. In the second step, the free 3'OH is used as a nucleophile to attack the phosphodiester bond of the opposite strand at the RSS-coding border in a direct transesterification reaction which results in two 5'-phosphorylated blunt signal ends and two coding ends with covalently sealed hairpin structure (Oettinger et al., 1990; McBlane et al., 1995). Coding joints can eventually become mature coding regions for Igs and TCRs, while depending on germline arrangement, DNA fragment with signal joints may be retained in the chromosome or may be permanently lost from the genome.

V(D)J recombination occurs almost exclusively according to 12/23 rule in

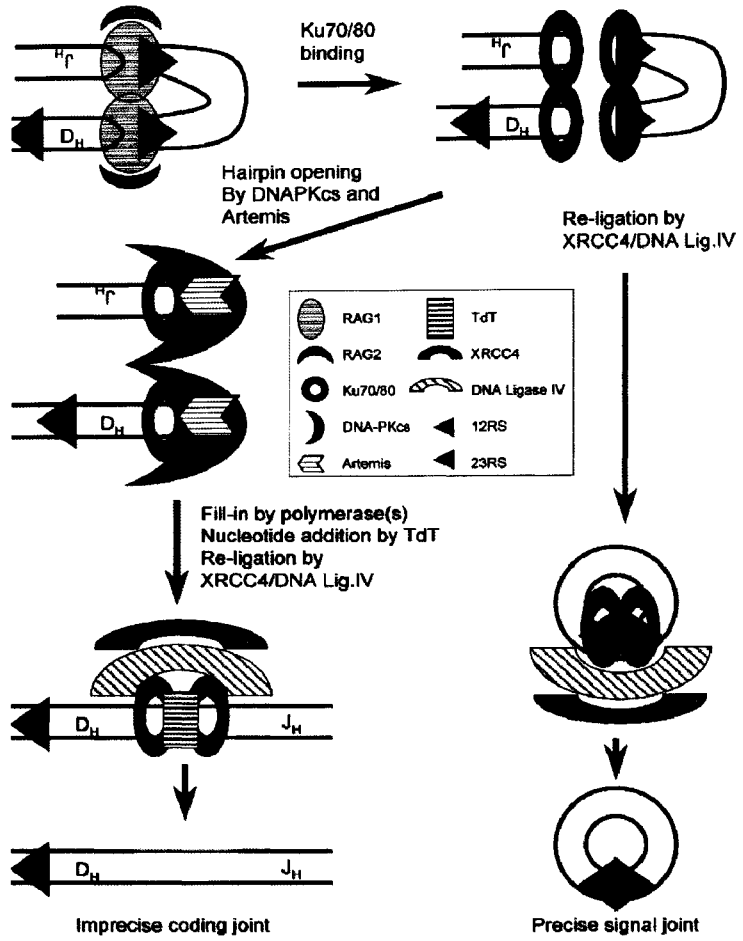
Figure 1. Schematic presentation of V(D)J recombination. V(D)J recombination is theoretically divided into two steps, cleavage and joining. Cleavage step is schematically shown in **A** and joining step is shown in **B**. **(A) The cleavage step of V(D)J recombination.** RAG1 and RAG2 proteins form a complex to initiate cleavage step by nicking at the border of RSS (recombination signal sequence) and coding flanks. 3' hydroxyl will attack the phosphodiester bond of the opposite strand at the RSS-coding border in a direct transesterification reaction, which results in two 5'-phosphorylated blunt signal ends and two coding ends with covalently sealed hairpin structure. Nicking can occur on a single RSS, however, hairpin structure can only occur in the synaptic complex, which is compose of a 12RSS and a 23RSS, termed V(D)J recombination 12/23 rule. **(B) The joining step of V(D)J recombination.** After cleavage, RAG1 and RAG2 proteins remain bound to the DNA ends in a post-cleavage complex. The DNA ds-breaks are processed and joined by the non-homologous end joining apparatus (NHEJ), which includes Ku70/Ku80 heterodimer, Artemis, DNA-PKcs, XRCC4, DNA ligase IV. Signal ends are relegated directly by XRCC4 and DNA ligase IV to form a precise signal joint. However, hairpin ends has to be opened first by Artemis and DNA-PKcs and then joined together with nucleotide loss and addition to be the coding regions for immunuoglobulin and T-cell receptor variable region. Figures A and B are introduced from Jung and Alt (2004).

A



A: Summary of RAG cleavage reaction

B



B: NHEJ pathway in V(D)J recombination

vivo. Although the mechanism enforcing the 12/23 rule remains poorly understood, it was shown to be imposed in the synapsis and coupled cleavage steps of V(D)J recombination in vivo and in vitro (Eastman et al., 1996; Steen et al., 1996; Hiom and Gellert, 1998; West and Lieber, 1998). RAG1/RAG2 complex was initially shown to follow the 12/23 rule in vitro (van Gent et al., 1996). High mobility group protein 1 or 2 (HMG1 or HMG2), ubiquitous abundantly expressed DNA-bending nuclear proteins, were also suggested to enhance the 12/23 rule (Bustin and Reeves, 1996; van Gent et al., 1997; Kim and Oettinger, 1998). Ku70/Ku80 and DNA-PKcs were recently found to modulate the RAG-mediated cleavage (Sawchuk et al., 2004). They were shown to enforce the 12/23 rule in combination with RAG complex and HMG1/2 by preferentially inhibiting both the 12/12 and 23/23 RSS cleavage. This suggests that Ku70/Ku80 and DNA-PKcs may play roles in precleavage complex before the joining step.

After cleavage, RAG1/RAG2 complex has been shown to associate with both coding ends and signal ends in a postcleavage complex in vivo and in vitro (Agrawal and Schatz, 1997; Hiom and Gellert, 1998; Jones and Gellert, 2001; Qiu et al., 2001; Yarnell Schultz et al., 2001; Huye et al., 2002; Tsai et al., 2002; Lee et al., 2004). The post-cleavage complex may recruit NHEJ joining factors (Leu et al., 1997; Ramsden et al., 1997; Lee et al., 2004); it may stabilize the joining procedure (Agrawal and Schatz, 1997; Hiom and Gellert, 1998); it may protect the ends from inappropriate insertion elsewhere in the genome (Fugmann et al., 2000a).

1.1.2 Joining step of V(D)J recombination

In vivo and in vitro experiments have shown that the joining step of V(D)J recombination is mediated by NHEJ apparatus, which includes Ku70/Ku80 heterodimer, DNA-PKcs, XRCC4 ,DNA ligase IV, Artemis and DNA polymerase μ (Bassing et al., 2002; Gellert, 2002; Ma et al., 2002; Mahajan et al., 2002; Bertocci et al., 2003; McElhinny and Ramsden, 2003).

1.1.2.1 Coding joint formation

Coding joints formation begins with hairpin end processing and is a complex and variable process. The opening of the hairpin end appears to be mediated by Artemis, recently identified factor of the NHEJ pathway in V(D)J recombination (Moshous et al., 2001). Artemis-deficient mice were found to be defective in coding joint formation with an accumulation of unresolved hairpin intermediates in thymocytes similar to the DNA-PK-deficient phenotype (Rooney et al., 2002). In vitro DNA-PKcs was shown to phosphorylate Artemis and form a complex with this phosphorylated Artemis, which leads to the activation of an endonuclease activity to nick the covalently sealed hairpin ends (Ma et al., 2002). Based on in vitro data and the Artemis-deficient phenotype, Artemis has been argued to be the major important hairpin-opening activity in V(D)J recombination although its function needs direct in vivo confirmation (Ma et al., 2002).

Coding joints are imprecise due to the addition or deletion of nucleotides after hairpin ends are opened, which leads to further "junctional" diversity in the variable

regions of the antigen specific receptors before ligation. There are two types of insertion of nucleotides in coding joints: non-templated (N nucleotide addition) and templated (P-nucleotide addition) (Lafaille et al., 1989; McCormack et al., 1989). The mechanism of the nucleotides loss from the coding end is much less known.

1.1.2.2 Signal joint formation

Signal joints formation mediated by two heptamer sequences end to end fusion is simple and usually precise. DNA-PKcs and Artemis were shown not to be required for signal joint formation (Bassing et al., 2002; Lieber et al., 2003). Although signal joint formation is simple, coding ends are processed and joined faster and signal ends are relatively long-lived in vivo. The mechanism for this is still unknown. However, in vivo and in vitro studies showed that, after coding joints formation, RAG proteins remain bound tightly to the synapsis signal ends, which was implicated to be cell-cycle regulated and restricted to the G_0/G_1 phase (Jiang et al., 2004). Furthermore, signal joints can also be recleaved by RAG proteins to regenerate the two signal ends, which may explain the abundance of signal ends in the lymphocyte precursors (Neiditch et al., 2002).

1.1.3 Regulation of V(D)J recombination

V(D)J recombination is tightly regulated with respect to developmental timing, lineage specificity and allelic exclusion . It is regulated at least by two means: one is by

the expression level of RAG1 and RAG2 proteins and the second is by the limited access of the recombination machinery to certain RSS sites.

Expression of RAG1 and RAG2 proteins is tightly limited to specific developmental subsets of B and T lymphocytes (Schatz et al., 1989; Oettinger et al., 1990; Grawunder et al., 1995; Monroe et al., 1999; Yu et al., 1999). The half-life of RAG1 has been shown to be prolonged in the absence of RAG2, suggesting that RAG1 is more prone to be degraded when present in the complex with RAG2 (Grawunder et al., 1996). The amount of RAG2 protein was found to vary greatly through the cell cycle, with high levels in G1 phase and low in the S, G2 and M phases without much change of RAG2 mRNA. The down regulation of RAG2 protein level was suggested to be controlled by phosphorylation followed by degradation (Lin and Desiderio, 1993). RAG2 preferentially accumulates in the G_0/G_1 phase, which could explain why V(D)J recombination and V(D)J breaks occur predominantly in the G_0/G_1 phase (Schlissel et al., 1993; Lin and Desiderio, 1995).

Although RAG1 and RAG2 expression limits V(D)J recombination activity to developing lymphocytes, they alone are not sufficient for the regulation. For example, fibroblasts in which RAG1 and RAG2 are expressed can rearrange artificial substrates but not any of their own antigen receptor loci. In addition, apart from the Ig and TCR loci, there are still many sequences at random in the chromosome alike RSS at which cleavage should be avoided (Gellert, 2002). Accessibility of gene segment was shown to be regulated by growth factors, chromatin remodeling complexes or transcriptional

factors such as E2A proteins (E12 or E47), Swi/Snf remodeling complex, OcaB and Pax5 (Bain et al., 1994; Zhuang et al., 1994; Lin and Grosschedl, 1995; Corcoran et al., 1998; Kwon et al., 2000; Romanow et al., 2000; Hesslein et al., 2003; Krangel, 2003).

1.2 RAG1 and RAG2 proteins overview

RAG1 and RAG2 proteins were identified by their capacity to synergestically confer V(D)J recombination to non-lymphoid cells in 1990's (Oettinger et al., 1990). RAG1 and RAG2 gene loci are the nearest neighbors ever characterized in vertebrates and are separated by 3-15kb within the same genomic locus, depending on different species. These genes are implied to have originated from an ancestral transposon and have been introduced into the vertebrate genome as the evolution of the vertebrate adaptive immune system (Agrawal et al., 1998; Hiom et al., 1998). Genes coding for RAG1 and RAG2 are both located on chromosome 11 in human and chromosome 2 in mouse (Ichihara et al., 1992; Oettinger et al., 1992; Sherrington et al., 1992). RAG1 cDNA clones predict RAG1 as a 119kDa protein and both human and mouse RAG2 protein are predicted to be 65kDa (Schatz et al., 1989; Oettinger et al., 1990; Lin and Desiderio, 1994). Both RAG1 and RAG2 coding sequences are highly conserved among a wide range of species that carry out V(D)J recombination (Schatz et al., 1989; Sadofsky et al., 1994).

Core RAG1 and core RAG2 were determined to be the minimal regions required for efficient RSS cleavage in V(D)J recombination (Sadofsky et al., 1993;

Cuomo and Oettinger, 1994; Sadofsky et al., 1994; Kirch et al., 1996). Compared to full length, core RAG1 and core RAG2 are more soluble and are more readily isolated (Cuomo and Oettinger, 1994; Sadofsky et al., 1994). Therefore, these truncated forms of RAG1 and RAG2 are largely employed in the studies of V(D)J recombination in vivo and in vitro. However, the non-core regions of RAG1 and RAG2 are evolutionary highly conserved and are important for V(D)J recombination activity in vivo. For instance, non-core region of RAG1 is important for IgH locus D_H-to-J_H rearrangement, whereas the C-terminus of RAG2 is important for V_H-to-DJ_H rearrangement (Roman et al., 1997; Kirch et al., 1998). In addition, non-core regions also may function to suppress the "non-standard" RAG activity, like to suppress the RAG-mediated hybrid joining reaction in which a signal end can join with a coding end that is created from a different RSS, and to suppress RAG-mediated transpositions in vitro (Sekiguchi et al., 2001; Elkin et al., 2003; Tsai and Schatz, 2003; Swanson et al., 2004).

Endogenous RAG1 and RAG2 proteins coexist in a complex in primary thymocytes as well as when expressed in adherent cells, where they are localized at the periphery of the nucleus (Spanopoulou et al., 1995). Both RAG1 and RAG2 have their own nuclear localization signal. Nuclear localization of RAG1 protein is contributed by four basic domains ($BI_{aa141-146}$, $BII_{aa218-236}$, $BIII_{aa243-249}$, $BIV_{aa826-840}$) and a nuclear localization signal in the fifth basic domain ($BV_{aa969-973}$) (Cortes et al., 1994; Cuomo et al., 1994). For RAG2, a region spanning from amino acid residue 491 to 527 in the C-terminal directs its nuclear localization (Corneo et al., 2002).

Either RAG1 or RAG2 deficient mice have a severe combined immunodeficiency with a complete lack of circulating T and B cells due to their inability to assemble antigen receptor gene segments and early block in both B and T cell maturation (Mombaerts et al., 1992; Shinkai et al., 1992). In addition, some mutations that alter the amino acid sequence of RAG1 or RAG2 have been shown to disturb B cell generation and to partially disturb T cell generation as unfolded in Omenn syndrome in humans (Schwarz et al., 1996; Villa et al., 1998; Villa et al., 1999).

1.2.1 The Primary sequence, domain and structure of RAG1

RAG1 protein is a relatively large protein with 1040 amino acids in mouse and with 1043 amino acids in human. The murine form of core RAG1 include amino acid residue from 384 to 1008 (Sadofsky et al., 1993) (Figure 2). Although both RAG1 and RAG2 are required to initiate V(D)J recombination, the catalytic active residues required for both nicking and hairpin formation were only found in RAG1, which includes two aspartate (D) at position 600 and 708, and a glutamate (E) at position 962, comprising a DDE motif which are thought to coordinate one or more divalent cations in the active site (Kim et al., 1999; Landree et al., 1999; Fugmann et al., 2000b; De and Rodgers, 2004).

1.2.1.1 Non- core regions of RAG1

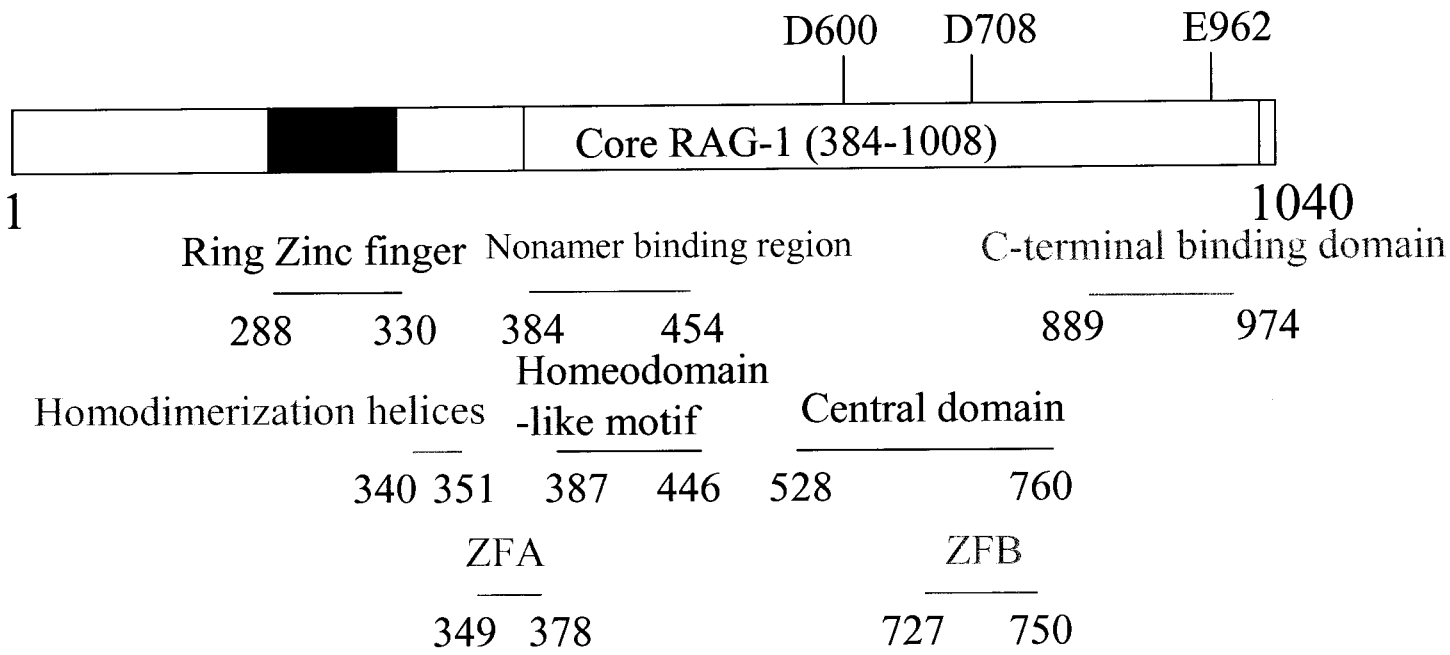
The non-core regions have been conserved throughout evolution and were found to influence the catalytic efficiency of V(D)J recombination and also potentially serve as accessory and/or regulatory functions (Noordzij et al., 2000; Santagata et al., 2000; Jones and Gellert, 2003).

The zinc-binding dimerization domain (ZDD) is the only domain in the non-core region of RAG1 that has been biophysically and structurally characterized (Rodgers et al., 1996; Bellon et al., 1997). ZDD, including a C3HC4 zinc-binding motif (zinc ring finger) that binds two zinc ions and a C2H2 zinc finger sequence (ZFA) that binds another zinc ion, is a highly specific homodimerization domain. Ring zinc finger has been shown to possess E3 ubiquitin ligase activity that promotes RAG1 autoubiquitylation at a highly conserved K residue K233 (Jones and Gellert, 2003; Yurchenko et al., 2003). However, the physiological substrate of ubiquitination activity catalyzed by RAG1 has yet to be established.

1.2.1.2 Nonamer binding region (NBR)

A N-terminal region of core RAG1 spanning from amino acid residue 384-454 is referred to as the nonamer binding region, which has been shown to recognize RSS nonamer sequence specifically by surface plasmon resonance and an in vivo one-hybrid system (Spanopoulou et al., 1995; Difilippantonio et al., 1996). It is the only well-characterized DNA binding domain in RAG1.The region spanning from amino acid

Figure 2. Schematic representation of RAG1 protein. Murine RAG1 protein is composed of 1040 amino acids. Core RAG1, which has been shown to be the minimal regions required for V(D)J recombination, is mapped to span from amino acid residue 384 to 1008. Two aspartate residues at position 600 and 708, and a glutamate at position 962 comprises a DDE motif, which are thought to coordinate one or more divalent cations in the active site. In the non-core region, there is a zinc-binding dimerization domain (ZDD) spanning from amino acid residue 265 to 380. ZDD, including a C3HC4 zinc-binding motif (zinc ring finger) spanning from amino acid residue 288 to 330 and a C2H2 zinc finger sequences (ZFA) spanning from amino acid residue 349 to 378 is a highly specific homodimerization domain in RAG1. Moreover, the zinc ring finger in RAG is also shown to posses an E3 ligase activity, which can autoubiquitinylate RAG1 itself and some other substrates that are not yet established. Core RAG1 is consisted of three DNA-binding region: nonamer binding region, central domain and C-terminal binding domain. NBR, spanning from amino acid residue 384 to 454 can bind to nonamer sequence of RSS with high specificity and affinity. The region spanning from amino acid residue 387 to 446 is required for nonamer binding and is shown to be most closely similar to the homeodomain of the Salmonella typhimurium Hin invertase. Amino acid residue from 528 to 760 is defined to be the central domain, which is involved in RSS-heptamer binding and RAG2-interaction. Within this domain, there is another zinc finger (ZFB), which has been shown to interact with RAG2. The C-terminal binding domain, which spans from amino acid residue 889 to 974 has been found to bind coding flank region in a non-sequence-specific manner cooperatively with high affinity and has been shown to be involved in the oligomerization of core RAG1.


D600
D708
E962
Core RAG-1 (384-1008)
1
1040
Ring Zinc finger
Nonamer binding region
C-terminal binding domain
288
330
384
454
889
974
Homeodomain
-like motif
Homodimerization helices
Central domain
340
351
387
446
528
760
ZFA
ZFB
349
378
727
750

residue 387 to 446 required for RSS nonamer binding of RAG1 has been highly conserved during evolution and is most closely similar to the homeodomain of the Salmonella typhimurium Hin invertase, which mediates flagellar variation (Simon et al., 1980). Deletions and point mutations in this homeodomain-like region disrupt RAG1 DNA binding activity and further functional analysis of RAG1 site-directed mutants revealed that NBR was also important for the joining step in V(D)J recombination, as R401/R402, E423, R440 were shown to exhibit a wild-type cleavage activity but severe joining defects in both coding and signal joint formation (Huye et al., 2002).

1.2.1.3 The central domain

The RAG1 central domain was identified to be a topologically independent domain by limited proteolysis and matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry (Arbuckle et al., 2001). It contains two active-site residues (D600 and D708) and also a classic C2H2 zinc finger motif, termed Zinc finger B (ZFB) (Rodgers et al., 1996). The isolated central domain was shown to be capable to bind specifically to double-stranded (ds) RSS heptamer sequence with very weak binding affinity. Analysis revealed that the central domain only shows a three-fold specificity for the heptamer sequence over nonsequence-specific DNA (Aidinis et al., 2000). ZFB was previously determined to be the predominant region for recruiting RAG2 protein to RAG1 and is involved in non-specific DNA binding (Aidinis et al., 2000; Arbuckle et al., 2001). In addition, ZFB was also found to make contribution to

RAG1-ssRSS binding, as central domain without ZFB binds to ssRSS inefficiently and with a three- or four-fold affinity difference (Peak et al., 2003). RAG1/RAG2 were shown to cleave more efficiently on ssRSS flanking double-stranded coding region and maintain in a more stable complex with nicked RSS (Cuomo et al., 1996; Ramsden et al., 1996; Grawunder and Lieber, 1997). Hence, it is proposed that the central domain plays an important role in the different affinity for different forms of RSS generated during V(D)J recombination.

The central domain has been implicated in different steps in V(D)J recombination, as point mutations in this domain has been found to be defective in DNA-binding, RSS nicking and joining, and lead to immunodeficiency diseases, such as SCID and Omenn's syndrome (Huye et al., 2002; Tsai et al., 2002).

1.2.1.4 The C-terminal domain

Similar to the central domain, the C-terminal domain was also identified to be a topologically independent domain in core RAG1 by limited proteolysis and MALDI-TOF mass spectrometry. The C-terminal domain was found to bind dsDNA in a nonsequence-specific manner cooperatively and with high affinity (Arbuckle et al., 2001). In addition, DNA-protein cross-linking studies has further defined a region from amino acid residue 889 to 974 as the coding flank binding region (Mo et al., 2001). The C-terminal domain was also found to be involved in self-association, which most likely contributes to the oligomerization of core RAG1, as core RAG1 has been found to form

a dimer or a tetramer in the complex with RSS sequence (Rodgers et al., 1999; Ciubotaru et al., 2003; Godderz et al., 2003).

Point mutations in the C-terminal domain has demonstrated that C-terminal domain was important for the joining step in V(D)J recombination; however, interestingly all those joining mutants have specific defects in coding joint formation without a significant impairment in signal joints (Huye et al., 2002).

1.2.2 The Primary sequence, domain and structure of RAG2

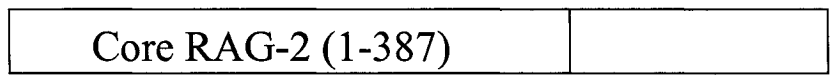
RAG2 protein has not been reported to share detectable similarity with any other protein in the data bank, however RAG2 entire sequence is highly conserved through evolution from pufferfish to humans. Both human and mouse RAG2 are composed of 527 amino acids and core RAG2 was defined from amino acid residue 1 to 387 (Figure 3A) (Sadofsky et al., 1994). The conservation of the non-core region of RAG2 is slightly higher than the core region.

1.2.2.1 Core RAG2 region

Hydrophobic cluster analysis (HCA) suggested that RAG2 is constituted of two distinct globular domains separated by a hinge region. The N-terminal globular domain corresponds to the core region (Figure 3B and 3C). This region is composed of a six-fold repeat of a 50-residue motif related to kelch/mipp motif that forms a four-stranded twisted antiparallel β sheet and is arranged in a circular formation like blades of a

Figure 3. Schematic representation and the proposed structure model of RAG2 protein. RAG2 itself doesn't appear to bind to RSS but it was found to help recognize or cleave distorted DNA intermediates and play an essential role in the joining step. **(A) The schematic representation of RAG2 protein.** Murine RAG2 protein is composed of 527 amino acids and is schematically divided into four regions: core RAG2 spanning from residue 1 to 387; hinge region spanning from residue 355 to 408; PHD domain spanning from amino acid residue 417 to 484 and a C-terminal domain in RAG2 from amino acid residue 491 to 527. Region spanning from amino acid residue 314 to 371 is the minimal domain defined in RAG2 for interacting with RAG1. NLS of RAG2 is localized at the extreme C-terminus, which is spanning from amino acid residue 491 to 527. **(B) and (C) The ribbon representation of an approximate template for the proposed human RAG2(1-349) six-bladed β-propeller structure. The β-propeller is viewed along (B) or perpendicular to (C) its axis.** Each blade of the propeller (labeled *a-f*; the *b* blade is *circled*) consists of a four-stranded β-sheet (β-strands are labeled *1* (inner strand) to *4* (outer strand). The sheets are organized in a circular formation like blades of a propeller or turbine. The closure of the turbine is formed by a 1 + 3 combination (N-terminal strand β_4 and three C-terminal strands β_1 to β_3). Two wide surfaces are provided on each side of the propeller by loops connecting strand β_2 to β_3 and β_4 to β_1 the longer loops, on the one hand (top) and strand β_1 to β_2 and β_3 to β_4, the shorter loops on the other hand (bottom). Figures in B and C are adapted from Corneo et al (2000).

A



1 527

RAG1-interacting domain

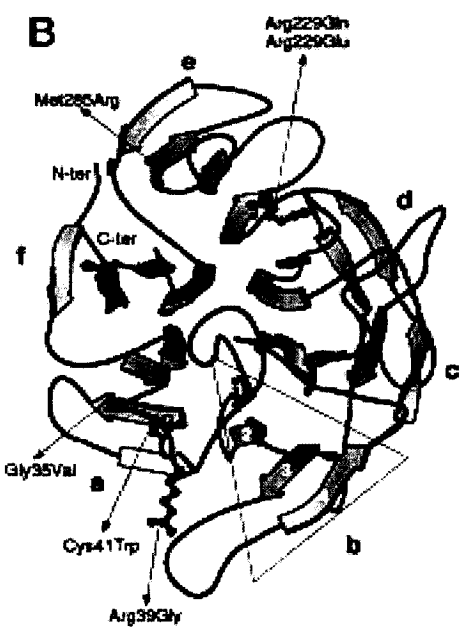
314 371

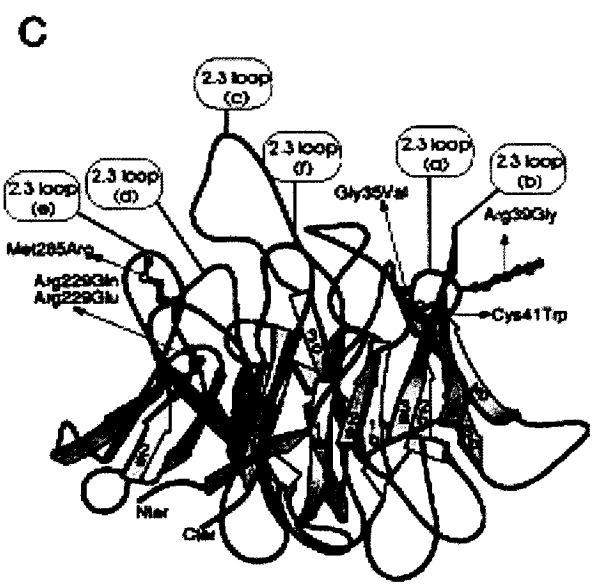
PHD domain

417 484

Nuclear localization

491 527





propeller (Callebaut and Mornon, 1998; Corneo et al., 2000). This proposed core-RAG2 structure model could play a crucial role in the tight complex formed by RAG1/RAG2 proteins and RSS sequences (Corneo et al., 2000).

The RAG1-interacting region has been mapped in the core RAG2 from amino acid residue 314 to 371, which coincides with the sixth kelch motif of the core RAG2. This sixth kelch motif alone can coprecipitate RAG1, but it can not activate the DNA recognition and hydrolytic activity, which implicated that the first five kelth motifs are also very important for the activity of the RAG1/RAG2 complex (Aidinis et al., 2000).

1.2.2.2 The PHD domain

C-terminal region of RAG2 from amino acid residue 417 to 484 was shown to display a homology with zinc-finger domain, constituted by seven cysteines and a histidine C4HC3, termed PHD (plant homeodomain), TTC (trithorax consensus) or LAP (leukemia-associated protein), which has been identified in a series of chromatin-associated proteins and is thought to be involved in protein-protein interactions (Callebaut and Mornon, 1998).

1.2.2.3 The C-terminal domain of RAG2

RAG2 protein lacking the very last 17aa (511-527) has been shown to lose nuclear localization, but the 17aa alone was not sufficient on its own to drive nuclear localization (Corneo et al., 2002). The C-terminal domain of RAG2 spanning from

amino acid residue 491 to 527 was determined to be the minimal necessary region to drive RAG2 nuclear localization (Corneo et al., 2002).

1.2.3 DNA binding activity of RAG1 and RAG2

RAG1 plays a major role in the recognition of RSS. Core RAG1 has been shown to bind to 12RSS and 23RSS independently of RAG2 with similar affinity but to bind tighter to 12RSS over 23RSS when RAG2 is present (Rodgers et al., 1999). HMG1 and HMG2 have been reported to facilitate the formation of the stable complex of RAG1/RAG2 with 23RSS (van Gent et al., 1997). By using electrophoretic mobility shift assays (EMSA), core RAG1 was found to have significant specificity for both the nonamer and heptamer sequence, but has greater specificity for the nonamer (Rodgers et al., 1999). Moreover, protein-DNA cross-linking studies implicated that both RAG1 and RAG2 proteins can contact with the coding flank that borders RSS heptamer, while the C-terminal of RAG1 has been shown to form the strongest contact to the coding flank (Eastman et al., 1999; Swanson and Desiderio, 1999; Mo et al., 2001). Although RAG2, in the absence of RAG1, has been reported to either bind to RSS weakly with no sequence specificity or not at all, RAG2 is required for the cleavage step. RAG2 has been suggested to induce the conformation change of RAG1 and enhance RAG1 binding to RSS and therefore form a stable RAG1/RAG2/RSS complex (Akamastsu and Oettinger, 1998; Mo et al., 1999). In addition, the RAG1/RAG2 complex has been suggested to preferentially assemble on an isolated 12RSS and then incorporate a naked

23RSS to assemble the synaptic complex (Jones and Gellert, 2002). However, the stoichimometry of the components in the RAG1/RAG2/RSS complex remains unresolved.

1.3 Ku protein overview

Ku protein identified as an autoantigen was named by the first 2 letters of scleroderma-polymyositis overlap syndrome patient in which it was identified (Mimori et al., 1981). Ku is a heterodimeric protein composed of Ku70 subunit and Ku80 subunit (Mimori and Hardin, 1986; Mimori et al., 1986). The gene encoding Ku70 is localized to chromosome 12 in human and chromosome 15 in mouse and the gene encoding Ku80 is localized to chromosome 2 in human and chromosome 1 in mouse (Cai et al., 1994; Koike et al., 1996). KARP-1 (Ku86 autoantigen related protein-1) is an alternate form of Ku80. It is expressed from the hKu80 gene locus with an extra 9kDa polypeptide to the N-terminus of Ku80 (Myung et al., 1997).

The Ku70/Ku80 heterodimer is a predominantly nuclear protein. Each of the two subunits has its own NLS (Koike et al., 1999a; Koike et al., 1999b; Koike et al., 1999c; Koike et al., 2000; Bertinato et al., 2001; Koike et al., 2001). Ku70 or Ku80 can either translocates to nucleus independently or as a heterodimer.

Ku is a prolific DNA binding protein. It binds to many DNA forms nonspecifically and with strong affinity, including double-stranded (ds) DNA ends, DNA nicks, gaps or bubbles (Blier et al., 1993; Falzon et al., 1993). It has also been shown to

bind to over 20 DNA sequences in a sequence-specific manner (Schild-Poulter et al., 2004).

Ku is involved in multiple cellular processes, including DNA double strand break repair pathway (NHEJ) (Smider et al., 1994; Errami et al., 1996; Baumann and West, 1998), V(D)J recombination (Smider et al., 1994; Errami et al., 1996; Lieber et al., 2003), telomere maintenance (Boulton and Jackson, 1996; Bailey et al., 1999; Tuteja and Tuteja, 2000), DNA replication (Shakibai et al., 1996; Ruiz et al., 1999; Schild-Poulter et al., 2003; Park et al., 2004), transcriptional regulation (Giffin et al., 1997; Camara-Clayette et al., 1999) and the maintenance of genomic stability (Difilippantonio et al., 2000; Li et al., 2002). While Ku is required for these processes, the precise mechanisms of its action are still unclear.

1.3.1 The primary sequence, domain and structure of Ku70 and Ku80

Ku has been identified in several species, including vertebrates, insects, yeast and even in plants and bacteria (Dynan and Yoo, 1998; Aravind and Koonin, 2001; Doherty et al., 2001; Tamura et al., 2002). The primary sequence and structure of both Ku subunit are well-conserved during evolution from yeast to humans (Paesen et al., 1996; Dynan and Yoo, 1998; Yagura and Sumi, 1999).

The crystal structure of the Ku70/Ku80 heterodimer bound to 14bp duplex DNA end was obtained with a Ku80 C-terminus truncation at a resolution of 2.5 Å. It unveiled the overall structure of the dimer as a form of a basket, with a large base extending in a

loop-structure reminiscent of a handle (Walker et al., 2001)(Figure 4A). Although Ku70 and Ku80 subunits do not share extensive sequence homology, they exhibit structural similarities: each subunit consists of an N-terminal α/β domain, a central β-barrel domain and a helical C-terminal arm (Figure 4B).

1.3.1.1 N-terminal α/β domain

N-terminal α/β domain in each subunit is composed of a six-bladed β-sheet which corresponds to the von Willebrand factor A (vWA) domain (Aravind and Koonin, 2000). N-terminal α/β domains are located on each side of the ring structure formed by the folding of both Ku70 and Ku80 β-sheet extensions. This domain makes scarce contribution to the dimer formation.

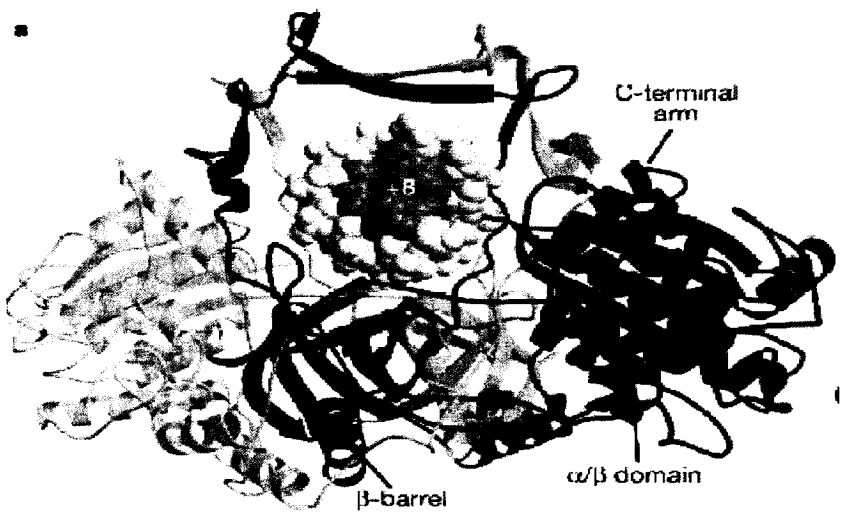
1.3.1.2 The core region of Ku70 and Ku80

Identification of prokaryotic Ku protein has allowed to define a central core region in both subunits of eukaryotic Ku proteins (Aravind and Koonin, 2001) (Figure 4B). The core region in each subunit consists of a central β-barrel domain and a helical C-terminal arm. This core region is critical for heterodimerization and DNA-binding (Wu and Lieber, 1996; Jin and Weaver, 1997; Cary et al., 1998; Wang et al., 1998b)

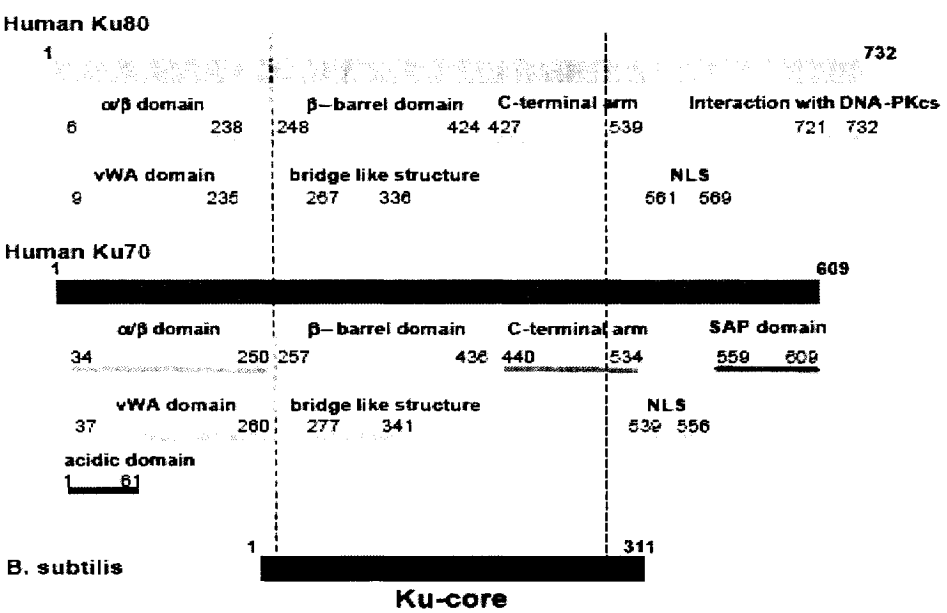
The central β-barrel domain consists of seven β strands and forms a symmetrical round barrel in each subunit. The two β-barrel domains together form a relatively flat base that holds approximately two turns of DNA.

Figure 4. Structure of the Ku dimer. (A) The crystal structure of Ku70/Ku80 bound to a structured 14bp duplex DNA molecule. The N-terminal α/β domain is composed of a six-stranded β-sheet in both Ku70 and Ku80 subunit and it lies at the periphery of the Ku70/Ku80 heterodimer. The amino edge of the β-sheet is proximal and the carboxy edge is distal to the DNA-binding grove. The central β-barrel domain is composed of seven β-strands and forms part of the DNA binding channel. A bridge-like structure extends from amino acid residue 277 to 341 in Ku70 and from 267 to 336 in Ku80 and encircles approximately 3 to 4bp of DNA that sits on the fat base formed by two β-barrel domains of each subunit. Ku70 is shown in red and Ku80 is shown in yellow and the duplex DNA in grey. **(B) Schematic representation of the human Ku70 and Ku80 subunit.** Ku70 subunit is composed of 609 amino acids. N-terminal α/β domain is from amino acid residue 34 to 250, which including a von willebrand factor A-like (vWA) domain from amino acid residue 37 to 260. The central β-barrel domain maps from amino acid residue 257 to 436. C-terminal arm is localized from amino acid 440 to 534.NLS of Ku70 is mapped from amino acid residue 539 to 556. SAP (SAF-A/B, Acinus and Plas) domain is mapped from amino acid residue 572 to 769, which corresponds to the C-terminal Ku-80 independent DNA binding region and contains a helix-loop-helix (HLH) structure. Ku80 subunit is composed of 732 amino acids. Like Ku70 subunit, Ku80 subunit also contains a vWA-like domain from amino acid residue 9 to 235, a central β-barrel domain from amino acid residue 248 to 424 and a C-terminal arm from amino acid residue 427 and 539. NLS of Ku80 is mapped from amino acid residue 561 to 569. The C-terminal 28 amino acid residues are involved in interaction with the catalytic subunit of DAN-dependent protein kinase (DNA-PKcs). The dash line defines the core region of Ku that displays structure homology between eukaryotic and prokaryotic Ku proteins. Figure in A are taken from Walker et al. (2001) and figure in B are taken from Schild-Poulter et al. (2004).

A


C-terminal arm
α/β domain
β-barrel

B


Human Ku80
1
732
α/β domain
6
238
β–barrel domain
248
424
C-terminal arm
427
539
Interaction with DNA-PKcs
721
732
vWA domain
9
235
bridge like structure
267
336
NLS
561
569
Human Ku70
1
609
α/β domain
34
250
β–barrel domain
257
436
C-terminal arm
440
534
SAP domain
559
609
vWA domain
37
260
bridge like structure
277
341
NLS
539
556
acidic domain
1
61
1
311
B. subtilis
Ku-core

The C-terminal arm in the core region is formed by α–helices and folded around the β-barrel of the opposite subunit. The C-terminal arm is particularly important for the stability of Ku70/Ku80 heterodimer complex (Jin and Weaver, 1997; Osipovich et al., 1997).

1.3.1.3 C-terminus of Ku70 and Ku80

Both of the C-terminus of Ku70 and Ku80 subunit outside of the core region of are not important for Ku70/Ku80 heterodimerization (Cary et al., 1998; Wang et al., 1998a; Wang et al., 1998b).

NMR study of the C-terminal region of Ku70 reveals a helix-extended loop-helix structure, which corresponds to a SAP domain (Zhang et al., 2001). Amino acid residues from 536 to 609 in Ku70 constitute a Ku80-independent DNA binding domain (Wang et al., 1998a; Wang et al., 1998b).

NMR study of the C-terminal region of Ku80 revealed a globular domain spanning from amino acid residue 592 to 709 (Harris et al., 2004). The extreme C-terminal of Ku80 constitutes the DNA-PKcs-interacting domain, which is an unstructured, random coil conformation in the solution and could potentially form an α–helix when binding to DNA-PKcs (Harris et al., 2004).

1.3.2 Ku and DNA-PK kinase activity

DNA-PKcs, the catalytic subunit of the DNA-PK holoenzyme, is the only

member of the phosphatidylinositol-3-kinase (PI3K)-related family that is exclusive to higher eukaryotes (Gottlieb and Jackson, 1993; Hoekstra, 1997; Rouse and Jackson, 2002). PI3K-related family doesn't have lipid kinase activity but shares a kinase-homology domain and the extreme C-terminus tail with the PI3K family, a ubiquitous lipid kinases family capable of phosphorylating phosphoinositides at the 3-hydroxyl (Katso et al., 2001). Ku70/Ku80 acts as the DNA binding component of DNA-PK holoenzyme (Gottlieb and Jackson, 1993; Giffin et al., 1996). Although DNA-PKcs is able to interact with DNA ends and some short oligonucleotides at the absence of Ku (Yaneva et al., 1997; Hammarsten and Chu, 1998), Ku is required to recruit DNA-Pkcs to DNA ends and activate DNA-PK activity under physiological conditions (DeFazio et al., 2002; Weterings et al., 2003).

1.3.3 Integration of Ku to V(D)J recombination

Ku is required for N-region diversity. Both Ku70 and Ku80 can associate with TdT (Mahajan et al., 1999; Purugganan et al., 2001; Sandor et al., 2004). Ku70/Ku80 is essential for the formation of both coding joints and signal joints. Ku70/Ku80 was found to remain associated with signal ends and RAG1/RAG2 complex, DNA-PKcs and XRCC4 in the postcleavage complex (Agrawal and Schatz, 1997). Signal joint formation is completely inhibited and coding joint formation is reduced in the absence of Ku (Cortes et al., 1996; Weis-Garcia et al., 1997).

1.4 Objectives:

Unpublished data in our laboratory have recently demonstrated that full-length RAG proteins can physically interact with Ku in vivo via coimmunoprecipitation assays. This RAG-Ku complex is a complicated protein-protein complex, as Ku70/Ku80 is a heterodimer and RAG1 forms a complex with RAG2 to mediate the efficient cleavage in V(D)J recombination. Ku70/Ku80 is the initiator of the NHEJ pathway by binding to the DNA double-stranded breaks. The exciting finding of the interaction between RAG and Ku may suggest a new model of the integration of NHEJ apparatus into the V(D)J recombination. The Ku-RAG interaction may help to recruit the NHEJ factors to the DNA-ends for efficient end processing, which facilitates DNA repair response and promotes cell survival. **The overall goal of my project has been to define the Ku-interacting domains in both RAG1 and RAG2 with the long term objective of linking their interaction to specific function in V(D)J recombination.** (1) Define Ku-interacting regions in both RAG1 and RAG2 proteins by truncating RAG1 and RAG2 from both N- and C- terminus; (2) Confirm the results from the deletion mapping assay by expressing the isolated potential Ku-interacting regions as a fusion protein and testing to binding to Ku; (3) Find the RAG mutants that are not able to interact with Ku by site-directed mutagenesis; (4) Do in vivo experiment to link the Ku-RAG interaction to specific function in V(D)J recombination. In this study, although the Ku-interacting region in RAG2 remains to be precisely determined, I have successfully defined a 100 aa peptide in the central domain of RAG1 protein that is sufficient for binding to Ku.

2. Materials and Methods

2.1 Plasmids and vectors

The mammalian expression vectors pRAG1 (M2CD7) and pRAG2 (R2RCD2) encoding murine full-length RAG1 and RAG2 were kindly provided by Dr. M.A. Oettinger (Harvard University, Boston, MA). These vectors were grown in *E. coli* MC1061/P3 (Invitrogen) and large scale plasmid DNA were prepared by using alkaline lysis followed by cesium chloride/ ethidium bromide gradient centrifugation.

The baculovirus expression vector VBB2-86Ku and VBB2-Kup70tH[6] encoding the Ku80 subunit and a hexahistidine-tagged Ku70 subunit of the human Ku were generously provided by Dr. J.D. Capra (University of Texas Southwestern School of Medicine, Dallas, TX, USA). These expression vectors were propagated in *Spodoptera frugiperda* cell line Sf9 (ATCC CRL1711).

Several of the plasmids employed here were made previously by Dr. Caroline Schild-Poulter in the laboratory. pET-R2 contains the core region of murine RAG2 inserted into pET-30c vector with a hexahistidine-tag on N- teminus. pET-MBP-R1 was created by subcloning of maltose-binding protein(MBP) fusion core RAG1 into pET-30b vector, which contains a hexahistidine-tag only on N-terminus. pET-MBP was created by the same strategy and the same backbone vector (pET-30b). These vectors were expressed in *E. coli* BL21 and small scale of plasmid DNA was prepared by QIAprep Spin Miniprep Kit (QiaGEN, Mississauga, ON). pET-30b and pET-30c vectors were purchased from Novagen (Mississaga, ON).

2.2 Preparation of DNA constructs

2.2.1Preparation of DNA fragments for in vitro translation

All RAG1 fragments used for in vitro translation were generated by PCR amplification using pRAG1 (M2CD7) as the template (using the PTC-200 Peltier Thermocycler, MJ research, Reno, CA), except for the fragments R1 $_{1\text{-}501}$ and R1 $_{1\text{-}755}$ that were generated by restriction enzyme digestion with SphI and AfeI, respectively. RAG1 C-terminal mutant cDNAs (R1 $_{1\text{-}677}$, R1 $_{1\text{-}596}$, R1 $_{1\text{-}554}$ and R1 $_{1\text{-}549}$) were produced by PCR amplification using the same T7Fw primer and their corresponding Rev primers. RAG1 N-terminal mutant cDNA (R1 $_{400\text{-}1040}$, R1 $_{499\text{-}1040}$, R1 $_{592\text{-}1040}$ and R1 $_{660\text{-}1040}$) was produced by PCR amplification using the same R1Rev1054 primer and its corresponding Fw primers (see **Appendix I** for primer sequences).

Similar to RAG1, RAG2 C-terminal mutant cDNAs (R2 $_{1\text{-}328}$, R2 $_{1\text{-}278}$, R2 $_{1\text{-}242}$ and R2 $_{1\text{-}164}$) were generated by restriction enzyme digestion on pRAG2 (R2RCD2) with StyI, PvuII, Acc65I and NspI separately. RAG2 N-terminal mutant cDNAs (R2 $_{132\text{-}532}$, R2 $_{161\text{-}532}$, R2 $_{173\text{-}532}$ and R2 $_{389\text{-}532}$) were generated by PCR amplification using the same R2Rev554 primer and their corresponding Fw primers (see **Appendix I** for primer sequences).

All DNA fragments used for in vitro translation were purified by phenol/chloroform followed by a standard ethanol precipitation and then finally were suspended in Rnase-free water (QiaGEN, Mississauga, ON).

2.2.2 Construction of plasmids pMAL-R1$_{499-763}$, pMAL-R1$_{573-763}$, pMAL-R1$_{660-763}$, pMAL-R1$_{792-1013}$

pMAL-RAG1 $_{499-763}$ was cloned by PCR amplification (using the PTC-200 Peltier Thermocycler, MJ research, Reno, CA). RAG1 fragment from amino acid residue 499 to 763 of RAG1 was generated by PCR amplification using pRAG1 (M2CD7) as a template. The primers employed here were XbaI-R1Fw499 and Rev763stop (see **Appendix I** for primer sequences). The PCR product was purified by electrophoresis on 1% agarose gel and extracted by using the Gel Extraction Kit (QiaGEN, Mississauga, ON). The purified PCR product was then digested with XbaI. pMAL-c2 vector was first digested with HindIII followed by filling-in with Klenow and then further digested with XbaI. The digested PCR product was then ligated into XbaI and HindIII sites of pMAL-c2 vector at 16°C for 16 h by T4 DNA ligase in a 20 μl reaction. 1/10 of the ligation reaction products were transformed into *E. coli* DH5α by electroporation using an *E. coli* Pulser (Bio-Rad Laboratories). Plasmid DNA from transformed *E. coli* was prepared by QIAprep Spin Miniprep Kit (QiaGEN, Mississauga, ON), a standard alkaline lysis protocol and was sent to be sequenced. All enzymes used for cloning were purchased from New England Biolabs (Mississauga, ON). Constructs of pMAL-R1$_{573-763}$, pMAL-R1$_{660-763}$ and pMAL-R1$_{792-1013}$ were created by using the same strategy and the same vector backbone (pMAL-c2).

2.2.3 Preparation of RAG1 K405A/H406A/R407A mutant

RAG1 K405A/H406A/R407A was generated by site-directed mutagenesis using pET-MBP-R1 as the backbone. The primers employed here were R1Fw K405A/H406A/R407A and R1Rev K405A/H406A/R407A (see Appendix I for primer sequences). PCR of 21 cycles at 95°C, 30 sec, 55°C 1 min, 68°C, 14 min were done in 20 µl of reaction mixture containing 20 mM Tris-HCl (pH8.8), 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM Mg SO_4, 1% Triton X-100, 100 ug/ml nuclease-free bovine serum albumin (BSA), 100 µM each of dNTPs, 15 ng pET-MBP-R1 and 7.5 ng of each primers, 1.25 U of PfuTurbo polymerase. After checking on 1% agarose gel, the PCR product was directly digested with 0.5 µl DpnI (New England Biolab), which is used to digest the parental DNA template. 1 µl of mixture was then transformed into *E. coli* DH5α by electroporation using an *E. coli* Pulser (Bio-Rad Laboratories). Plasmid DNA from transformed *E. coli* was prepared by QIAprep Spin Miniprep Kit (QiaGEN, Mississauga, ON). It was initially checked by digestion with endonuclease EarI and was then sequenced.

2.3 Large scale of plasmid (pRAG1 and pRAG2) preparations by alkaline lysis followed by cesium chloride/ ethidium bromide gradient centrifugation

A 500 ml bacterial culture was pelleted by centrifugation at 5000 rpm for 10 min at 4°C, resuspended by vortexing and pippetting up and down in 20ml ice cold GTE buffer containing 25 mM Tris-HCl (pH8.0), 10 mM EDTA, 50 mM glucose with 50

mg/ml lysozyme. After 10 min incubation on ice, 40 ml lysis buffer containing 200 mM NaOH and 1% SDS was added and mixed gently by inverting followed by incubation on ice for 10 min. 30 ml 3 M KOAc (pH 5.5) was added and the suspension was kept on ice for another 10 min. After centrifugation at 6000 rpm for 20 min at 4°C, the supernatant was filtered through 4 layers of cheesecloth. 90 ml isopropanol was added, mixed well and the solution was incubated at room temperature for 15 min followed by centrifugation at 5000 rpm for 20 min at 4°C. The supernatant was discarded and the pellet was drained well at room temperature. The pellet was then resuspended in 15 ml 1X TE buffer containing 10 mM Tris-HCl (pH 8.0) and 1 mM EDTA and transferred to a 40 ml Oakridge tube. 10ml 5 M LiCl was added and incubated on ice for 20 min. After the suspension was centrifuged at 7000 rpm for 15 min at 4°C, the supernatant was transferred to a 250 ml bottle and precipitated with 50 ml 100% ethanol for 30 min on ice. The suspension was centrifuged at 7000 rpm for 20 min at 4°C and the pellet was drained well by inverting the bottle. The pellet was resuspended in 2.8 ml TE buffer, 3.08 g CsCl and 30 µl ethidium bromide (2 mg/ml). The solution was transferred to a polypropylene ultracentrifuge tube (Beckman), the tube was sealed and centrifuged at 95 Krpm overnight at 22°C in a Beckman TLN 100 fixed angle rotor. The following day, the supercoiled DNA band was extracted from the tube by syringe using an 18 gauge needle. The supercoiled DNA was transferred to another polypropylene ultracentrifuge tube, filled to the top with CsCl solution (Refractive index = 1.397-1.399), sealed and centrifuged again at 95 Krpm for 6 h at 22°C in a Beckman TLN 100 fixed angle rotor.

The supercoiled DNA was removed as above and ethidium bromide was extracted using water/ salt saturated isoporpanol. When the ethidium bromide was no longer detectable in either phase, the aqueous phase was transferred to dialysis tubing and dialyzed against 2L 1X TE buffer at 4°C for overnight with 2 buffer changes in interim. The DNA quality and quantity was determined by measuring the absorbance at 260 nm and 280 nm and by running on an agarose gel to assess supercoiled DNA. DNA preps were then precipitated with 1/10 volume of NaOAc and 2.5 volumes of ice-cold 100% ethanol and stored in aliquots in -20°C.

2.4 Cell culture and infections

2.4.1 Cell culture

Human Jurkat T cells (ATCC CRL-1990) were maintained in RPMI-1640 (Gibco-BRL) with 10% FBS (JRH Bioscience) and 2 mM L-glutamine at 37°C with 5% CO_2. Cells were cultured by adding fresh medium every 2 to 3 days and were maintained at a density between $1X10^5$ and $1X\ 10^6$ cells/ml. The Sf9 cells were grown in TNM-FH made from Grace's insect medium TNM (Gibco-BRL) supplemented with yeastolate (3.3 mg/ml) and 10% heat inactivated fetal bovine serum (FBS) (JRH Biosciences). They were seeded at a density of $8X10^5$ cells/cm^2, maintained at 27°C and passed at confluency in 1 to 3-5 dilutions.

2.4.2 Infection of Sf9 cells with the baculovirus expression vector VBB2-86Ku and

*VBB2-Kup*70tH[6]

Sf9 cells were seeded at a density of $8X10^5$ cells/cm^2 and allowed to attach for 1h, and then gently washed with serum-free TNM. VBB2-86Ku and VBB2-Kup70tH[6] were then co-infected into Sf9 cells at a multiplicity of infection (MOI) of 5 to 10. Following incubation at 27°C for 1 h, the baculovirus innoculum was aspirated and discarded and fresh TNM-FH was added. After infection for 3 days, Sf9 cells were scraped off the dish with a rubber policeman, washed 3 times with ice-cold 1X phosphate-buffered saline (PBS) and stored on ice for purification of recombinant Ku70/Ku80. This preparation was provided by Mr.DominicVallée in the laboratory.

2.5 Preparation of whole cell extracts and nuclear extracts

2.5.1 Preparation of Jurkat T cell whole cell extracts (WCE)

Jurkat T cell WCEs were used for in vitro protein binding assay with in vitro translated ^{35}S labeled full-length RAG1/RAG2 proteins and their deletion mutants. Jurkat T cells were centrifuged at 3000 rpm for 3 min and then the cell pellets were washed three times with 5-10 volumes of ice-cold 1X PBS. Cell lysis was performed at $5X10^7$/ml at 4°C for 25 min on ice with agitation in WCE buffer containing 150 mM NaCl, 1 mM EDTA, 50 mM Hepes (pH7.4), 10% glycerol, 1% Nonidet P-40, 1 mM phenylmethylsulfonyl fluoride (PMSF) and 0.5 mM dithiothreitol (DTT). After centrifugation at 13 Krpm for 30 min at 4°C, the supernatant was collected and stored in aliquots at –80°C. Protein concentration of the supernatant was measured with Bradford

Assay (BioRad).

2.5.2 Preparation of Jurkat T cell nuclear extracts (NE)

Jurkat T cell NEs were used for in vitro binding assay with purified maltose-binding protein fusion RAG1 proteins (MBP-RAG1s) and were prepared following the method described previously (Andrews and Faller, 1991). Jurkat T cells were centrifuged at 3000 rpm for 3 min and then the cell pellets were washed three times with 5-10 volumes of ice-cold 1X PBS. The cell pellets were then resuspended by pipetting in Buffer A (300 µl per 3X 10^6 cells) containing 10 mM Hepes (pH7.9), 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT and 0.2 mM PMSF. The samples were allowed to swell on ice for 6 min and vortexed at the maximum position for 10 sec and then were centrifuged at 13 Krpm for 1 min at 4°C. The nuclei pellets were then resuspended with 60 µl of Buffer C containing 20 mM Hepes (pH7.9), 420 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 25% glycerol, 0.5 mM DTT and 0.2 mM PMSF and incubated on ice for 25 min for high salt extraction. Cellular debris was removed by centrifuging at 13 Krpm for 5 min at 4°C and the supernatant (NE from Jurkat T cells) was dialyzed against binding buffer containing 25 mM Hepes (pH7.4), 60 mM KCl, 1.2 mM EDTA, 2 mM DTT, 0.2 mM PMSF, 0.1% Nonidet P-40, 12% glycerol for 6 to 8 h with 3 buffer changes in interim. The dialyzed nuclear extracts were quantified to be 1 to 2 µg/µl and stored in aliquots at –80°C.

2.6 Protein concentration determination

The Bradford Assay was used to determine the protein concentration of whole cell extracts and nuclear extracts. A small volume of extract was added to 800 µl of water followed by addition of 200 µl of Bio-Rad Protein Assay Dye Reagent in the cuvette and was mixed by inverting 2-3 times on parafilm. The absence of each sample was measured at 595 nm on a spectrophotometer (Biochrom Ultrospec 3000). The amount of protein present in the extract was determined by comparison to a bovine serum albumin (BSA) standard curve (2 µg, 5 µg, 10 µg).

2.7 Protein expression and purification

2.7.1 Expression of His-tagged RAG1 and RAG2 proteins and MBP fusion RAG1 proteins

1 ng of each plasmid was transformed into *E. coli* BL21 by electroporation using an *E. coli* Pulser (Bio-Rad Laboratories). Single colony was inoculated from corresponding selective plates and incubated in 5 ml LB with shaking at 37°C for 4 to 5 h until the OD_{595} reaches 0.6 to 0.8. Then isopropyl-beta-D-thiogalactopyranoside (IPTG) was added into the bacteria culture at 0.3 mM of final concentration to induce the gene expression. All MBP fusion RAG1 proteins were induced with shaking at 37°C for 2 h while all his-tagged RAG1 and RAG2 proteins were induced at 4°C for overnight. Bacteria cells were then centrifuged at 5000 rpm for 15 min at 4°C. For MBP fusion RAG1 proteins, bacteria pellets were resuspended with Column buffer containing 20

mM Tris-HCl (pH 7.4), 200 mM NaCl, 1 mM EDTA, 1 mM PMSF, 10 mM DTT and 0.1% NP-40 while for his-tagged RAG1 and RAG2 proteins, pellets were resuspended with 1X Extraction buffer (pH7.0) containing 50 mM Sodium phosphate, 300 mM NaCl, 0.1% NP-40 and 1 mM PMSF. The bacteria cells were then sonicated 3 times for 20 sec with a 30sec interval on ice by using a Branson Sonifier 450 on constant duty cycle (output control =1). Following sonication, bacteria lysate was centrifuged at 13 Krpm for 5 min at 4°C and the supernatant was then collected and store in aliquots at -80°C for further purification.

2.7.2 Purification of his-tagged RAG1 and RAG2 proteins from bacteria lysate

His-tagged RAG1 and RAG2 proteins including pET-MBP-R1, pET-MBP-R1 K405A/H406A/R407A, and pET-R2 were purified from bacteria lysate by using the TALON^{TM} resin according to the supplier's instruction with modifications (Clontech). 60 µl of slurry TALON^{TM} resin (50% w/v) was transferred to a 1.5 ml sterile Eppendorf tube and centrifuged at 3000 rpm for 2 min at 4°C. After removal of the supernatant, the 50% slurry beads were then washed twice by gently vortex with 10 bed volumes of ice-cold 1 X Extraction buffer (pH7.0) containing 50 mM Sodium phosphate, 300 mM NaCl, 0.1% NP-40 and 1 mM PMSF. After spinning at 3000 rpm for 2 min at 4°C, the supernatant was discarded and the prepared bacteria lysate was then added to the resin and rotated on the wheel for 30 min at 4°C. After centrifugation at 14 Krpm for 1 min at 4°C, the supernatant was removed and the resin was washed three times with 1 X

Extraction/ Wash buffer (pH8.0) containing 50 mM sodium phosphate, 300 mM NaCl, 0.1% NP-40 and 1 mM PMSF. For the MBP fusion RAG1 binding assay, after the last wash with 1 X Extraction/ Wash buffer (pH8.0), 40 µl of Elution buffer (pH 7.0) containing 150 mM imidazole, 50 mM sodium phosphate, 300 mM NaCl, 1 mM PMSF was added to the protein-resin complex. After gentle vortex and centrifugation twice at 14 Krpm for 2 min, the supernatants (eluted his-tagged MBP fusion RAG1) were pooled together and stored with 20% glycerol at -80°C for further use. The purity of the proteins was assessed by electrophoresis in 8% sodium dodecyl sulfate (SDS) - polyacrylamide gel followed by rapid coomassie blue staining.

2.7.3 Purification of MBP fusion RAG1 proteins

MBP fusion RAG1 proteins including MBP fusion $R1_{499-763}$, $R1_{573-763}$, $R1_{660-763}$, $R1_{792-1013}$ were purified from bacteria lysate by using the amylose resin according to the supplier's instruction with modifications (New England Biolab). 100 µl of slurry amylose resin (50% w/v) was transferred to a 1.5 ml sterile Eppendorf tube and centrifuged at 3000 rpm for 2 min at 4°C. After removal of the supernatant, the 50% slurry beads were washed twice by vortex with 10 bed volumes of ice-cold Column buffer containing 20 mM Tris-HCl (pH 7.4), 200 mM NaCl, 1 mM EDTA, 1 mM PMSF, 10 mM DTT and 0.1% NP-40. After spinning at 3000 rpm for 1 min at 4°C, the supernatant was discarded and the prepared bacteria lysate was then added to the resin and rotated on the wheel for 2 h at 4°C. After centrifugation at 3000 rpm for 1 min at

4°C, the supernatant was removed and the resin was washed three times with Column buffer containing 20 mM Tris-HCl (pH 7.4), 200 mM NaCl, 1 mM EDTA, 1 mM PMSF, 10 mM DTT and 0.1% NP-40. After the last wash, the resin was resuspended with Elution buffer containing 20 mM Tris-HCl (pH 7.4), 200 mM NaCl, 10 mM maltose, 1 mM EDTA, 1 mM PMSF, 10 mM DTT and 0.1% NP-40 by vortex and incubated on wheel with agitation for 20 min at 4°C. After centrifugation at 3000 rpm for 1 min at 4°C, the supernatant (protein elution) was transferred to a 1.5 ml sterile Eppendorf tube and the resin was further eluted twice. All the protein eluates were pooled together. Purified proteins were stored in 20% glycerol -80°C for further in vitro binding assay. The purity of the proteins was assessed by of electrophoresis in 8% (SDS) - polyacrylamide gel followed by rapid coomassie blue staining.

2.7.4 Purification of recombinant Ku70/Ku80 from Sf9 cells

Baculo Ku70/Ku80 from Sf9 cells were purified at 4°C by column chromatography according to instruction with modification of His•Bind Resin and Buffer Kit (Novagen). 200 µl of His•Bind Resin (50% w/v) was transferred to a small polypropylene column and allowed to pack under gravity flow. The column was first washed twice with 6 bed volumes of deionized water, then was charged twice with 5 bed volumes of 1 X Charge buffer containing 50 mM $NiSO_4$ and was equilibrated twice with 6 bed volumes of 1 X Binding buffer containing 500 mM NaCl, 20 mM Tris-HCl (pH 7.9) and 5 mM imidazole. The Sf9 whole cell extracts diluted with 4 volumes of binding

buffer was then loaded. The column was then washed twice with 10 bed volumes of 1 X Binding buffer containing 500 mM NaCl, 20 mM Tris-HCl (pH 7.9) and 5 mM imidazole and twice with 6 bed volumes of 1 X Wash buffer containing 500 mM NaCl, 20 mM Tris-HCl (pH 7.9) and 60 mM imidazole. The Baculo Ku70/Ku80 was eluted by fractions from the column by adding 6 bed volumes of 1 X Elution buffer containing 1 M imidazole, 20 mM Tris-HCl (pH7.9) and 500 mM NaCl. After checking on SDS gel, the eluates were pooled together, concentrated and then dialyzed against binding buffer containing 25 mM Hepes (pH7.9), 60 mM KCl, 1.2 mM EDTA, 2 mM DTT, 0.2 mM PMSF, 0.1% Nonidet P-40, 12% glycerol for 6 to 8 h with 3 buffer changes in interim. Purified recombinant Ku70/Ku80 was stored with 30% glycerol in aliquots in -80°C for further use. The purity of the proteins was assessed by sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS-PAGE) followed by silver staining and was further confirmed by western blotting with an anti-Ku70 antibody (N3H10 Ab-4, NeoMarkers) and an anti-Ku80 antibody (Ab-2, NeoMarkers).

2.8 In vitro translation of full-length and truncated RAG1 and RAG2 proteins

Full-length and truncated RAG1 and RAG2 proteins were translated in vitro in the presence of ^{35}S methionine using the T7 coupled reticulate lysate system according to the manufacturer's recommended protocol with modification (Promega). 1 μg of pRAG1 (M2CD7) and 0.8 μg of pRAG2 (R2RCD2) vectors were used as DNA templates for full-length RAG1 and full-length RAG2 protein translation in a 50 μl reaction containing

25 μl rabbit reticulocyte lysate, 2 μl TnT reaction buffer, 1 μl amino acid mixture minus methionine, 5 μl ^{35}S labeled methionine, 1 μl T7 RNA polymerase, 1 μl RNasin® ribonuclease inhibitor and Rnase-free water. The preparation of truncated mutants of RAG1 and RAG2 was described elsewhere in the text. PCR DNA templates were used at the volume of 6 to 8 μl (about 200 ng total) in a 25 μl reaction containing 12.5 μl rabbit reticulocyte lysate, 1 μl TnT reaction buffer, 0.5 μl amino acid mixture minus methionine, 2.5 μl ^{35}S labeled methionine, 0.5 μl T7 RNA polymerase, 0.5 μl RNasin® ribonuclease inhibitor and Rnase-free water. After coupled transcription/translation at 30°C for 1.5 h, 2 μl of translation product was resolved on 10% or 15% SDS-polyacrylamide gel depending on the size of the deletion mutant. Gels were dried for 2 h at 80°C on a slow heating program and then the translation products were visualized with PhosphorImager software and the densitometry was performed to quantify the relative protein translation level by using ImageQuant (Molecular Dynamics). After correcting for the background, the absolute volume of each deletion mutant for further binding assay was calculated as compared to the other protein translation level.

2.9 Direct protein binding assay

Direct binding of his-tagged MBP fusion core RAG1 and his-tagged core RAG2 to purified Ku70/Ku80 were performed at 4°C for 2 h in the binding buffer containing 25 mM Hepes (pH7.4), 60 mM KCl, 0.2 mM PMSF, 0.1% NP-40, 12% glycerol up to 200 μl final volume with 200 μg/ml ethidium bromide. 10% of purified Ku70/Ku80 was

saved as a control for input. Purified his-tagged MBP (product of pET-MBP) and his-tag alone (product of pET vector) on TALON^{TM} resins were used as the negative controls for his-tagged MBP fusion core RAG1 and his-tagged core RAG2 respectively. After washes with binding buffer, RAG-Ku complex was resuspended in SDS loading buffer and resolved on 8% SDS-polyarylamide gel. The proteins were transferred onto immunoblot PVDF membranes for 1 h at 100V. Then the membranes were immunoblotted with an anti-Ku70 antibody (N3H10 Ab-4, NeoMarkers) at the dilution of 1/1000(v/v) for 2 h at room temperature and the immunoreactive proteins were detected by using the Western Lightning Chemiluminescence Reagent (Amersham Pharmacia Biotech). To check the input proteins, the membranes were rehybridized with an anti-MBP antibody (MBP-c18 sc-808, Santa Cruz Biotechnology) at the dilution of 1/1000 (v/v) and an anti-RAG2 antibody (RAG2-M300 sc-5600, Santa Cruz Biotecnology) at the dilution of 1/500 (v/v).

2.10 In-vitro protein binding assay (deletion mapping assay)

300 μg of WCEs from Jurkat T cells were used per binding reaction. WCEs in the WCE buffer containing 150 mM NaCl, 1 mM EDTA, 50 mM Hepes (pH7.4), 10% glycerol, 1% NP-40, 1 mM PMSF and 0.5 mM DTT were first precleared on wheel with 10μl 50% slurry protein A-Sepharose beads (Sigma) at 4°C for 30 min. After gentle centrifugation at 3000 rpm at 4°C for 1 min, the supernatant was collected and Ku70/Ku80 was then immunoprecipitated from the supernatant at 4°C for 2 h from with 2 μg/ml anti-Ku70 antibody (N3H10 Ab-4, NeoMarkers) followed by incubation on

wheel at 4°C for 30 min with 20 μl protein A-Sepharose beads. The Sepharose beads were washed 3 times with WCE buffer followed by twice with binding buffer containing 25 mM Hepes (pH7.4), 60 mM KCl, 0.2 mM PMSF, 0.1% NP-40, 12% glycerol. In vitro translated ^{35}S labeled proteins were diluted in binding buffer and 10% of diluted ^{35}S labeled proteins were saved for use as input. Binding of in vitro translated full-length and truncated RAG1 and RAG2 proteins to Ku70/Ku80 immobilized on protein A-Sepharose beads was performed at 4°C for 2 h in a final volume of 200 μl binding buffer in the presence of 200 μg/ml ethidium bromide with Sepharose beads alone as negative controls. After extensive wash 3 times with 500 μl binding buffer, RAG-Ku complex was resuspended in SDS loading dye and then resolved on a 10% or 15% SDS-polyacrylamide gel. Gels were dried for 2 h at 80°C on a slow heating program and then the specifically bound RAG proteins were visualized with PhosphorImager software.

2.11 MBP fusion protein binding assay

Binding of eluted MBP fusion RAG1 proteins to Ku70/Ku80 from 60 μg NE from Jurkat T cells was performed at 4°C for 2 h in a final volume of 200μl binding buffer containing 25 mM Hepes (pH7.4), 60 mM KCl, 0.2 mM PMSF, 2 mM DTT, 0.1% NP-40, 12% glycerol and in the presence of 200 μg/ml ethidium bromide followed by RAG1-Ku70/Ku80 complex immunoprecipitation on protein A-Sepharose beads by 1.5 μg/ml anti-Ku70 antibody (N3H10 Ab-4, NeoMarkers) at 4°C for 2 h from the binding reaction. After the Sepharose beads were extensively washed 3 times with the binding

buffer, RAG1-Ku complex was then resuspended in SDS loading buffer, resolved on 8% SDS-polyacrylamide gel and transferred onto immunoblot PVDF membranes for 16 h at 30V. Then the membranes were immunoblotted with an anti-MBP antibody (MBP-c18 sc-808, Santa Cruz Biotechnology) at the dilution of 1/1000 (v/v) for 2 h at room temperature and the immunoreactive proteins were detected by using the Western Lightning Chemiluminescence Reagent (Amersham Pharmacia Biotech). To check the immunoprecipitated Ku70/Ku80, the membranes were rehybridized with an anti-Ku70 antibody (N3H10 Ab-4, NeoMarkers) at the dilution of 1/1000 (v/v) for 2 h at room temperature.

2.12 RAG1 point mutation binding assay

Binding of MBP fusion RAG1 point mutant protein to Ku70/Ku80 from 60 μg nuclear extract from Jurkat T cells were performed at 4°C for 2 h in the binding buffer containing 25 mM Hepes (pH7.4), 60 mM KCl, 0.2 mM PMSF, 0.1% NP-40, 12% glycerol up to 200 μl final volume with 200 μg/ml ethidium bromide. 10% of purified MBP fusion RAG1 point mutant protein was saved for input. MBP was used as the negative control. After washed 3 times with binding buffer, RAG-Ku complex was resuspended in SDS loading buffer and resolved on 8% SDS-polyarylamide gel. The proteins were transferred onto immunoblot PVDF membranes for 16 h at 30V. Then the membranes were immunoblotted with an anti-Ku70 antibody (N3H10 Ab-4, NeoMarkers) at the dilution of 1/1000 (v/v) for 2 h at room temperature and the immunoreactive

proteins were detected by using the Western Lightning Chemiluminescence Reagent (Amersham Pharmacia Biotech). The membranes were rehybridized with an anti-MBP antibody (MBP-c18 sc-808, Santa Cruz Biotechnology) at the dilution of 1/1000 (v/v) to check input proteins.

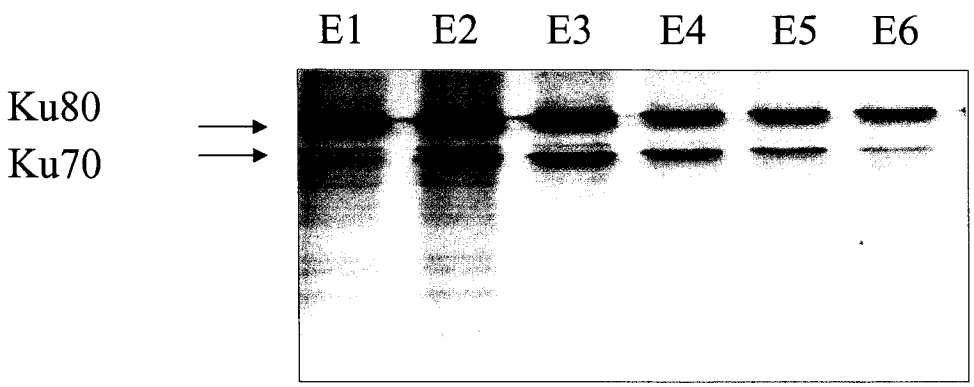
3. Results:

3.1 The interactions of RAG1/RAG2 with Ku70/Ku80 are direct and are mediated by the core domains

Previous work in our laboratory has shown that full-length RAG1/RAG2 proteins can interact with Ku70/Ku80 in vivo via coimmunoprecipitation assays (Cui et al. in preparation). Core RAG1 and RAG2 proteins have been shown to be active in the cleavage step of V(D)J recombination (Cuomo and Oettinger, 1994; Sadofsky et al., 1994). Moreover, compared to full-length RAG1 and RAG2 proteins, core RAG1/RAG2 were found to be more soluble and much easier to isolate. To investigate whether the core regions of RAG1/RAG2 mediate the Ku-RAG interactions and whether the interactions are direct, purified Ku and core RAGs were employed for an in vitro direct binding assay.

Recombinant Ku70/Ku80 was purified by column chromatography with His•Bind Resin from Sf9 insect cells that were infected with baculovirus of VBB2-Kup70tH^{6}and VBB2-86Ku. The purity of the recombinant Ku70/Ku80 was determined by silver staining and shown in Figure 5. As proteins expressed in bacteria can be easily purified and are widely used for protein-protein interaction assay, core RAG2 cloned in pET-30c vector was expressed in *E. coli* BL21 and purified on TALON^{TM} resin by the his-tag present on the N-terminal of core RAG2 protein. Similar attempts to express core RAG1 cloned in pET-30c were unsuccessful. Maltose binding protein has been found to increase the expression and the solubility of recombinant

Figure 5. Elution of purified Ku70/Ku80. The Sf9 insect cells were infected with recombinant Baculovirus of VBB2-Kup70tH^{6}and VBB2-86Ku and incubated for 3 days. Recombinant Ku70/Ku80 was purified by column chromatography with His•Bind Resin from the whole cell extracts of Sf9 cells, which were kindly provided by Dominic Vallée. The presence of purified Ku70/Ku80 was assessed in fractions 1 to 6 (E1 to E6). Each elution is in a volume of 300 µl. The purity of the proteins was assessed by of electrophoresis in 8% sodium dodecyl sulfate (SDS) - polyacrylamide gel followed by silver staining.


E1
E2
E3
E4
E5
E6
Ku80
Ku70

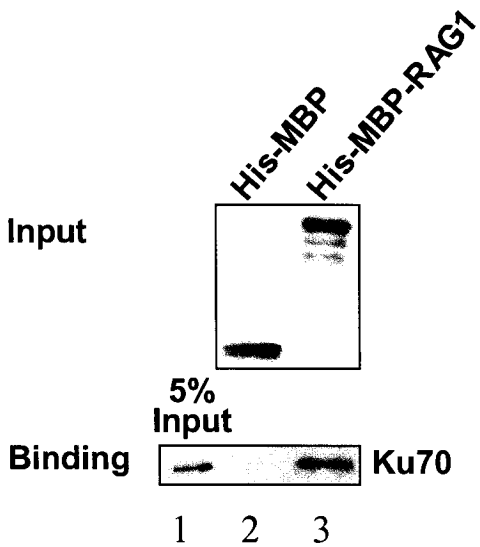
fusion RAG proteins (Rodgers et al., 1996). Therefore, maltose binding protein (MBP) fusion core RAG1 was cloned in pET-30b vector and was successfully expressed and purified on TALON^{TM} resin.

Previous work in our laboratory has shown that both RAG1 and RAG2 proteins are in the complex with Ku70/K80 via coimmunoprecipitation assays (Cui et al. in preparation). To examine whether RAG1 and/or RAG2 can directly interact with Ku, purified core RAG1 and RAG2 protein were separately tested for binding to purified Ku. To test the direct binding of core RAG1 proteins with Ku70/Ku80, his-tagged MBP fusion core RAG1 bound to TALON^{TM} resin were incubated with purified Baculo recombinant Ku70/Ku80. Binding assays were performed in the presence of ethidium bromide. EtBr prevents protein-DNA interactions which could provide a false impression of protein-protein interaction, whereas the two proteins are only bound to the same molecule of DNA (Lai and Herr, 1992). The presence of the Ku protein was detected with an anti-Ku70 antibody.

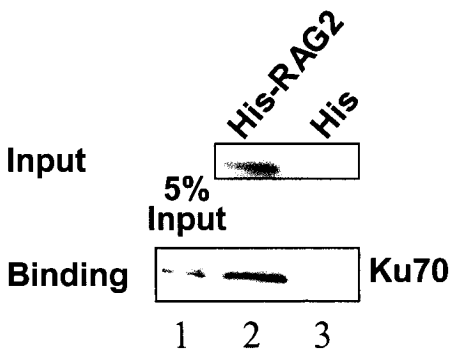
As shown in Figure 6A, his-tagged MBP fusion core RAG1 associated with Ku70/Ku80 and the interaction was DNA independent (lane 3). His-tagged MBP purified on TALON^{TM} resin was employed as a negative control and did not show any non-specific interaction with recombinant Ku70/Ku80. To further examine the input of his-tagged MBP fusion RAG1 and his-tagged MBP, the binding membrane was stripped first and rehybridized with an anti-MBP antibody. As also shown in Figure 6A input (lanes 2 and 3), his-tagged MBP fusion RAG1 and his-tagged MBP were used as similar

Figure 6. Interaction of purified core RAG1 and core RAG2 proteins with purified recombinant Ku. (A) Core RAG1 interacts with Ku. His-tagged MBP fusion core RAG1 was tested for binding with purified Baculo Ku70/Ku80 (lane 3). The presence of Ku70/Ku80 was detected with an anti-Ku70 antibody (lanes 1 to 3). His-tagged MBP was employed as a negative control (lane 2). Ethidium bromide (200 μg/ml) was included in the reactions in lanes 2 and 3. Binding of core RAG1 to Ku70/Ku80 was compared with 5% of Baculo Ku input (lane 1). Input of his-tagged MBP fusion core RAG1 and his-tagged MBP were detected with an anti-MBP antibody. **(B) Core RAG2 interacts with Ku.** His-tagged core RAG2 was tested for binding with purified Baculo Ku70/Ku80 (lane 2). The presence of Ku70/Ku80 was detected with an anti-Ku70 antibody (lanes 1 to 3). PET-30c vector alone was employed as a negative control (lane 3). Ethidium bromide (200 μg/ml) was included in the reactions in lanes 2 and 3. Binding of core RAG2 to Ku70/Ku80 was compared with 5% of Baculo Ku input (lane 1). Input of his-tagged core RAG2 and pET-30c vector were detected with an anti-RAG2 antibody.

A


His-MBP
His-MBP-RAG1
Input
5%
Input
Binding
Ku70
1 2 3

B


His-RAG2
His
Input
5%
Input
Binding
Ku70
1 2 3

amount. Therefore, I concluded that Ku70/Ku80 directly associates with core RAG1 and the Ku-RAG1 interaction is DNA independent.

Using the same direct binding assay, his-tagged core RAG2 bound to TALON^{TM} resin was also examined for interaction with purified Ku70/Ku80 in the presence of ethidium bromide (Figure 6B). pET-30c, the empty vector expressing a 92 amino acid his-tagged peptide, was employed as a negative control. This peptide showed no binding to Ku70/Ku80. To further examine the input of his-tagged RAG2, The membrane was rehybridized with an anti-RAG2 antibody to verify the presence of his-tagged RAG2 (Figure 6B input lanes 2 and 3). These findings demonstrate that, similar to core RAG1, core RAG2 can also associate with Ku70/Ku80 directly and the interaction between RAG2 and Ku is DNA independent.

In summary, both core RAG1 and core RAG2 can directly interact with Ku70/Ku80 and the interactions are DNA independent.

3.2 A DNA binding defective mutant in NBR can associate with Ku70/Ku80

NBR, the only well-characterized DNA binding domain in RAG1, has been shown to bind specifically to nonamer sequence of RSS (Spanopoulou et al., 1995; Difilippantonio et al., 1996). RAG1 K405A/H406A/R407A was shown to have severe defects in DNA-binding, nicking and hairpin formation on a radiolabeled 12RSS oligonucleotide substrate (Huye et al., 2002). In order to examine whether nonamer binding activity is dispensable for the Ku-RAG1 interaction, RAG1

K405A/H406A/R407A was purified on TALON™ His resin and was tested for binding with Ku70/Ku80 from the NEs of Jurkat T cells in the presence of ethidium bromide. The presence of Ku70/Ku80 was detected with an anti-Ku70 antibody.

As shown in Figure 7, his-tagged MBP fusion RAG1 K405A/H406A/R407A was able to interact with Ku70/Ku80 as well as his-tagged MBP fusion core RAG1 protein (lanes 2 and 3).

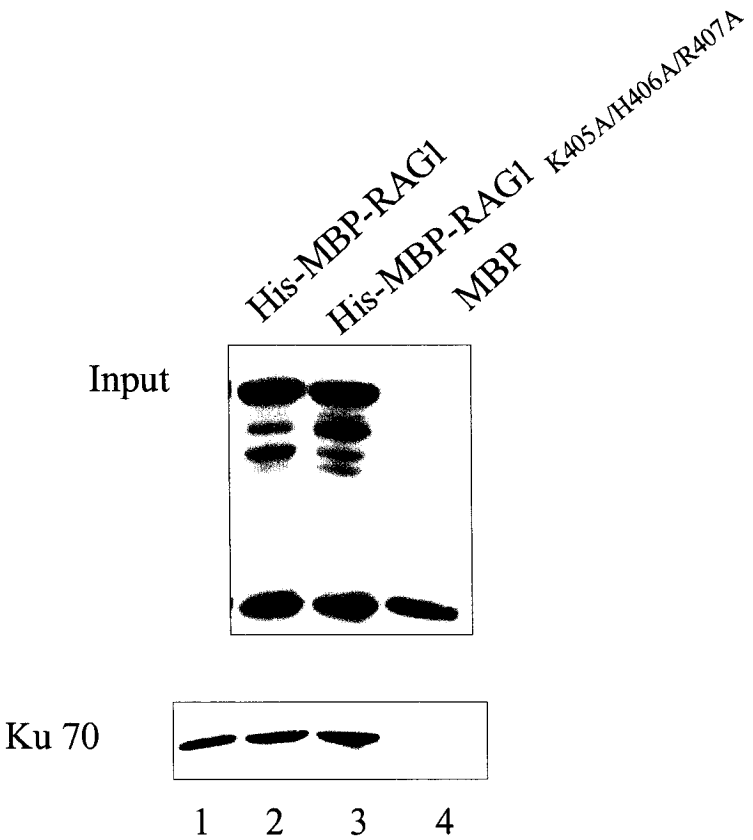
In summary, Ku-interaction mediated through the core region of RAG1 is independent of the DNA binding activity of RAG1. As Ku70/Ku80 has been previously demonstrated in our laboratory to recognize heptamer sequence of the RSS, this mutant would be used in future experiment to determine if Ku can recruit the RAG1 DNA-binding defective mutant to the RSS and therefore recover the DNA cleavage by RAG1/RAG2.

3.3 Mapping of Ku-interacting region in RAG1 protein

In order to define the Ku-interacting region in RAG1, an in vitro protein binding assay was employed, in which RAG1 deletion mutants were in vitro translated as ^{35}S labeled protein by using T7 TnT® rabbit reticulocyte lysate kit and then were tested for binding with immunoprecipitated Ku from WCE from Jurkat T cells. Binding assays were performed in the presence of EtBr to prevent possible non-specific protein-DNA interactions (Lai and Herr, 1992).

DNA constructs of RAG1 deletion mutants were all created from

Figure 7. DNA-binding defective mutant RAG1 K405A/H406A/R407A interacts with Ku. RAG1 K405A/H406A/R407A was created by site-directed mutagenesis using pET-MBP-R1 as the backbone. RAG1 K405A/H406A/R407A was purified on TALON^{TM} His resin and was tested for binding with Ku protein from 60 μg Jurkat T cell nuclear extracts. Ku binding was assessed by Western Blot analysis using a Ku70 antibody (lane 3). His-tagged MBP core RAG1 was used as a positive control (lane 2) and MBP as a negative control (lane 4). The binding of Ku was compared with 10% of RAG1 protein input.

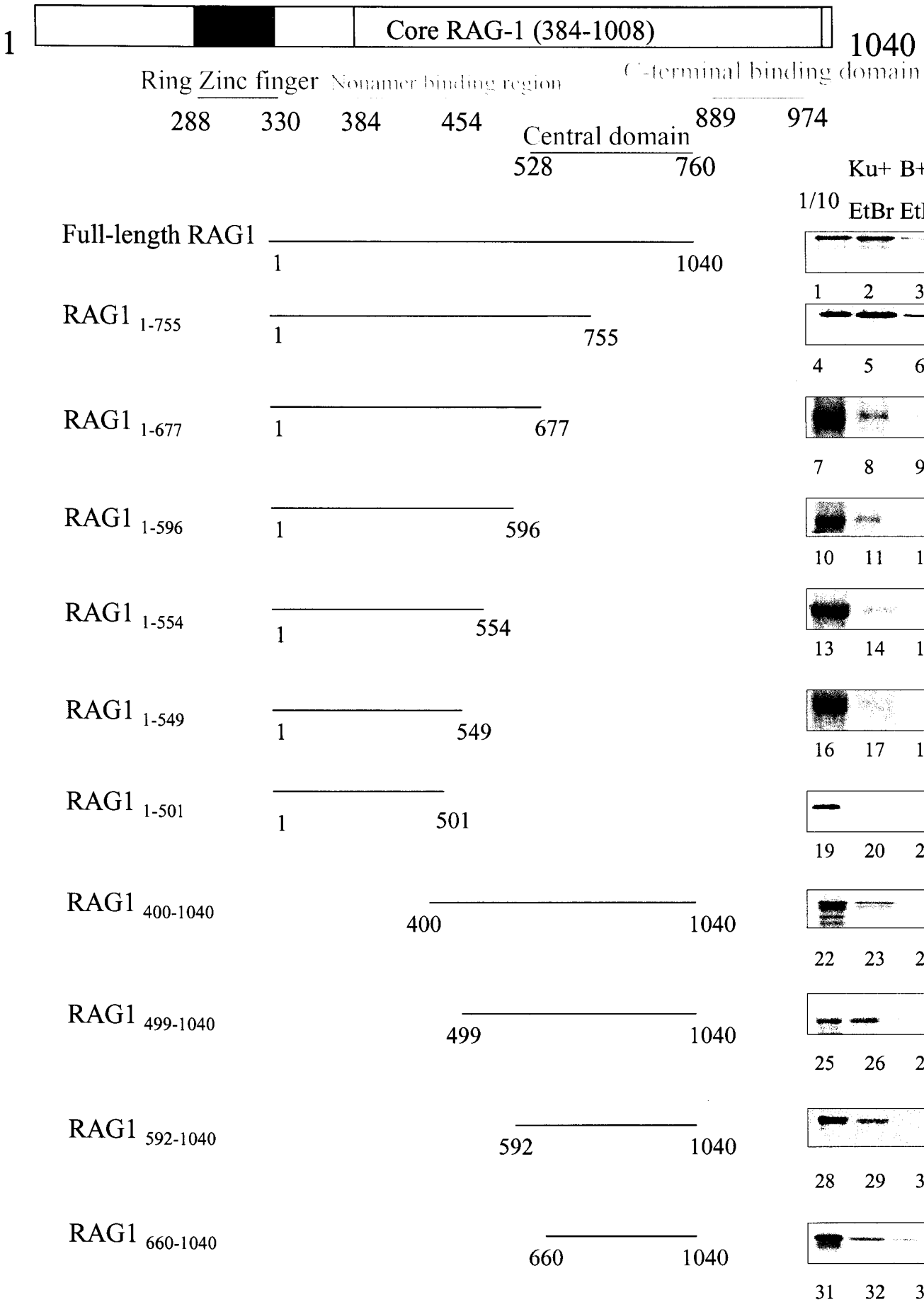

His-MBP-RAG1
His-MBP-RAG1 K405A/H406A/R407A
MBP
Input
Ku 70
1
2
3
4

pRAG1(M2CD7) vector which is under the control of a T7 promoter. To obtain RAG1 $_{1\text{-}755}$ and RAG1 $_{1\text{-}501}$, the pRAG1 vector was digested with restriction enzyme AfeI and SphI, respectively. All the other N-terminus and C-terminus deletion mutants of RAG1 were generated by PCR amplification with 5' primers containing the T7 promoter sequence (see **Appendix I** for primer sequences). The deletion mutants were then in vitro translated as ^{35}S labeled proteins and tested for binding to Ku. As a positive control for this system, full-length RAG1 encoded by pRAG1 (M2CD7) vector was shown to associate with Ku (Figure 8 lanes 1 to 3), which is consistent with previous work by Dr. Caroline Schild-Poulter in our laboratory.

As shown in Figure 8 lanes 4 to 6 and lanes 19 to 21, RAG1 deletion up to C-terminus amino acid residue 755 can still interact with Ku70/Ku80, while further deletion to amino acid residue 501 abolished its ability to bind to Ku70/Ku80. In addition, RAG1 $_{1\text{-}755}$ yielded similar Ku-binding to the full-length RAG1 (Figure 8 lanes 1 to 3 and 4 to 6).

To exclude that the lost of binding ability of RAG1 $_{1\text{-}501}$ to Ku was due to the disruption of the secondary structure of this mutant and also to determine the N-terminal boundary of Ku-interacting region in RAG1, N-terminal deletions up to amino acid residue 400 and 499 were tested for binding to Ku. As expected, both of the mutants were able to interact with Ku70/Ku80 (lanes 22 to 24 and 25 to 27) and the binding of RAG1 $_{499\text{-}1040}$ was similar to that of RAG1 $_{1\text{-}755}$ and full-length RAG1. This suggests that the Ku-interacting domain in RAG1 is localized between amino acid residue 500 and

Figure 8. Summary of full-length RAG1 and RAG1 mutants binding to Ku70/Ku80. Above is a brief schematic representation of RAG1 protein. Left is the schematic representation of RAG1 proteins used in this assay. The numbering refers to the amino acid context. Right is the summary of the Ku-binding results of each individual RAG1 constructs. Full-length RAG1 and RAG1 mutants were in vitro translated as ^{35}S labeled proteins and were tested for binding with immunoprecipitated Ku from 300 μg Jurkat T cell extracts in the presence of ethidium bromide (200 μg/ml). RAG1-Ku complexes were resolved on SDS- polyacrylamide gels and the RAG proteins were visualized with PhosphorImager software (lanes 2, 5, 8, 11, 14, 23, 26, 29 and 32). Protein A-Sepharose beads (B+ EtBr) were employed as negative control.


1
Core RAG-1 (384-1008)
1040
Ring Zinc finger
Nonamer binding region
C-terminal binding domain
288
330
384
454
889
974
Central domain
528
760
Ku+ B+
1/10 EtBr EtBr
Full-length RAG1
1
1040
1 2 3
RAG1 1-755
1
755
4 5 6
RAG1 1-677
1
677
7 8 9
RAG1 1-596
1
596
10 11 12
RAG1 1-554
1
554
13 14 15
RAG1 1-549
1
549
16 17 18
RAG1 1-501
1
501
19 20 21
RAG1 400-1040
400
1040
22 23 24
RAG1 499-1040
499
1040
25 26 27
RAG1 592-1040
592
1040
28 29 30
RAG1 660-1040
660
1040
31 32 33

755, which corresponds to the central domain in RAG1 spanning from amino acid residue 528 to 760 (Arbuckle et al., 2001).

To map the C-terminal boundary of the Ku-interacting domain in RAG1 in more detail, several C-terminal deletion mutants between amino acid residues 500 and 755 were created and tested for binding to Ku70/Ku80. RAG1 C-terminal deletion mutant RAG1 $_{1\text{-}677}$ was still able to interact with Ku70/Ku80 but the binding efficiency was reduced (lanes 7 to 9). Further C-terminal deletion mutant RAG1 $_{1\text{-}596}$ yielded similar Ku-binding to RAG1 $_{1\text{-}677}$ (lanes 10 to 12). These findings suggest that C-terminal deletion further than amino acid residue 755 resulted in severe decrease of Ku binding. A C-terminal deletion mutant RAG1 $_{1\text{-}554}$ could barely associate with Ku70/Ku80 while the binding of RAG1 $_{1\text{-}549}$ mutant was hardly detectable (lanes 13 to 18).

These findings indicate that: (a) Deletion of the region between amino acid residue 677 and 755 of RAG1 substantially reduced the binding to Ku70/Ku80, as shown by the severe decrease of Ku binding between RAG1 $_{1\text{-}755}$ and RAG1 $_{1\text{-}677}$ shown in Figure 8 lanes 4 to 6 and 7 to 9; (b) The region spanning amino acid residue 1 to 549 is not involved in interacting with Ku70/Ku80.

To further define the N-terminal boundary of the Ku-interacting region in core RAG1, two more N-terminal deletion mutants RAG1 $_{592\text{-}1040}$ and RAG1 $_{660\text{-}1040}$ were further tested. As shown in lanes 28 to30 and 31 to 33, both RAG1 $_{592\text{-}1040}$ and RAG1 $_{660\text{-}1040}$ mutants can associate with Ku70/Ku80.

In summary, (a) the region spanning from amino acid residue 550 to 755 that corresponds to the central domain of RAG1 is suggested to be involved in the interaction with Ku70/K80; (b) The region spanning from amino acid 555 to 677 is suggested to help to stabilize RAG1 binding to Ku.

3.4 The region spanning from amino acid residue 660 to 763 in RAG1 is involved in the interaction with Ku

In order to refine the determination of the Ku-interacting region in RAG1, three isolated peptides RAG1 $_{499\text{-}763}$, RAG1 $_{573\text{-}763}$ and RAG1 $_{660\text{-}763}$ were expressed as MBP fusion proteins. They were purified on amylose resin and eluted out from the resin. The purity of these RAG1 fusion proteins was determined by coomassie staining and is shown in Figure 9. In the presence of ethidium bromide, they were tested for binding to Ku70/Ku80 from the NEs of Jurkat T cells followed by immunoprecipitation of Ku. The presence of RAG-Ku complex was then detected with anti-MBP antibody.

The three RAG1 peptides used in this study are schematically shown in Figure 10A. His-tagged MBP fusion core RAG1 was employed as the positive control for this assay. For reasons still unclear, this construct can not be efficiently purified on amylose resin. Therefore, it was purified on TALON^{TM} His resin and then eluted. This may have to do with the presence of the 6X histine tail at the N-terminal of MBP fusion core RAG1, which could potentially interfere with its purification on amylose resin. Purified MBP alone was employed as a negative control.

Figure 9. Elution of purified MBP fusion RAG1 mutants. RAG1 $_{499\text{-}763}$, RAG1 $_{573\text{-}763}$ and RAG1 $_{660\text{-}763}$ were expressed as MBP fusion RAG1 proteins. They were purified on amylose resin and were eluted with 10mM maltose. 8 μl MBP fusion RAG1 $_{499\text{-}763}$ (lane 1), 10 μl MBP fusion RAG1 $_{573\text{-}763}$ (lane 2), 10 μl MBP fusion RAG1 $_{660\text{-}763}$ (lane 3), 4 μl MBP (lane 4) and 6 μl MBP fusion RAG1 $_{792\text{-}1013}$ (lane 5) were resolved in 8% sodium dodecyl sulfate (SDS)- polyacrylamide gel followed by coomassie staining.

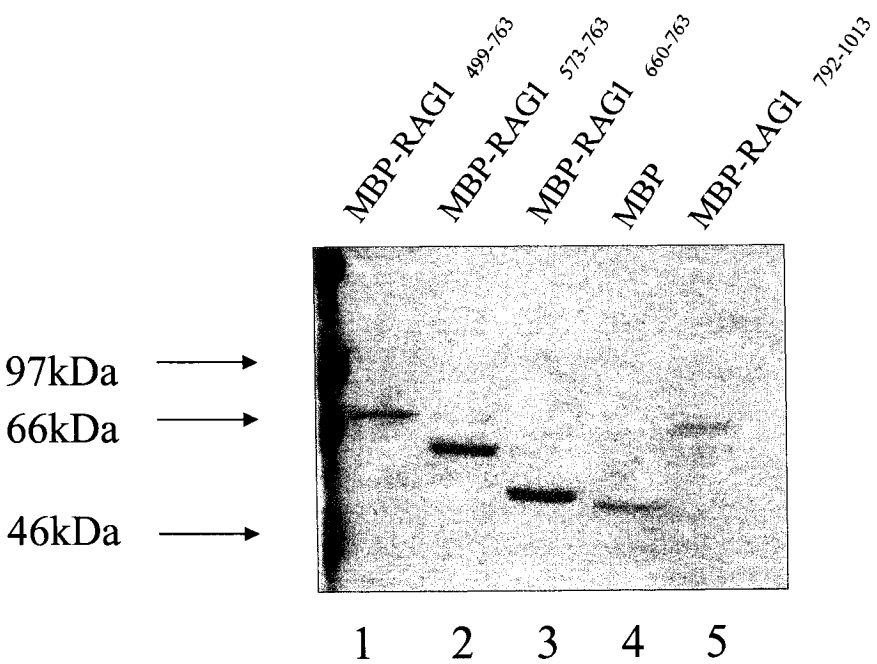
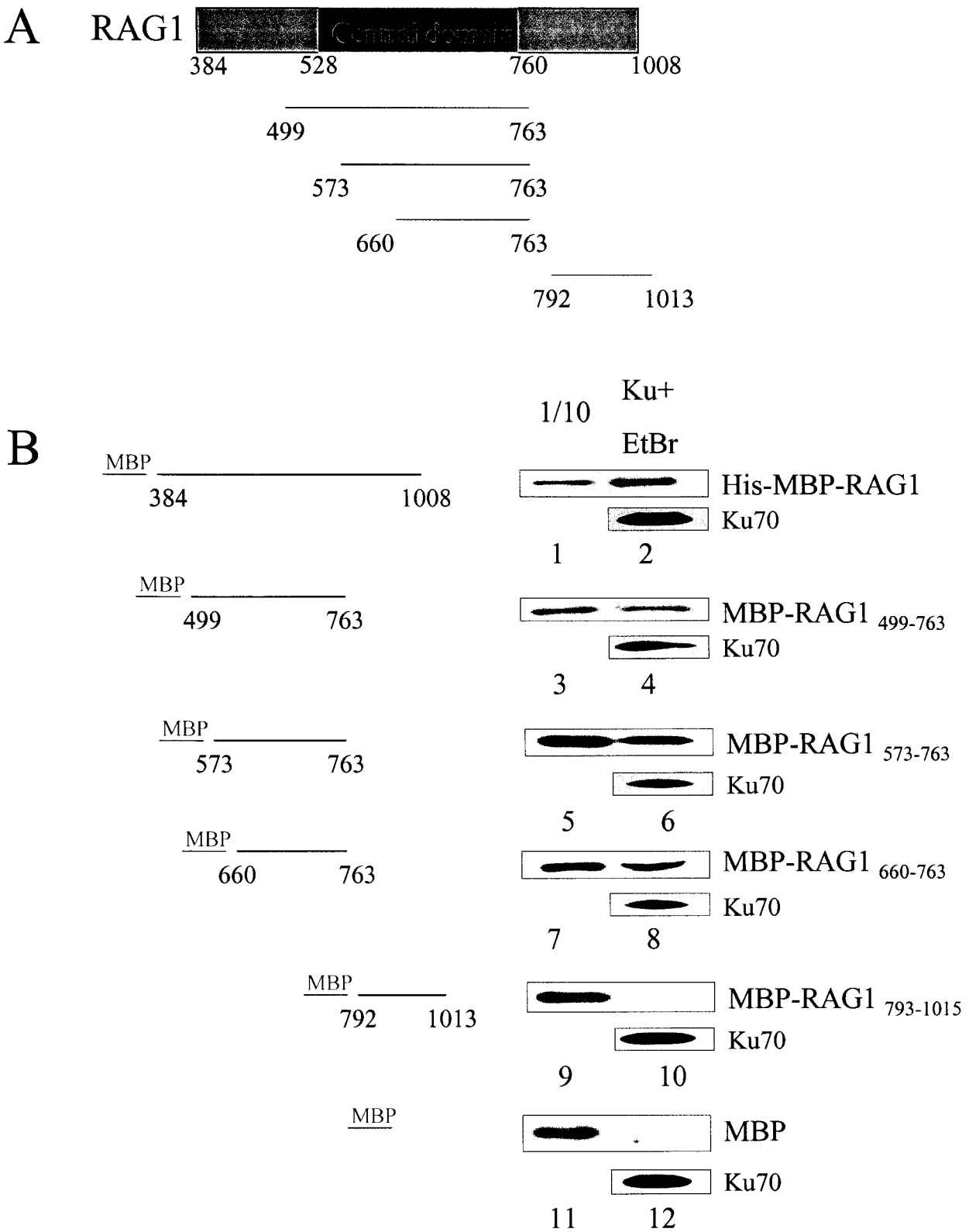

MBP-RAG1 499-763
MBP-RAG1 573-763
MBP-RAG1 660-763
MBP
MBP-RAG1 792-1013
97kDa
66kDa
46kDa
1
2
3
4
5

Figure 10. The region spanning from amino acid residue 660 to 763 interacts with Ku. (A) Schematic of RAG1 core region and the constructs used in the MBP fusion RAG1 binding assay. The numbering refers to the amino acid context. **(B) Summary of binding of MBP fusion RAG1 proteins to Ku.** After elution from the amylose resin, purified MBP fusion RAG1 mutants RAG1 $_{499-763}$, RAG1 $_{573-763}$, RAG1 $_{660-763}$ and RAG1 $_{792-1013}$ were allowed to bind to Ku from 60 μg Jurkat T cell nuclear extracts followed by Ku immunoprecipitation with a Ku70 antibody. Ethidium bromide (200μg /ml) was included throughout the binding reactions. The presence of MBP fusion RAG1 mutant was detected with an anti-MBP antibody (lanes 4, 6 and 8). His-tagged MBP fusion core RAG1 was employed as a positive control (lane 2) while MBP alone as a negative control (lane 12). The binding with Ku was compared with 10% of the input MBP fusion RAG1 proteins. Immunoprecipitated Ku was detected with an anti-Ku70 antibody.


A
RAG1
Central domain
384
528
760
1008
499
763
573
763
660
763
792
1013
B
Ku+
1/10
EtBr
MBP
384
1008
His-MBP-RAG1
Ku70
1
2
MBP
499
763
MBP-RAG1 499-763
Ku70
3
4
MBP
573
763
MBP-RAG1 573-763
Ku70
5
6
MBP
660
763
MBP-RAG1 660-763
Ku70
7
8
MBP
792
1013
MBP-RAG1 793-1015
Ku70
9
10
MBP
MBP
Ku70
11
12

As shown in Figure 10B lanes 4, 6 and 8, three MBP fusion RAG1 mutants RAG1 $_{499\text{-}763}$, RAG1 $_{573\text{-}763}$ and RAG1 $_{660\text{-}763}$ were all able to associate with Ku70/Ku80 with similar binding efficiency, while MBP alone did not show any binding to Ku70/Ku80 (lane 12).

C-terminal region spanning from amino acid residue 889 to 974 was defined to be the coding flank binding region (Mo et al., 2001). Point mutations in the C-terminal domain have demonstrated that C-terminal domain is important for joining step in V(D)J recombination (Huye et al., 2002). In order to examine whether C-terminal DNA binding domain of RAG1 would be an additional domain binding to Ku70/Ku80, the peptide of RAG1 $_{792\text{-}1013}$, which contains the intact coding flank binding region in RAG1 was expressed as MBP fusion protein. It was allowed to interact with Ku70/Ku80 from Jurkat NE in the presence of ethidium bromide. Following the immunoprecipitation of Ku from the binding reaction, the RAG-Ku complex was detected with an anti-MBP antibody. As shown in Figure 10B lane 10, MBP fusion RAG1 $_{792\text{-}1013}$ did not show any interaction with Ku70/Ku80. This finding suggests that coding flank binding region is not likely to be involved in the interaction between RAG1 and Ku.

In summary, these experiments define that the region spanning from amino acid residue 660 to 763 in the core region of RAG1 is sufficient for interaction with Ku.

3.5 Mapping of Ku-interacting domain in RAG2 protein

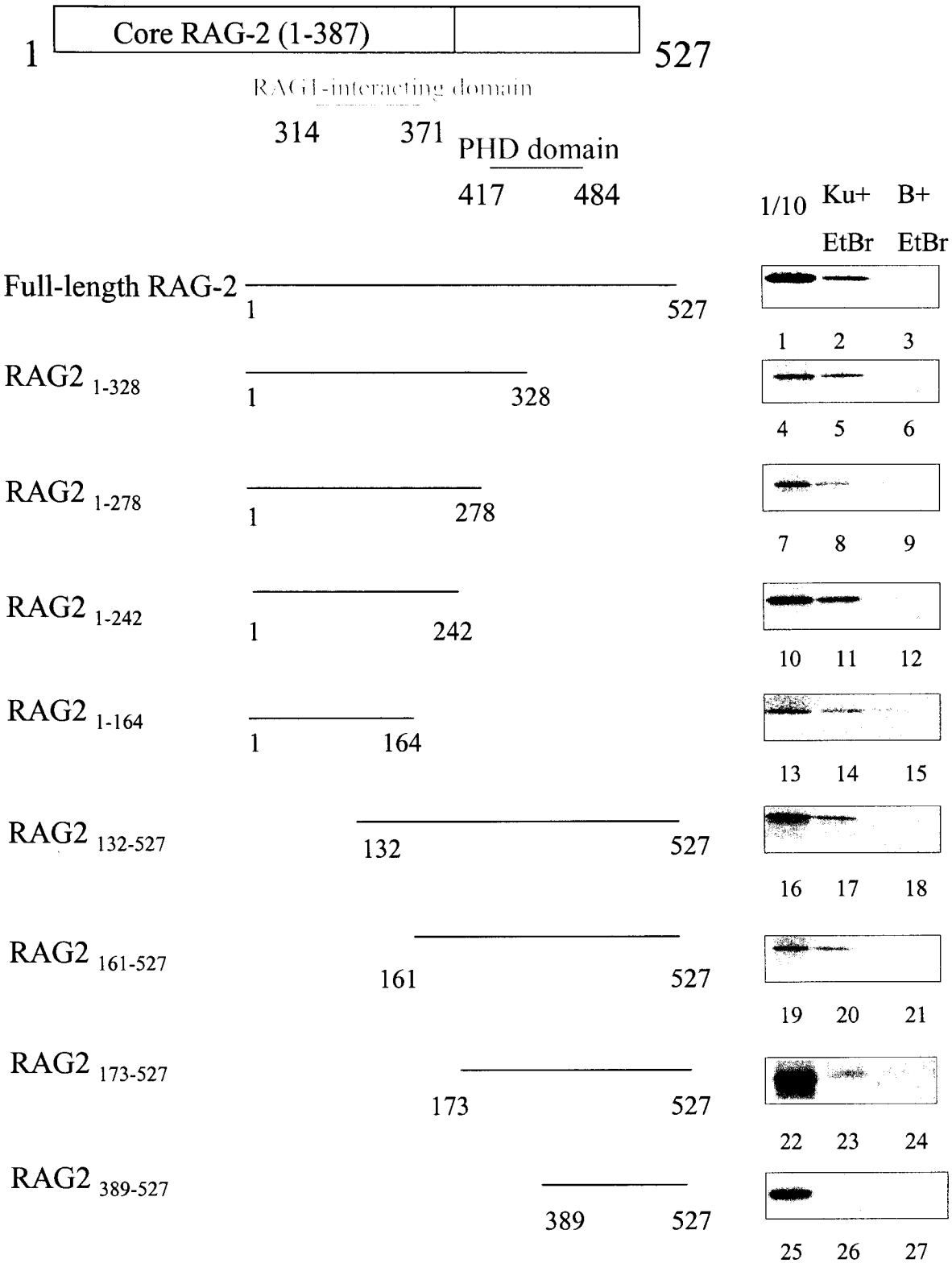
Previous work in our laboratory has shown that RAG2 alone can also interact

with Ku70/Ku80 in vitro (Cui et al. in preparation). To study the Ku-interacting region in RAG2, RAG2 mutants were in vitro translated and tested for binding with Ku in the same way I performed the binding assays just described for RAG1.

DNA constructs of RAG2 deletion mutants were all created from pRAG2 (R2RCD2) vector that is under the control of a T7 promoter. The mutant cDNA was either generated by restriction enzyme digestion or produced by PCR with primers containing a T7 promoter sequence (see **Appendix I** for primer sequences). The deletion mutants were then in vitro translated as ^{35}S labeled proteins and tested for binding to Ku. As a positive control for this system, full-length RAG2 encoded by pRAG2 (R2RCD2) vector was shown to associate with Ku (Figure 11 lanes 1 to 3).

In the initial experiments, progressive C-terminal truncations of RAG2 were examined. As shown in Figure 11 lanes 4 to 6, RAG2 C-terminal deletion mutant up to amino acid residue 328 can bind to Ku70/Ku80. In addition, the binding of RAG2 $_{1-328}$ was similar to that of full-length RAG2. This result suggests that the Ku-interacting domain in RAG2 remains intact in the RAG2 $_{1-328.}$ However, when RAG2 C-terminal was truncated to amino acid residue 278, the binding to Ku70/Ku80 was substantially decreased (lanes 7 to 9), which suggested the involvement of the region spanning amino acid residue 278 to 328 in the Ku-RAG2 interaction. Unexpectedly, further deletion up to amino acid residue 242 resulted in a recovery of Ku70/Ku80 binding to an efficiency similar to that of full-length RAG2 protein (lanes 1 to 3 and 10 to 12). This binding result of RAG2 $_{1-242}$ suggests that the decrease of RAG2 $_{1-278}$ binding could be due to an

Figure 11. Summary of full-length RAG2 and RAG2 mutants binding to Ku70/Ku80. Above is a brief schematic representation of RAG2 protein. Left is the schematic representation of RAG2 proteins used in this assay. The numbering refers to the amino acid context. Right is the summary of the Ku-binding results of each individual RAG2 constructs. Full-length RAG2 and RAG2 mutants were in vitro translated as ^{35}S labeled proteins and were tested for binding to immunoprecipitated Ku from 300 μg Jurkat T cell extracts in the presence of ethidium bromide (200 μg/ml). RAG2-Ku complexes were resolved on SDS- polyacrylamide gel and the RAG proteins were visualized with PhosphorImager (lanes 2, 5, 8, 11, 14, 17, 20 and 23). Protein A-Sepharose beads (B+ EtBr) were employed as negative control. Binding of full-length RAG2 and RAG2 mutants were compared with the 10% of ^{35}S labeled protein input.


1
Core RAG-2 (1-387)
527
RAG1-interacting domain
314
371
PHD domain
417
484
1/10
Ku+
EtBr
B+
EtBr
Full-length RAG-2
1
527
1 2 3
RAG2 1-328
1
328
4 5 6
RAG2 1-278
1
278
7 8 9
RAG2 1-242
1
242
10 11 12
RAG2 1-164
1
164
13 14 15
RAG2 132-527
132
527
16 17 18
RAG2 161-527
161
527
19 20 21
RAG2 173-527
173
527
22 23 24
RAG2 389-527
389
527
25 26 27

inhibitory region between amino acid residue 242 and 278.

In order to further define the Ku-interacting region within the region of RAG2 $_{1\text{-}242}$, RAG2 C-terminal was further deleted to amino acid residue 164. As shown in Figure 11 lanes 13 to 15, RAG2 $_{1\text{-}164}$ was still able to interact with Ku70/Ku80, however, compared to the full-length RAG2 and RAG2 $_{1\text{-}242}$, the binding was weaker. Two possibilities could explain the decreased binding of RAG2 $_{1\text{-}164}$. One is that the major Ku-binding domain in RAG2 would be localized in the region between amino acid residue 164 and 242. Therefore, the deletion of this part would result in the decrease in binding to Ku. The second possibility would be that the deletion up to amino acid residue 164 destroyed the overall structure of the potential Ku-interacting domain localized in the first 242 amino acids at the N-terminal of RAG2 protein.

Due to the size labeling limitation, RAG2 C-terminal deletion further than residue 164 was not feasible. Therefore, complementary N-terminal deletions of RAG2 protein were employed to refine my definition of the Ku-interacting domain in RAG2 protein. As shown in Figure 11 lanes 16 to 18, deletion of the N-terminal 132 amino acids did not alter the RAG2 binding to Ku and the binding of this construct appeared similar to that of RAG2 $_{1\text{-}242}$. This suggests that the Ku-interacting domain in RAG2 is present in the region spanning from amino acid residue 132 to 527. Further N-terminal deletion up to amino acid 161 still remained the ability to interact with Ku (lanes 19 to 21). However, as shown in lanes 22 to 24, when the first 173 amino acids at the N-terminal of RAG2 were deleted, the association of Ku and RAG2 was greatly

diminished. These findings suggest that the region from amino acid residue 161 to 173 might be the major region that mediates the interaction between RAG2 and Ku70/Ku80 in the region between amino acid residue 132 and 242. Alternatively, this region might be important for the structure support for the Ku-interacting domain in RAG2 that is localized in the region spanning from amino acid residue 132 to 242.

Finally, the region spanning from amino acid residue 389 to 527 was tested for binding to Ku70/Ku80. As expected, RAG2 $_{389\text{-}527}$ was incapable of interaction with Ku70/Ku80 (lanes 25 to 27), which suggests that the non-core region of RAG2 from amino acid residue 387 to 527 is not involved in interaction with Ku. This result further confirmed that the interaction between RAG2 and Ku is mediated by the core region of RAG2.

In summary, according to the binding results from both N-terminal and C-terminal truncations of RAG2, the region spanning from amino acid residue 132 to 242 was defined to be required for the interaction between RAG2 and Ku70/Ku80. This region maps to the 3rd and 4th kelch-motif repeats of the predicted RAG2 structure model. However, the further examination of the Ku-interacting region in RAG2 proved not to be feasible, as both RAG2 $_{1\text{-}132}$ and RAG2 $_{242\text{ to }389}$ constructs that I prepared could not be efficiently in vitro translated (data not shown).

4. Discussion

4.1 Definition of Ku-interacting domain in RAG1 and RAG2 proteins

In the present study, I have successfully set up an in vitro protein-protein interaction system to examine the Ku-interacting domains in both RAG1 and RAG2 proteins. I have been able to show that both core RAG1 and core RAG2 proteins can interact directly with Ku and these interactions are DNA-independent. Although the region in RAG2 that is both necessary and sufficient for Ku binding remains to be precisely determined, I have been able to define that the peptide spanning from amino acid residue 660 to 763 in the central domain of RAG1 was sufficient for the interaction with Ku.

Previous work in our laboratory showed that Ku can interact with most homeodomain proteins (Schild-Poulter et al., 2001). However, unexpectedly, both RAG1 $_{1\text{-}501}$ and RAG1 $_{1\text{-}549}$ mutants, which contain the intact homeodomain-like motif spanning from amino acid residue 387 to 446 in the NBR, were unable to interact with Ku in this study. This suggests that it was unlikely that Ku could interact with the RAG1 homeodomain-like motif. The primary sequence of RAG1 homeodomain is more homologous to the homeodomain of the Salmonella typhimurium Hin invertase than to other homeodomain structures (Spanopoulou et al., 1996). Compared to the eukaryotic homeodomain protein, RAG1 homeodomain-like motif is shown to maintain a helical configuration of the corresponding helix II between Hin and eukaryotic homeodomain proteins, but contains a longer turn between helix II and helix III. The homeodomain

structure difference between RAG1 and other homeodomain proteins may explain why the homeodomain-like motif is not required for Ku binding. However, it is also possible although much less likely that, with the additional amino acid residues, the homeodomain-like motif in RAG1 is not properly folded.

The Ku-interacting domain spanning from amino acid residue 660 to 763 of RAG1 defined in this study is localized in the central domain of RAG1, which contains the ZFB (Figure 2). The ZFB, spanning from amino acid residue 722 to 760, has been suggested to be the predominant region for recruiting RAG2 protein to RAG1 and is also involved in non-specific DNA binding (Aidinis et al., 2000; Arbuckle et al., 2001).

The RSS heptamer itself does not appear to provide a significant interface for DNA binding by RAG1/RAG2 complex but it is critical for DNA cleavage by the RAG1/RAG2 complex (Aidinis et al., 2000; Peak et al., 2003). DNA cleavage by RAG1/RAG2 requires the unpairing of base pairs at the border of the RSS heptamer and the coding sequences (Cortes et al., 1996; Ramsden et al., 1996). Previous work in our laboratory has shown that Ku can physically interact with RAG1 and RAG2 proteins via coimmunoprecipitation assays (Cui et al. in preparation). These findings suggested Ku may participate in V(D)J recombination process before the joining step which is mediated by the NHEJ apparatus (Bassing et al., 2002; Gellert, 2002; Ma et al., 2002; Mahajan et al., 2002; Bertocci et al., 2003; McElhinny and Ramsden, 2003).

The results shown here suggest that the close proximity of the Ku-interacting domain and the RAG2-interacting domain in RAG1 protein has direct relevance to the

occupancy of heptamer of RSS and the RSS cleavage by RAG1/RAG2. The interaction between RAG and Ku might activate the helicase activity of Ku. The helicase activity of Ku would induce the structure transitions at 5' of the heptamer or locally unwind the DNA helix (Tuteja et al., 1994; Ramsden et al., 1996; Ochem et al., 1997), which allows for an enhanced RSS heptamer binding by the RAG1/RAG2 complex and therefore enhances the RSS cleavage RSS.

The definition of the Ku-interacting region in RAG2 proved to be more challenging and the exact surface(s) of interaction remains to be determined. The binding results of RAG2 $_{1\text{-}328}$ and RAG2 $_{389\text{-}532}$ suggested that the Ku-interacting region was localized within the N-terminal 328 amino acid residues in the core region of RAG2 protein. Upon deletion of the C-terminal domain of RAG2 up to amino acid residue 278, the binding to Ku was strongly reduced. Unexpectedly, a further C-terminal deletion up to amino acid residue 242 recovered its binding with Ku. There are at least two possibilities that could explain this result. One is that the region between amino acid residue 242 and 278 has an inhibitory effect on the folding of the RAG2 $_{1\text{-}278}$ mutant and deletion of this region would expose the Ku binding domain in RAG2. The second possibility would be that the deletion up to amino acid residue 278 disrupted the fifth repeat in the core RAG2 spanning from amino acid residue 254 to 305, which might be a potential domain in the C-terminal region of RAG2 protein that is involved in the interaction with Ku. Based on the finding that RAG2 $_{132\text{-}527}$ can bind to Ku similarly to RAG2 $_{1\text{-}242,}$ we conclude that the Ku-interacting domain in RAG2 is localized in the

region between amino acid residue 132 and 242. Further attempts to clarify these results by testing the binding activity of RAG2 $_{1\text{-}132}$ and RAG2 $_{242\text{-}389}$ failed, as the peptides were not efficiently in vitro translated (data not shown).

RAG2 $_{161\text{-}532}$ and RAG2 $_{173\text{-}532}$ were employed to further determine the Ku-interacting region. While RAG2 $_{161\text{-}532}$ can associate with Ku similarly to RAG2 $_{132\text{-}527}$ and RAG2 $_{1\text{-}242}$, RAG2 $_{173\text{-}532}$ binds to Ku with a strongly reduced efficiency. It is possible that the region between amino acid residue 161 and 173 is the Ku-interacting region in the core region. Alternatively, this region is important for the folding of Ku-interacting structure in the region between amino acid residue 132 and 242.

The core region of RAG2 was suggested to be composed of six kelch/mipp internal repeats to form a six-bladed β -propeller –like 3D structure (Callebaut and Mornon, 1998). This structure has been suggested to be involved in mediating protein-protein interaction and protein-DNA interaction through the relatively flat surfaces presented on both sides of the structure formed by loops linking the β strands in each repeats (Callebaut and Mornon, 1998).

Due to the complexity of the proposed 3D structure of core RAG2, so far the RAG1 interacting domain in RAG2 spanning from amino acid residue 314 to 371 is the only defined protein-interacting domain in core RAG2 and this region coincides with the sixth kelch motif (Aidinis et al., 2000). Thus, it is possible that the other five repeats in the RAG2 N-terminus might be available to interact with other proteins involved in V(D)J recombination.

The amino acid positions 61, 128, 201, 253, 305 and 351 were suggested to define the six kelch motifs in RAG2 (Callebaut and Mornon, 1998; Aravind and Koonin, 1999). The region spanning from amino acid residue 132 to 242 was shown to be important for the interaction with Ku and this region maps with the 3rd and 4th repeat motif of the core RAG2 protein. In addition, region between amino acid 161 and 173 is localized within the predicted 2-3 loop of the 3rd repeat, which is the longest 2-3 loop of the six repeats (Callebaut and Mornon, 1998). This observation raises the possibility that the 2-3 loop of the 3rd repeat may directly mediate the RAG2 and Ku interaction in the region between amino acid residue 132 and 242. Due to the complexity of RAG2 structure, deletion of this loop might destroy the overall structure of RAG2. Therefore, in the further experiment, it is worthwhile to express the 3rd and the 4th repeat of motifs of RAG2 as a fusion protein and test for binding to Ku and then find the point mutants in the 2-3 loop of the 3rd repeat that are unable to interact with Ku.

4.2 Potential Implication of RAG-Ku interaction for the formation of synaptic complex between a 12RSS and a 23RSS

RAG1 can bind to an individual 12RSS or 23RSS in the complex with the RAG2 protein and forms a presynaptic complex in vitro, which is catalytically incompetent (Hiom and Gellert, 1997; Hiom and Gellert, 1998; Jones and Gellert, 2002; Mundy et al., 2002). It has been suggested that the assembled protein complex subsequently formed a competent synaptic complex by incorporating a second 23RSS or

12RSS that is free of RAG1/RAG2 complex (Jones and Gellert, 2002). RAG1/RAG2 proteins are active for DNA cleavage when bound to a single RSS in vitro but display a reduced cleavage efficiency compared with the cleavage of a paired 12/23 RSS complex (Eastman et al., 1996; Sawchuk et al., 1997; Jones and Gellert, 2002). RAG1/RAG2 complemented with 293T WCEs has been shown to strictly obey the 12/23 rule in vitro (Sawchuk et al., 1997). In a recent study, Ku70/Ku80 is found to be required to preferentially inhibit the cleavage activity of the 12/12 and 23/23 substrates via an immunodepletion assay. Furthermore, by using purified RAG1/RAG2, HMG1 /HMG2, Ku70/Ku80 and DNA-PKcs in a cis-cleavage assay, the 12/23 rule is reconstituted in vitro. Therefore, Ku70/Ku80 and DNA-PKcs are suggested to enforce the 12/23 rule in vitro by preferentially inhibiting both the 12/12 and 23/23 RSS cleavage (Sawchuk et al., 2004).

The involvement of Ku70/Ku80 and DNA-PKcs in the 12/23 rule might be mediated by the interaction between RAG and Ku. Interaction of Ku with RAG1/RAG2 complex may cause a conformation change on 12RSS and 23RSS, which may result in the preference in binding to a 12RSS by the RAG1/RAG2 complex. RAG1/RAG2 complex has been shown to be more active on a 12RSS than on a 23RSS (Kwon et al., 2000). Evidence suggests that the synaptic complex includes two subunits of RAG2 protein and at least a dimer, or probably a trimer or tetramer of RAG1 proteins (Swanson and Desiderio, 1999; Landree et al., 2001; Mundy et al., 2002). It is possible that when a tetramer complex of RAG1 homodimer and two subunits of RAG2 proteins bind to the

nonamer sequence of a 12RSS, one of the binding site of the tetramer RAG1/RAG2 is locked. Then the specific conformation of the RAG1/RAG2/Ku complex on the 12RSS would only allow the incorporation of a 23RSS to accommodate the other binding site of RAG1/RAG2. The RAG1/RAG2 complex would then mediate the coupled cleavage in the synaptic complex, which supports the synaptic complex ordered assembly model by Jones and Gellert (2002).

4.3 Implication of the RAG-Ku interaction for the integration of NHEJ apparatus into V(D)J recombination

V(D)J recombination is conceptually divided into two steps including the recognition and cleavage of RSS by RAG1/RAG2 proteins and the joining step of the signal and coding ends by the NHEJ pathway. After cleavage, RAG1 and RAG2 complex were shown to associate with both coding ends and signal ends in a postcleavage complex in vivo and in vitro (Agrawal and Schatz, 1997; Hiom and Gellert, 1998; Jones and Gellert, 2001; Qiu et al., 2001; Yarnell Schultz et al., 2001; Huye et al., 2002; Tsai et al., 2002; Lee et al., 2004). Moreover, the recent study by Lee et al. (2004) has suggested that the RAG1/RAG2 complex was involved in channeling DNA DSBs to the NHEJ machinery. However, how the NHEJ pathway is integrated into the V(D)J recombination remains to be elucidated.

It is possible that Ku associates with RAG proteins even before RAGs recognize and cleave at RSS sequence. RAG1/RAG2 make contacts with the phosphate

backbone of the RSS only on one side of the DNA helix (Swanson and Desiderio, 1998), which could potentially allow Ku to gain access to the RSS sequence from the contralateral side of the helix. When the DNA double-stranded breaks occur, Ku would already be present on the DNA and could quickly recruit the other NHEJ factors to start the joining process. This quick link between the cleavage and joining process in V(D)J recombination via the interaction of RAG and Ku could facilitate the DNA damage response and promote cell survival.

Alternatively, the interaction of RAG and Ku may also help to remodel the postcleavage complex. For example, RAG1 and RAG2 proteins have been suggested to be the substrates of DNA-PKcs (Lin et al., 1999). The recruitment of DNA-PKcs to the complex of RAG/Ku/DNA might mediate the phosphorylation of the RAG proteins, promoting their degradation, which would dissemble the postcleavage complex and therefore promote the access of DNA ends to the other NHEJ factors.

4.4 Future experiments

In this study, the region in RAG1 spanning from amino acid residue 660 and 763 has been shown to be sufficient to interact with Ku. However, we still wonder whether this region is necessary to mediate Ku-RAG1 interaction. This could be done by in vitro protein binding assays with RAG1 constructs bearing internal deletion of amino acid residue 660 to 763 to be tested for binding to Ku from Jurkat T cell nuclear extracts. If the region is confirmed to be necessary for interaction with Ku, ZFB (aa728-760), the

predominant RAG2-interacting region in RAG1, will be deleted from this region and the mutant of RAG1 $_{660\text{-}727}$ would be tested for binding to Ku. The ultimate goal of these experiments will be to define a minimal region which would be further targeted by site-directed mutagenesis. RAG1 mutants no longer interacting with Ku will be employed for further study of Ku-RAG interaction in V(D)J recombination.

This study has shown that a DNA-binding defective mutant in the NBR can still interact with Ku. One important question to be addressed is if and how the interaction between RAG and Ku is involved in DNA cleavage process. One way to answer this question is through in vitro RSS cleavage assay in which a complex extrachromosomal recombination DNA substrate such as pJH200 that contains endogenous V, D or J gene segments is employed. To test whether the facilitation of DNA cleavage is through the Ku-RAG interaction, DNA-binding defective mutant RAG1 K405A/H406A/R407A which was found to be capable of interacting with Ku in this study will be employed in in vitro cleavage assay with the pJH200 DNA substrate. This result would demonstrate that the Ku-RAG interaction is capable of recruiting the RAG1/RAG2 complex to the RSS.

Earlier studies have shown that the 12/23 rule is not strictly enforced when the RAG1/RAG2 complex alone is employed for cleavage of extrachromosomal recombination DNA substrate in vitro, however, this rule can be reinforced by supplementing with cellular extracts (Sawchuk et al., 1997). In addition, a recent study has determined that Ku and DNA-PK were necessary to enforce the 12/23 rule in the

DNA cleavage step by RAG1/RAG2 proteins in this vitro system (Sawchuk et al., 2004). To investigate if Ku-RAG interaction is involved in the enforcement of 12/23 rule, in vitro cleavage assays are to be performed by purified RAG2 protein and wild type RAG1 or RAG1 Ku-interacting defective mutant with Jurkat T cell whole cell extracts on extrachromosomal recombination DNA substrate containing either 12/23 RSS, 12/12 RSS or 23/23RSS.

The region in RAG2 between amino acid residue 132 and 242 seems to be involved in the interaction with Ku, which maps with the 3rd and 4th repeat of the core RAG2. RAG1-interaction domain has been defined to coincides the sixth kelch motif in core RAG2, which suggests that individual of the other five kelch motifs may also be involved in interaction with other components of the V(D)J recombination (Aidinis et al., 2000). Further analysis of Ku-interacting region of RAG2 would involve expressing the 3rd and 4th repeat as a fusion protein and testing for interacting with Ku.

4.5 Summary and Conclusions

The results in this study present the first detailed analysis of Ku-interacting domain definition in both RAG1 and RAG2 proteins. Core regions of RAG1 and RAG2 are shown to physically interact with purified recombinant Ku in solution and the interaction between RAG and Ku are DNA-independent.

Both N-terminal and C-terminal deletion mutants of RAG1 protein were examined to interact with Ku using in vitro protein binding assay. There were several

findings in this study: (1) The region between amino acid residue 1 to 549 was not involved in the interaction with Ku, which suggested that the N-terminal non-core region and the NBR in the core region including the homeodomain-like motif were not involved in the interaction with Ku; (2) The region spanning from amino acid residue between 550 and 755 was defined to associate with Ku, which suggested that the central domain of RAG1 was involved in the interaction with Ku. The Ku-interacting domain between amino acid residue 660 and 763 was further confirmed to be sufficient to interact with Ku; (3) The region spanning from amino acid residue 792 to 1013 was found incapable to associate with Ku, which suggested that the coding flank binding region at the C-terminal of RAG1 was not involved in the interaction with Ku. In addition, a DNA-binding defective mutant in the NBR of RAG1 was shown to interact with Ku. Due to the secondary and tertiary structure of the RAG2 protein, the precise Ku-interacting region was not precisely defined. However, the 3rd and the 4th kelch motifs in the core region of RAG2 were found likely to be involved in the interaction with Ku. In conclusion, research into the detailed Ku-interacting domain mapping in both RAG1 and RAG2 proteins is of fundamental importance in order to understand the role of the RAG-Ku interaction V(D)J recombination.

References

Agrawal, A., Eastman, Q. M., and Schatz, D. G. (1998). Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. Nature *394*, 744-751.

Agrawal, A., and Schatz, D. G. (1997). RAG1 and RAG2 form a stable postcleavage synaptic complex with DNA containing signal ends in V(D)J recombination. Cell *89*, 43-53.

Aidinis, V., Dias, D. C., Gomez, C. A., Bhattacharyya, D., Spanopoulou, E., and Santagata, S. (2000). Definition of minimal domains of interaction within the recombination-activating genes 1 and 2 recombinase complex. J Immunol *164*, 5826-5832.

Akamastsu, Y., and Oettinger, M. A. (1998). Distinct Roles of RAG1 and RAG2 in Binding the V(D)J Recombination Signal Sequences. Mol Cell Biol *18*, 4670-4678.

Aravind, L., and Koonin, E. V. (2000). SAP - a putative DNA-binding motif involved in chromosomal organization. Trends Biochem Sci *25*, 112-114.

Aravind, L., and Koonin, E. V. (2001). Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. Genome Res *11*, 1365-1374.

Arbuckle, J. L., Fauss, L. A., Simpson, R., Ptaszek, L. M., and Rodgers, K. K. (2001). Identification of two topologically independent domains in RAG1 and their role in macromolecular interactions relevant to V(D)J recombination. J Biol Chem *276*, 37093-37101.

Bailey, S. M., Meyne, J., Chen, D. J., Kurimasa, A., Li, G. C., Lehnert, B. E., and Goodwin, E. H. (1999). DNA double-strand break repair proteins are required to cap the ends of mammalian chromosomes. Proc Natl Acad Sci U S A *96*, 14899-14904.

Bain, G., Maandag, E. C., Izon, D. J., Amsen, D., Kruisbeek, A. M., Weintraub, B. C., Krop, I., Schlissel, M. S., Feeney, A. J., van Roon, M., and et al. (1994). E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. Cell *79*, 885-892.

Bassing, C. H., Swat, W., and Alt, F. W. (2002). The mechanism and regulation of chromosomal V(D)J recombination. Cell *109 Suppl*, S45-55.

Baumann, P., and West, S. C. (1998). DNA end-joining catalyzed by human cell-free extracts. Proc Natl Acad Sci U S A *95*, 14066-14070.

Bellon, S. F., Rodgers, K. K., Schatz, D. G., Coleman, J. E., and Steitz, T. A. (1997). Crystal structure of the RAG1 dimerization domain reveals multiple zinc-binding motifs including a novel zinc binuclear cluster. Nat Struct Biol *4*, 586-591.

Bertinato, J., Schild-Poulter, C., and Hache, R. J. (2001). Nuclear localization of Ku antigen is promoted independently by basic motifs in the Ku70 and Ku80 subunits. J Cell Sci *114*, 89-99.

Bertocci, B., De Smet, A., Berek, C., Weill, J. C., and Reynaud, C. A. (2003). Immunoglobulin kappa light chain gene rearrangement is impaired in mice deficient for DNA polymerase mu. Immunity *19*, 203-211.

Blackwell, T. K., and Alt, F. W. (1989). Molecular characterization of the lymphoid V(D)J recombination activity. J Biol Chem *264*, 10327-10330.

Blier, P. R., Griffith, A. J., Craft, J., and Hardin, J. A. (1993). Binding of Ku protein to DNA. Measurement of affinity for ends and demonstration of binding to nicks. J Biol Chem *268*, 7594-7601.

Boulton, S. J., and Jackson, S. P. (1996). Identification of a Saccharomyces cerevisiae Ku80 homologue: roles in DNA double strand break rejoining and in telomeric maintenance. Nucleic Acids Res *24*, 4639-4648.

Bustin, M., and Reeves, R. (1996). High-mobility-group chromosomal proteins: architectural components that facilitate chromatin function. Prog Nucleic Acid Res Mol Biol *54*, 35-100.

Cai, Q. Q., Plet, A., Imbert, J., Lafage-Pochitaloff, M., Cerdan, C., and Blanchard, J. M. (1994). Chromosomal location and expression of the genes coding for Ku p70 and p80 in human cell lines and normal tissues. Cytogenet Cell Genet *65*, 221-227.

Callebaut, I., and Mornon, J. P. (1998). The V(D)J recombination activating protein RAG2 consists of a six-bladed propeller and a PHD fingerlike domain, as revealed by sequence analysis. Cell Mol Life Sci *54*, 880-891.

Camara-Clayette, V., Thomas, D., Rahuel, C., Barbey, R., Cartron, J. P., and Bertrand, O. (1999). The repressor which binds the -75 GATA motif of the GPB promoter contains

Ku70 as the DNA binding subunit. Nucleic Acids Res *27*, 1656-1663.

Cary, R. B., Chen, F., Shen, Z., and Chen, D. J. (1998). A central region of Ku80 mediates interaction with Ku70 in vivo. Nucleic Acids Res *26*, 974-979.

Ciubotaru, M., Ptaszek, L. M., Baker, G. A., Baker, S. N., Bright, F. V., and Schatz, D. G. (2003). RAG1-DNA binding in V(D)J recombination. Specificity and DNA-induced conformational changes revealed by fluorescence and CD spectroscopy. J Biol Chem *278*, 5584-5596.

Corcoran, A. E., Riddell, A., Krooshoop, D., and Venkitaraman, A. R. (1998). Impaired immunoglobulin gene rearrangement in mice lacking the IL-7 receptor. Nature *391*, 904-907.

Corneo, B., Benmerah, A., and Villartay, J. P. (2002). A short peptide at the C terminus is responsible for the nuclear localization of RAG2. Eur J Immunol *32*, 2068-2073.

Corneo, B., Moshous, D., Callebaut, I., de Chasseval, R., Fischer, A., and de Villartay, J. P. (2000). Three-dimensional clustering of human RAG2 gene mutations in severe combined immune deficiency. J Biol Chem *275*, 12672-12675.

Cortes, P., Weis-Garcia, F., Misulovin, Z., Nussenzweig, A., Lai, J. S., Li, G., Nussenzweig, M. C., and Baltimore, D. (1996). In vitro V(D)J recombination: signal joint formation. Proc Natl Acad Sci U S A *93*, 14008-14013.

Cortes, P., Ye, Z. S., and Baltimore, D. (1994). RAG-1 interacts with the repeated amino acid motif of the human homologue of the yeast protein SRP1. Proc Natl Acad Sci U S A *91*, 7633-7637.

Cuomo, C. A., Kirch, S. A., Gyuris, J., Brent, R., and Oettinger, M. A. (1994). Rch1, a protein that specifically interacts with the RAG-1 recombination-activating protein. Proc Natl Acad Sci U S A *91*, 6156-6160.

Cuomo, C. A., Mundy, C. L., and Oettinger, M. A. (1996). DNA sequence and structure requirements for cleavage of V(D)J recombination signal sequences. Mol Cell Biol *16*, 5683-5690.

Cuomo, C. A., and Oettinger, M. A. (1994). Analysis of regions of RAG-2 important for V(D)J recombination. Nucleic Acids Res *22*, 1810-1814.

De, P., and Rodgers, K. K. (2004). Putting the pieces together: identification and

characterization of structural domains in the V(D)J recombination protein RAG1. Immunol Rev *200*, 70-82.

DeFazio, L. G., Stansel, R. M., Griffith, J. D., and Chu, G. (2002). Synapsis of DNA ends by DNA-dependent protein kinase. Embo J *21*, 3192-3200.

Difilippantonio, M. J., McMahan, C. J., Eastman, Q. M., Spanopoulou, E., and Schatz, D. G. (1996). Rag-1 mediates signal sequence recognition and recruitment of Rag-1 in V(D)J recombination. Cell *87*, 253-262.

Difilippantonio, M. J., Zhu, J., Chen, H. T., Meffre, E., Nussenzweig, M. C., Max, E. E., Ried, T., and Nussenzweig, A. (2000). DNA repair protein Ku80 suppresses chromosomal aberrations and malignant transformation. Nature *404*, 510-514.

Doherty, A. J., Jackson, S. P., and Weller, G. R. (2001). Identification of bacterial homologues of the Ku DNA repair proteins. FEBS Lett *500*, 186-188.

Dynan, W. S., and Yoo, S. (1998). Interaction of Ku protein and DNA-dependent protein kinase catalytic subunit with nucleic acids. Nucleic Acids Res *26*, 1551-1559.

Eastman, Q. M., Leu, T. M., and Schatz, D. G. (1996). Initiation of V(D)J recombination in vitro obeying the 12/23 rule. Nature *380*, 85-88.

Eastman, Q. M., Villey, I. J., and Schatz, D. G. (1999). Detection of RAG protein-V(D)J recombination signal interactions near the site of DNA cleavage by UV cross-linking. Mol Cell Biol *19*, 3788-3797.

Elkin, S. K., Matthews, A. G., and Oettinger, M. A. (2003). The C-terminal portion of RAG2 protects against transposition in vitro. Embo J *22*, 1931-1938.

Errami, A., Smider, V., Rathmell, W. K., He, D. M., Hendrickson, E. A., Zdzienicka, M. Z., and Chu, G. (1996). Ku86 defines the genetic defect and restores X-ray resistance and V(D)J recombination to complementation group 5 hamster cell mutants. Mol Cell Biol *16*, 1519-1526.

Falzon, M., Fewell, J. W., and Kuff, E. L. (1993). EBP-80, a transcription factor closely resembling the human autoantigen Ku, recognizes single- to double-strand transitions in DNA. J Biol Chem *268*, 10546-10552.

Fugmann, S. D., Lee, A. I., Shockett, P. E., Villey, I. J., and Schatz, D. G. (2000a). The RAG proteins and V(D)J recombination: complexes, ends, and transposition. Annu Rev

Immunol *18*, 495-527.

Fugmann, S. D., Villey, I. J., Ptaszek, L. M., and Schatz, D. G. (2000b). Identification of two catalytic residues in RAG1 that define a single active site within the RAG1/RAG2 protein complex. Mol Cell *5*, 97-107.

Gellert, M. (2002). V(D)J recombination: RAG proteins, repair factors, and regulation. Annu Rev Biochem *71*, 101-132.

Giffin, W., Kwast-Welfeld, J., Rodda, D. J., Prefontaine, G. G., Traykova-Andonova, M., Zhang, Y., Weigel, N. L., Lefebvre, Y. A., and Hache, R. J. (1997). Sequence-specific DNA binding and transcription factor phosphorylation by Ku Autoantigen/DNA-dependent protein kinase. Phosphorylation of Ser-527 of the rat glucocorticoid receptor. J Biol Chem *272*, 5647-5658.

Giffin, W., Torrance, H., Rodda, D. J., Prefontaine, G. G., Pope, L., and Hache, R. J. (1996). Sequence-specific DNA binding by Ku autoantigen and its effects on transcription. Nature *380*, 265-268.

Godderz, L. J., Rahman, N. S., Risinger, G. M., Arbuckle, J. L., and Rodgers, K. K. (2003). Self-association and conformational properties of RAG1: implications for formation of the V(D)J recombinase. Nucleic Acids Res *31*, 2014-2023.

Gottlieb, T. M., and Jackson, S. P. (1993). The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen. Cell *72*, 131-142.

Grawunder, U., Leu, T. M., Schatz, D. G., Werner, A., Rolink, A. G., Melchers, F., and Winkler, T. H. (1995). Down-regulation of RAG1 and RAG2 gene expression in preB cells after functional immunoglobulin heavy chain rearrangement. Immunity *3*, 601-608.

Grawunder, U., and Lieber, M. R. (1997). A complex of RAG-1 and RAG-2 proteins persists on DNA after single-strand cleavage at V(D)J recombination signal sequences. Nucleic Acids Res *25*, 1375-1382.

Grawunder, U., Schatz, D. G., Leu, T. M., Rolink, A., and Melchers, F. (1996). The half-life of RAG-1 protein in precursor B cells is increased in the absence of RAG-2 expression. J Exp Med *183*, 1731-1737.

Hammarsten, O., and Chu, G. (1998). DNA-dependent protein kinase: DNA binding and activation in the absence of Ku. Proc Natl Acad Sci U S A *95*, 525-530.

Harris, R., Esposito, D., Sankar, A., Maman, J. D., Hinks, J. A., Pearl, L. H., and Driscoll, P. C. (2004). The 3D solution structure of the C-terminal region of Ku86 (Ku86CTR). J Mol Biol *335*, 573-582.

Hesslein, D. G., Pflugh, D. L., Chowdhury, D., Bothwell, A. L., Sen, R., and Schatz, D. G. (2003). Pax5 is required for recombination of transcribed, acetylated, 5' IgH V gene segments. Genes Dev *17*, 37-42.

Hiom, K., and Gellert, M. (1997). A stable RAG1-RAG2-DNA complex that is active in V(D)J cleavage. Cell *88*, 65-72.

Hiom, K., and Gellert, M. (1998). Assembly of a 12/23 paired signal complex: a critical control point in V(D)J recombination. Mol Cell *1*, 1011-1019.

Hiom, K., Melek, M., and Gellert, M. (1998). DNA transposition by the RAG1 and RAG2 proteins:a possible source of oncogenic translocations. Cell *94*, 463-470.

Hoekstra, M. F. (1997). Responses to DNA damage and regulation of cell cycle checkpoints by the ATM protein kinase family. Curr Opin Genet Dev *7*, 170-175.

Huye, L. E., Purugganan, M. M., Jiang, M. M., and Roth, D. B. (2002). Mutational analysis of all conserved basic amino acids in RAG-1 reveals catalytic, step arrest, and joining-deficient mutants in the V(D)J recombinase. Mol Cell Biol *22*, 3460-3473.

Ichihara, Y., Hirai, M., and Kurosawa, Y. (1992). Sequence and chromosome assignment to 11p13-p12 of human RAG genes. Immunol Lett *33*, 277-284.

Jiang, H., Ross, A. E., and Desiderio, S. (2004). Cell cycle-dependent accumulation in vivo of transposition-competent complexes between recombination signal ends and full-length RAG proteins. J Biol Chem *279*, 8478-8486.

Jin, S., and Weaver, D. T. (1997). Double-strand break repair by Ku70 requires heterodimerization with Ku80 and DNA binding functions. Embo J *16*, 6874-6885.

Jones, J. M., and Gellert, M. (2001). Intermediates in V(D)J recombination: a stable RAG1/2 complex sequesters cleaved RSS ends. Proc Natl Acad Sci U S A *98*, 12926-12931.

Jones, J. M., and Gellert, M. (2002). Ordered assembly of the V(D)J synaptic complex ensures accurate recombination. Embo J *21*, 4162-4171.

Jones, J. M., and Gellert, M. (2003). Autoubiquitylation of the V(D)J recombinase protein RAG1. Proc Natl Acad Sci U S A *100*, 15446-15451.

Jung, D., and Alt, F. W. (2004). Unraveling V(D)J recombination:insights into gene regulation. Cell *116*, 299-311.

Katso, R., Okkenhaug, K., Ahmadi, K., White, S., Timms, J., and Waterfield, M. D. (2001). Cellular function of phosphoinositide 3-kinases: implications for development, homeostasis, and cancer. Annu Rev Cell Dev Biol *17*, 615-675.

Kim, D. R., Dai, Y., Mundy, C. L., Yang, W., and Oettinger, M. A. (1999). Mutations of acidic residues in RAG1 define the active site of the V(D)J recombinase. Genes Dev *13*, 3070-3080.

Kim, D. R., and Oettinger, M. A. (1998). Functional analysis of coordinated cleavage in V(D)J recombination. Mol Cell Biol *18*, 4679-4688.

Kirch, S. A., Rathbun, G. A., and Oettinger, M. A. (1998). Dual role of RAG2 in V(D)J recombination: catalysis and regulation of ordered Ig gene assembly. Embo J *17*, 4881-4886.

Kirch, S. A., Sudarsanam, P., and Oettinger, M. A. (1996). Regions of RAG1 protein critical for V(D)J recombination. Eur J Immunol *26*, 886-891.

Koike, M., Awaji, T., Kataoka, M., Tsujimoto, G., Kartasova, T., Koike, A., and Shiomi, T. (1999a). Differential subcellular localization of DNA-dependent protein kinase components Ku and DNA-PKcs during mitosis. J Cell Sci *112 (Pt 22)*, 4031-4039.

Koike, M., Ikuta, T., Miyasaka, T., and Shiomi, T. (1999b). Ku80 can translocate to the nucleus independent of the translocation of Ku70 using its own nuclear localization signal. Oncogene *18*, 7495-7505.

Koike, M., Ikuta, T., Miyasaka, T., and Shiomi, T. (1999c). The nuclear localization signal of the human Ku70 is a variant bipartite type recognized by the two components of nuclear pore-targeting complex. Exp Cell Res *250*, 401-413.

Koike, M., Matsuda, Y., Mimori, T., Harada, Y. N., Shiomi, N., and Shiomi, T. (1996). Chromosomal localization of the mouse and rat DNA double-strand break repair genes Ku p70 and Ku p80/XRCC5 and their mRNA expression in various mouse tissues. Genomics *38*, 38-44.

Koike, M., Shiomi, T., and Koike, A. (2000). Ku70 can translocate to the nucleus independent of Ku80 translocation and DNA-PK autophosphorylation. Biochem Biophys Res Commun *276*, 1105-1111.

Koike, M., Shiomi, T., and Koike, A. (2001). Dimerization and nuclear localization of ku proteins. J Biol Chem *276*, 11167-11173.

Krangel, M. S. (2003). Gene segment selection in V(D)J recombination:accessibility and beyond. Nat Immunol *4*, 624-630.

Kwon, J., Morshead, K. B., Guyon, J. R., Kingston, R. E., and Oettinger, M. A. (2000). Histone acetylation and hSWI/SNF remodeling act in concert to stimulate V(D)J cleavage of nucleosomal DNA. Mol Cell *6*, 1037-1048.

Lafaille, J. J., DeCloux, A., Bonneville, M., Takagaki, Y., and Tonegawa, S. (1989). Junctional sequences of T cell receptor gamma delta genes: implications for gamma delta T cell lineages and for a novel intermediate of V-(D)-J joining. Cell *59*, 859-870.

Landree, M. A., Wibbenmeyer, J. A., and Roth, D. B. (1999). Mutational analysis of RAG1 and RAG2 identifies three catalytic amino acids in RAG1 critical for both cleavage steps of V(D)J recombination. Genes Dev *13*, 3059-3069.

Lee, G. S., Neiditch, M. B., Salus, S. S., and Roth, D. B. (2004). RAG proteins shepherd double-strand breaks to a specific pathway, suppressing error-prone repair, but RAG nicking initiates homologous recombination. Cell *117*, 171-184.

Leu, T. M., Eastman, Q. M., and Schatz, D. G. (1997). Coding joint formation in a cell-free V(D)J recombination system. Immunity *7*, 303-314.

Lewis, S. M. (1994). The mechanism of V(D)J joining: lessons from molecular, immunological, and comparative analyses. Adv Immunol *56*, 27-150.

Li, G., Nelsen, C., and Hendrickson, E. A. (2002). Ku86 is essential in human somatic cells. Proc Natl Acad Sci U S A *99*, 832-837.

Lieber, M. R., Ma, Y., Pannicke, U., and Schwarz, K. (2003). Mechanism and regulation of human non-homologous DNA end-joining. Nat Rev Mol Cell Biol *4*, 712-720.

Lin, H., and Grosschedl, R. (1995). Failure of B-cell differentiation in mice lacking the transcription factor EBF. Nature *376*, 263-267.

Lin, J. M., Landree, M. A., and Roth, D. B. (1999). V(D)J recombination catalyzed by mutant RAG proteins lacking consensus DNA-PK phosphorylation sites. Mol Immunol *36*, 1263-1269.

Lin, W. C., and Desiderio, S. (1993). Regulation of V(D)J recombination activator protein RAG-2 by phosphorylation. Science *260*, 953-959.

Lin, W. C., and Desiderio, S. (1994). Cell cycle regulation of V(D)J recombination-activating protein RAG-2. Proc Natl Acad Sci U S A *91*, 2733-2737.

Lin, W. C., and Desiderio, S. (1995). V(D)J recombination and the cell cycle. Immunol Today *16*, 279-289.

Ma, Y., Pannicke, U., Schwarz, K., and Lieber, M. R. (2002). Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. Cell *108*, 781-794.

Mahajan, K. N., Gangi-Peterson, L., Sorscher, D. H., Wang, J., Gathy, K. N., Mahajan, N. P., Reeves, W. H., and Mitchell, B. S. (1999). Association of terminal deoxynucleotidyl transferase with Ku. Proc Natl Acad Sci U S A *96*, 13926-13931.

Mahajan, K. N., McElhinny, S. A. N., Michell, B. S., and Ramsden, D. A. (2002). Association of DNA polymerase mu (pol mu) with Ku and ligase IV: role for pol mu in end-joining double-strand break repair. Mol Cell Biol *22*, 5194-5202.

McBlane, J. F., van Gent, D. C., Ramsden, D. A., Romeo, C., Cuomo, C. A., Gellert, M., and Oettinger, M. A. (1995). Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. Cell *83*, 387-395.

McCormack, W. T., Tjoelker, L. W., Carlson, L. M., Petryniak, B., Barth, C. F., Humphries, E. H., and Thompson, C. B. (1989). Chicken IgL gene rearrangement involves deletion of a circular episome and addition of single nonrandom nucleotides to both coding segments. Cell *56*, 785-791.

McElhinny, S. A. N., and Ramsden, D. A. (2003). Polymerase mu is a DNA-directed DNA/RNA polymerase. Mol Cell Biol *23*, 2309-2315.

Mimori, T., Akizuki, M., Yamagata, H., Inada, S., Yoshida, S., and Homma, M. (1981). Characterization of a high molecular weight acidic nuclear protein recognized by autoantibodies in sera from patients with polymyositis-scleroderma overlap. J Clin

Invest *68*, 611-620.

Mimori, T., and Hardin, J. A. (1986). Mechanism of interaction between Ku protein and DNA. J Biol Chem *261*, 10375-10379.

Mimori, T., Hardin, J. A., and Steitz, J. A. (1986). Characterization of the DNA-binding protein antigen Ku recognized by autoantibodies from patients with rheumatic disorders. J Biol Chem *261*, 2274-2278.

Mo, X., Bailin, T., and Sadofsky, M. J. (1999). RAG1 and RAG2 cooperate in specific binding to the recombination signal sequence in vitro. J Biol Chem *274*, 7025-7031.

Mo, X., Bailin, T., and Sadofsky, M. J. (2001). A C-terminal region of RAG1 contacts the coding DNA during V(D)J recombination. Mol Cell Biol *21*, 2038-2047.

Mombaerts, P., Iacomini, J., Johnson, R. S., Herrup, K., Tonegawa, S., and Papaioannou, V. E. (1992). RAG-1-deficient mice have no mature B and T lymphocytes. Cell *68*, 869-877.

Monroe, R. J., Seidl, K. J., Gaertner, F., Han, S., Chen, F., Sekiguchi, J., Wang, J., Ferrini, R., Davidson, L., Kelsoe, G., and Alt, F. W. (1999). RAG2:GFP knockin mice reveal novel aspects of RAG2 expression in primary and peripheral lymphoid tissues. Immunity *11*, 201-212.

Moshous, D., Callebaut, I., de Chasseval, R., Corneo, B., Cavazzana-Calvo, M., Le Deist, F., Tezcan, I., Sanal, O., Bertrand, Y., Philippe, N., *et al.* (2001). Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human severe combined immune deficiency. Cell *105*, 177-186.

Mundy, C. L., Patenge, N., Matthews, A. G., and Oettinger, M. A. (2002). Assembly of the RAG1/RAG2 synaptic complex. Mol Cell Biol *22*, 69-77.

Myung, K., He, D. M., Lee, S. E., and Hendrickson, E. A. (1997). KARP-1: a novel leucine zipper protein expressed from the Ku86 autoantigen locus is implicated in the control of DNA-dependent protein kinase activity. Embo J *16*, 3172-3184.

Neiditch, N. B., Lee, G. S., Huye, L. E., Brandt, V. L., and Roth, D. B. (2002). The V(D)J recombinase efficiently cleaves and transposes signal joints. Mol Cell *9*, 871-878.

Noordzij, J. G., Verkaik, N. S., Hartwig, N. G., de Groot, R., van Gent, D. C., and van Dongen, J. J. (2000). N-terminal truncated human RAG1 proteins can direct T-cell

receptor but not immunoglobulin gene rearrangements. Blood *96*, 203-209.

Oettinger, M. A., Schatz, D. G., Gorka, C., and Baltimore, D. (1990). RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science *248*, 1517-1523.

Oettinger, M. A., Stanger, B., Schatz, D. G., Glaser, T., Call, K., Housman, D., and Baltimore, D. (1992). The recombination activating genes, RAG 1 and RAG 2, are on chromosome 11p in humans and chromosome 2p in mice. Immunogenetics *35*, 97-101.

Osipovich, O., Durum, S. K., and Muegge, K. (1997). Defining the minimal domain of Ku80 for interaction with Ku70. J Biol Chem *272*, 27259-27265.

Paesen, G. C., Zanotto, P. M., and Nuttall, P. A. (1996). A tick homologue of the human DNA helicase II 70-kDa subunit. Biochim Biophys Acta *1305*, 120-124.

Park, S. J., Ciccone, S. L., Freie, B., Kurimasa, A., Chen, D. J., Li, G. C., Clapp, D. W., and Lee, S. H. (2004). A positive role for the Ku complex in DNA replication following strand break damage in mammals. J Biol Chem *279*, 6046-6055.

Peak, M. M., Arbuckle, J. L., and Rodgers, K. K. (2003). The central domain of core RAG1 preferentially recognizes single-stranded recombination signal sequence heptamer. J Biol Chem *278*, 18235-18240.

Purugganan, M. M., Shah, S., Kearney, J. F., and Roth, D. B. (2001). Ku80 is required for addition of N nucleotides to V(D)J recombination junctions by terminal deoxynucleotidyl transferase. Nucleic Acids Res *29*, 1638-1646.

Qiu, J. X., Kale, S. B., Yarnell Schultz, H., and Roth, D. B. (2001). Separation-of-function mutants reveal critical roles for RAG2 in both the cleavage and joining steps of V(D)J recombination. Mol Cell *7*, 77-87.

Ramsden, D. A., McBlane, J. F., van Gent, D. C., and Gellert, M. (1996). Distinct DNA sequence and structure requirements for the two steps of V(D)J recombination signal cleavage. EMBO J *15*, 3197-3206.

Ramsden, D. A., Paull, T. T., and Gellert, M. (1997). Cell-free V(D)J recombination. Nature *388*, 488-491.

Rodgers, K. K., Bu, Z., Fleming, K. G., Schatz, D. G., Engelman, D. M., and Coleman, J. E. (1996). A zinc-binding domain involved in the dimerization of RAG1. J Mol Biol *260*,

70-84.

Rodgers, K. K., Villey, I. J., Ptaszek, L., Corbett, E., Schatz, D. G., and Coleman, J. E. (1999). A dimer of the lymphoid protein RAG1 recognizes the recombination signal sequence and the complex stably incorporates the high mobility group protein HMG2. Nucleic Acids Res *27*, 2938-2946.

Roman, C. A., Cherry, S. R., and Baltimore, D. (1997). Complementation of V(D)J recombination deficiency in RAG-1(-/-) B cells reveals a requirement for novel elements in the N-terminus of RAG-1. Immunity *7*, 13-24.

Romanow, W. J., Langerak, A. W., Goebel, P., Wolvers-Tettero, I. L., van Dongen, J. J., Feeney, A. J., and Murre, C. (2000). E2A and EBF act in synergy with the V(D)J recombinase to generate a diverse immunoglobulin repertoire in nonlymphoid cells. Mol Cell *5*, 343-353.

Rooney, S., Sekiguchi, J., Zhu, C., Cheng, H. L., Manis, J., Whitlow, S., DeVido, J., Foy, D., Chaudhuri, J., Lombard, D., and Alt, F. W. (2002). Leaky Scid phenotype associated with defective V(D)J coding end processing in Artemis-deficient mice. Mol Cell *10*, 1379-1390.

Rouse, J., and Jackson, S. P. (2002). Interfaces between the detection, signaling, and repair of DNA damage. Science *297*, 547-551.

Ruiz, M. T., Matheos, D., Price, G. B., and Zannis-Hadjopoulos, M. (1999). OBA/Ku86: DNA binding specificity and involvement in mammalian DNA replication. Mol Biol Cell *10*, 567-580.

Sadofsky, M. J. (2001). The RAG proteins in V(D)J recombination: more than just a nuclease. Nucleic Acids Res *29*, 1399-1409.

Sadofsky, M. J., Hesse, J. E., and Gellert, M. (1994). Definition of a core region of RAG-2 that is functional in V(D)J recombination. Nucleic Acids Res *22*, 1805-1809.

Sadofsky, M. J., Hesse, J. E., McBlane, J. F., and Gellert, M. (1993). Expression and V(D)J recombination activity of mutated RAG-1 proteins. Nucleic Acids Res *21*, 5644-5650.

Sandor, Z., Calicchio, M. L., Sargent, R. G., Roth, D. B., and Wilson, J. H. (2004). Distinct requirements for Ku in N nucleotide addition at V(D)J- and non-V(D)J-generated double-strand breaks. Nucleic Acids Res *32*, 1866-1873.

Santagata, S., Gomez, C. A., Sobacchi, C., Bozzi, F., Abinun, M., Pasic, S., Cortes, P., Vezzoni, P., and Villa, A. (2000). N-terminal RAG1 frameshift mutations in Omenn's syndrome: internal methionine usage leads to partial V(D)J recombination activity and reveals a fundamental role in vivo for the N-terminal domains. Proc Natl Acad Sci U S A *97*, 14572-14577.

Sawchuk, D. J., Mansilla-Soto, J., Alarcon, C., Singha, N. C., Langen, H., Bianchi, M. E., Lees-Miller, S. P., Nussenzweig, M. C., and Cortes, P. (2004). Ku70/Ku80 and DNA-dependent protein kinase catalytic subunit modulate RAG-mediated cleavage: implications for the enforcement of the 12/23 rule. J Biol Chem *279*, 29821-29831.

Schatz, D. G., Oettinger, M. A., and Baltimore, D. (1989). The V(D)J recombination activating gene, RAG-1. Cell *59*, 1035-1048.

Schild-Poulter, C., Hache, R. J., and Soubeyrand, S. (2004). Ku antigen: a versitile DNA binding protein with multiple cellular functions. In "Recent Research Developments in Dynamical Genetics" Research Signpost, Transworld Research Network Edtors VParisi, VDe Fonzo and F Aluffi-Pentini, 257-284.

Schild-Poulter, C., Matheos, D., Novac, O., Cui, B., Giffin, W., Ruiz, M. T., Price, G. B., Zannis-Hadjopoulos, M., and Hache, R. J. (2003). Differential DNA binding of Ku antigen determines its involvement in DNA replication. DNA Cell Biol *22*, 65-78.

Schlissel, M., Constantinescu, A., Morrow, T., Baxter, M., and Peng, A. (1993). Double-strand signal sequence breaks in V(D)J recombination are blunt, 5'-phosphorylated, RAG-dependent, and cell cycle regulated. Genes Dev *7*, 2520-2532.

Schwarz, K., Gauss, G. H., Ludwig, L., Pannicke, U., Li, Z., Lindner, D., Friedrich, W., Seger, R. A., Hansen-Hagge, T. E., Desiderio, S., *et al.* (1996). RAG mutations in human B cell-negative SCID. Science *274*, 97-99.

Sekiguchi, J. A., Whitlow, S., and Alt, F. W. (2001). Increased accumulation of hybrid V(D)J joins in cells expressing truncated versus full-length RAGs. Mol Cell *8*, 1383-1390.

Shakibai, N., Kumar, V., and Eisenberg, S. (1996). The Ku-like protein from Saccharomyces cerevisiae is required in vitro for the assembly of a stable multiprotein complex at a eukaryotic origin of replication. Proc Natl Acad Sci U S A *93*, 11569-11574.

Sherrington, P. D., Forster, A., Seawright, A., van Heyningen, V., and Rabbitts, T. H. (1992). Human RAG2, like RAG1, is on chromosome 11 band p13 and therefore not linked to ataxia telangiectasia complementation groups. Genes Chromosomes Cancer *5*, 404-406.

Shinkai, Y., Rathbun, G., Lam, K. P., Oltz, E. M., Stewart, V., Mendelsohn, M., Charron, J., Datta, M., Young, F., Stall, A. M., and et al. (1992). RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. Cell *68*, 855-867.

Simon, M., Zieg, J., Silverman, M., Mandel, G., and Doolittle, R. (1980). Phase variation: evolution of a controlling element. Science *209*, 1370-1374.

Smider, V., Rathmell, W. K., Lieber, M. R., and Chu, G. (1994). Restoration of X-ray resistance and V(D)J recombination in mutant cells by Ku cDNA. Science *266*, 288-291.

Spanopoulou, E., Cortes, P., Shih, C., Huang, C. M., Silver, D. P., Svec, P., and Baltimore, D. (1995). Localization, interaction, and RNA binding properties of the V(D)J recombination-activating proteins RAG1 and RAG2. Immunity *3*, 715-726.

Steen, S. B., Gomelsky, L., and Roth, D. B. (1996). Double-strand breaks, DNA hairpins, and the mechanism of V(D)J recombination. Curr Top Microbiol Immunol *217*, 66-77.

Swanson, P., Volkmer, D., and Wang, L. (2004). Full-length RAG2, and not full-length RAG-1, specifically suppresses RAG-mediated transposition, butnot hybrid joint formation or disintegration. J Biol Chem *279*, 4034-4044.

Swanson, P. C., and Desiderio, S. (1999). RAG-2 promotes heptamer occupancy by RAG-1 in the assembly of a V(D)J initiation complex. Mol Cell Biol *19*, 3674-3683.

Tamura, K., Adachi, Y., Chiba, K., Oguchi, K., and Takahashi, H. (2002). Identification of Ku70 and Ku80 homologues in Arabidopsis thaliana: evidence for a role in the repair of DNA double-strand breaks. Plant J *29*, 771-781.

Tonegawa, S. (1983). Somatic generation of antibody diversity. Nature *302*, 575-581.

Tsai, C. L., Drejer, A. H., and Schatz, D. G. (2002). Evidence of a critical architectural function for the RAG proteins in end processing, protection, and joining in V(D)J recombination. Genes Dev *16*, 1934-1949.

Tsai, C. L., and Schatz, D. G. (2003). Regulation of RAG1/RAG2-mediated transposition

by GTP and the C-terminal region of RAG2. Embo J *22*, 1922-1930.

Tuteja, R., and Tuteja, N. (2000). Ku autoantigen: a multifunctional DNA-binding protein. Crit Rev Biochem Mol Biol *35*, 1-33.

van Gent, D. C., Hiom, K., Paull, T. T., and Gellert, M. (1997). Stimulation of V(D)J cleavage by high mobility group proteins. Embo J *16*, 2665-2670.

van Gent, D. C., Ramsden, D. A., and Gellert, M. (1996). The RAG1 and RAG2 proteins establish the 12/23 rule in V(D)J recombination. Cell *85*, 107-113.

Villa, A., Santagata, S., Bozzi, F., Giliani, S., Frattini, A., Imberti, L., Gatta, L. B., Ocs, H. D., Schwarz, K., Notarangelo, L. D., *et al.* (1998). Partial V(D)J recmbination activity leads to Omenn syndrome. Cell *93*, 885-896.

Villa, A., Santagata, S., Bozzi, F., Imberti, L., and Notarangelo, L. D. (1999). Omenn syndrome: a disorder of Rag1 and Rag2 genes. J Clin Immunol *19*, 87-97.

Walker, J. R., Corpina, R. A., and Goldberg, J. (2001). Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. Nature *412*, 607-614.

Wang, J., Dong, X., Myung, K., Hendrickson, E. A., and Reeves, W. H. (1998a). Identification of two domains of the p70 Ku protein mediating dimerization with p80 and DNA binding. J Biol Chem *273*, 842-848.

Wang, J., Dong, X., and Reeves, W. H. (1998b). A model for Ku heterodimer assembly and interaction with DNA. Implications for the function of Ku antigen. J Biol Chem *273*, 31068-31074.

Weis-Garcia, F., Besmer, E., Sawchuk, D. J., Yu, W., Hu, Y., Cassard, S., Nussenzweig, M. C., and Cortes, P. (1997). V(D)J recombination: in vitro coding joint formation. Mol Cell Biol *17*, 6379-6385.

West, R. B., and Lieber, M. R. (1998). The RAG-HMG1 complex enforces the 12/23 rule of V(D)J recombination specifically at the double-hairpin formation step. Mol Cell Biol *18*, 6408-6415.

Weterings, E., Verkaik, N. S., Bruggenwirth, H. T., Hoeijmakers, J. H., and van Gent, D. C. (2003). The role of DNA dependent protein kinase in synapsis of DNA ends. Nucleic Acids Res *31*, 7238-7246.

Wu, X., and Lieber, M. R. (1996). Protein-protein and protein-DNA interaction regions within the DNA end-binding protein Ku70-Ku86. Mol Cell Biol *16*, 5186-5193.

Yagura, T., and Sumi, K. (1999). Molecular cloning and sequencing of cDNAs encoding homologues of human Ku70 and Ku80 autoantigen from Xenopus and their expression in various Xenopus tissues. Biochim Biophys Acta *1445*, 160-164.

Yaneva, M., Kowalewski, T., and Lieber, M. R. (1997). Interaction of DNA-dependent protein kinase with DNA and with Ku: biochemical and atomic-force microscopy studies. Embo J *16*, 5098-5112.

Yarnell Schultz, H., Landree, M. A., Qiu, J. X., Kale, S. B., and Roth, D. B. (2001). Joining-deficient RAG1 mutants block V(D)J recombination in vivo and hairpin opening in vitro. Mol Cell *7*, 65-75.

Yu, W., Misulovin, Z., Suh, H., Hardy, R. R., Jankovic, M., Yannoutsos, N., and Nussenzweig, M. C. (1999). Coordinate regulation of RAG1 and RAG2 by cell type-specific DNA elements 5' of RAG2. Science *285*, 1080-1084.

Yurchenko, V., Xue, Z., and Sadofsky, M. (2003). The RAG1 N-terminal domain is an E3 ubiquitin ligase. Genes Dev *17*, 581-585.

Zhang, Z., Zhu, L., Lin, D., Chen, F., Chen, D. J., and Chen, Y. (2001). The three-dimensional structure of the C-terminal DNA-binding domain of human Ku70. J Biol Chem *276*, 38231-38236.

Zhuang, Y., Soriano, P., and Weintraub, H. (1994). The helix-loop-helix gene E2A is required for B cell formation. Cell *79*, 875-884.

Appendix I

List of primers used in this study

T7 primer	5'-TAATACGACTCACTATAGGG-3'
T7-R1Fw400	5'-TAATACGACTCACTATAGGGCATGATGACGAGAAGGGCGCAGAAACATCGG-3'
T7-R1Fw499	5'-TAATACGACTCACTATAGGGCATGATGCCTTTGCATGCTCTTCGGAATGCC-3'
T7-R1Fw592	5'-TAATACGACTCACTATAGGGCATGATGACAGTGGTGGTAAAGGAGTCTTGC-3'
T7-R1Fw660	5'-TAATACGACTCACTATAGGGCATGATGGCAGATGAGTCTGACCATGAGACC-3'
R1Rev549	5'-TGGGTACTCATCCACAGAGGAGGC-3'
R1Rev554	5'-AGAGTCGTAGCGGAACCTCTTCGC-3'
R1Rev596	5'-CTTTACCACCACTGTGAAGGGACC-3'
R1Rev677	5'-GGCAATGAGGGGGCTTAGAATAGC-3'
R1Rev1056	5'-GAAGACCTCACTGCAACCCAAAGG-3'
T7-R2Fw132	5'-TAATACGACTCACTATAGGGGCATGATGGATGTCCCTGAACCCAGATACGGC-3'
T7-R2Fw161	5'-TAATACGACTCACTATAGGGCATGATGTACATGCCTTCTACCCAGAGAACC-3'
T7-R2Fw173	5'-TAATACGACTCACTATAGGGCATGATGAATAGTGTAGCTGACTGCCTACCC-3'
T7-R2Fw389	5'-TAATACGACTCACTATAGGGCATGATGGAAGCAACCAGTTTTGATGGTGACG-3'
R2Rev532	5'-GCAATATACCTGAGTCTGAGGGGC-3'
XbaI-R1Fw499	5'-CGTCTAGACCTTTGCATGCTCTTCGGAATGCC-3'
XbaI-R1Fw573	5'-CGTCTAGAGACATCTTGGAAGGCATGAGATCC-3'
XbaI-R1Fw660	5'-CGTCTAGAGCAGATGAGTCTGACCATGAGACC-3'
XbaI-R1Fw792	5'-CGTCTAGAGATGCGCTTCACTGTGACATTGGC-3'
R1Rev763stop	5'-ATTGGACCGCCAGACCTCATAGCG-3'
R1Rev1015stop	5'-GGTAAACCCAGAGCTTTTTAACGC-3'
R1FwK405A/ H406A/R407A	5'-CGAGAAGGGCGCAGGCAGCTGCTCTTCGGGAGCTCAAGATTC-3'
R1RevK405A/ H406A/R407A	5'-GAATCTTGAGCTCCCGAAGAGCAGCTGCCTGCGCCCTTCTCG-3'

Appendix II

Amino acid sequence of human RAG1 protein

1	MAASFPPTLGLSSAPDEIQHPHIKFSEWKFKLFRVRSFEKTPEEAQKEKK
51	DSFEGKPSLEQSPAVLDKADGQKPVPTQPLLKAHPKFSKKFHDNEKARGK
101	AIHQANLRHLCRICGNSFRADEHNRRYPVHGPVDGKTLGLLRKKEKRATS
151	WPDLIAKVFRIDVKADVDSIHPTEFCHNCWSIMHRKFSSAPCEVYFPRNV
201	TMEWHPHTPSCDICNTARRGLKRKSLQPNLQLSKKLKTVLDQARQARQRK
251	RRAQARISSKDVMKKIANCSKIHLSTKLLAVDFPEHFVKSISCQICEHIL
301	ADPVETNCKHVFCRVCILRCLKVMGSYCPSCRYPCFPTDLESPVKSFLSV
351	LNSLMVKCPAKECNEEVSLEKYNHHISSHKESKEIFVHINKGGRPRQHLL
401	SLTRRAQKHRLRELKLQVKAFADKEEGGDVKSVCMTLFLLALRARNEHRQ
451	ADELEAIMQGKGSGLQPAVCLAIRVNTFLSCSQYHKMYRTVKAITGRQIF
501	QPLHALRNAEKVLLPGYHHFEWQPPLKNVSSSTDVGIIDGLSGLSSSVDD
551	YPVDTIAKRFRYDSALVSALMDMEEDILEGMRSQDLDDYLNGPFTVVVKE
601	SCDGMGDVSEKHGSGPVVPEKAVRFSFTIMKITIAHSSQNVKVFEEAKPN
651	SELCCKPLCLMLADESDHETLTAILSPLIAEREAMKSSELMLELGGILRT
701	FKFIFRGTGYDEKLVREVEGLEASGSVYICTLCDATRLEASQNLVFHSIT
751	RSHAENLERYEVWRSNPYHESVEELRDRVKGVSAKPFIETVPSIDALHCD
801	IGNAAEFYKIFQLEIGEVYKNPNASKEERKRWQATLDKHLRKKMNLKPIM
851	RMNGNFARKLMTKETVDAVCELIPSEERHEALRELMDLYLKMKPVWRSSC
901	PAKECPESLCQYSFNSQRFAELLSTKFKYRYEGKITNYFHKTLAHVPEII
951	ERDGSIGAWASEGNESGNKLFRRFRKMNARQSKCYEMEDVLKHHWLYTSK
1001	YLQKFMNAHNALKTSGFTMNPQASLGDPLGIEDSLESQDSMEF

* The GeneBank accession number for the amino acid sequence of human RAG1 protein is AAA60248.

Appendix III

Amino acid sequence of human RAG2 protein

1	MSLQMVTVSNNIALIQPGFSIMNFDGQVFFFGQKGWPKRSCPTGVFHLDV
51	KHNHVKLKPTIFSKDSCYLPPLRYPATCTFKGSLESEKHQYIIHGGKTPN
101	NEVSDKIYVMSIVCKNNKKVTFRCTEKDLVGDVPEARYGHSINVVYSRGK
151	SMGVLFGGRSYMPSTHRTTEKWNSVADCLPCVFLVDFEFGCATSYILPEL
201	QDGLSFHVSIAKNDTIYILGGHSLANNIRPANLYRIRVDLPLGSPAVNCT
251	VLPGGISVSSAILTQTNNDEFVIVGGYQLENQKRMICNIISLEDNKIEIR
301	EMETPDWTPDIKHSKIWFGSNMGNGTVFLGIPGDNKQVVSEGFYFYMLKC
351	AEDDTNEEQTTFTNSQTSTEDPGDSTPFEDSEEFCFSAEANSFDGDDEFD
401	TYNEDDEEDESETGYWITCCPTCDVDINTWVPFYSTELNKPAMIYCSHGD
451	GHWVHAQCMDLAERTLIHLSAGSNKYYCNEHVEIARALHTPQRVLPLKKP
501	PMKSLRKKGSGKILTPAKKSFLRRLFD

* The GeneBank accession number for the amino acid sequence of human RAG2 protein is P55895.

Appendix IV

Amino acid sequence of human Ku70 subunit

1	MSGWESYYKTEGDEEAEEEQEENLEASGDYKYSGRDSLIFLVDASKAMFE
51	SQSEDELTPFDMSIQCIQSVYISKIISSDRDLLAVVFYGTEKDKNSVNFK
101	NIYVLQELDNPGAKRILELDQFKGQQGQKRFQDMMGHGSDYSLSEVLWVC
151	ANLFSDVQFKMSHKRIMLFTNEDNPHGNDSAKASRARTKAGDLRDTGIFL
201	DLMHLKKPGGFDISLFYRDIISIAEDEDLRVHFEESSKLEDLLRKVRAKE
251	TRKRALSRLKLKLNKDIVISVGIYNLVQKALKPPPIKLYRETNEPVKTKT
301	RTFNTSTGGLLLPSDTKRSQIYGSRQIILEKEETEELKRFDDPGLMLMGF
351	KPLVLLKKHHYLRPSLFVYPEESLVIGSSTLFSALLIKCLEKEVAALCRY
401	TPRRNIPPYFVALVPQEEELDDQKIQVTPPGFQLVFLPFADDKRKMPFTE
451	KIMATPEQVGKMKAIVEKLRFTYRSDSFENPVLQQHFRNLEALALDLMEP
501	EQAVDLTLPKVEAMNKRLGSLVDEFKELVYPPDYNPEGKVTKRKHDNEGS
551	GSKRPKVEYSEEELKTHISKGTLGKFTVPMLKEACRAYGLKSGLKKQELL
601	EALTKHFQD

* The GeneBank accession number for the amino acid sequence of human Ku70 subunit is P12956.

Appendix V

Amino acid sequence of human Ku80 subunit

1	MVRSGNKAAVVLCMDVGFTMSNSIPGIESPFEQAKKVITMFVQRQVFAEN
51	KDEIALVLFGTDGTDNPLSGGDQYQNITVHRHLMLPDFDLLEDIESKIQP
101	GSQQADFLDALIVSMDVIQHETIGKKFEKRHIEIFTDLSSRFSKSQLDII
151	IHSLKKCDISLQFFLPFSLGKEDGSGDRGDGPFRLGGHGPSFPLKGITEQ
201	QKEGLEIVKMVMISLEGEDGLDEIYSFSESLRKLCVFKKIERHSIHWPCR
251	LTIGSNLSIRIAAYKSILQERVKKTWTVVDAKTLKKEDIQKETVYCLNDD
301	DETEVLKEDIIQGFRYGSDIVPFSKVDEEQMKYKSEGKCFSVLGFCKSSQ
351	UQRRFFMGNQVLKVFAARDDEAAAVALSSLIHALDDLDMVAIVRYAYDKR
401	ANPQVGVAFPHIKHNYECLVYVQLPFMEDLRQYMFSSLKNSKKYAPTEAQ
451	LNAVDALIDSMSLAKKDEKTDTLEDLFPTTKIPNPRFQRLFQCLLHRALH
501	PREPLPPIQQHIWNMLNPPAEVTTKSQIPLSKIKTLFPLIEAKKKDQVTA
551	QEIFQDNHEDGPTAKKLKTEQGGAHFSVSSLAEGSVTSVGSVNPAENFRV
601	LVKQKKASFEEASNQLINHIEQFLDTNETPYFMKSIDCIRAFREEAIKFS
651	EEQRFNNFLKALQEKVEIKQLNHFWEIVVQDGITLITKEEASGSSVTAEE
701	AKKFLAPKDKPSGDTAAVFEEGGDVDDLLDMI

* The GeneBank accession number for the amino acid sequence of human Ku80 subunit is NP_066964.

RAG2 proteins. Through the course of my Master's, I obtained valuable experience in independent research. I became very familiar with techniques such as Mammalian cell culture, cloning techniques (PCR, site-directed mutagenesis, restriction enzyme digestion, ligation, electroporation, and small-, midi- and large-scale DNA preparation, SDS-PAGE, western-blotting, (co)immunoprecipitation, GST/MBP/His pull-down assay, fusion protein expression and purification, protein in vitro translation, in vitro protein binding assay.

2003 January to June **4th year Honours Student at Ocean University of China**
The focus of my Bachelor's thesis was to isolate, purify and detect Lymphocystis Disease Virus from the infected tissues and culture cell of Flounder (Paralichthys olibvaceus). I obtained valuable exprerience in cell culture, PCR, virus purification.

Scholarship and award

January 2004-December 2004	international scholarship awarded by university of Ottawa
September 2003-September 2004	admission scholarship awarded by university of Ottawa
September 2002-July 2003	annual scholarship for outstanding students awarded by Ocean university of China