V(D)J Recombination

ADVANCES IN EXPERIMENTAL MEDICINE AND BIOLOGY

Recent Volumes in this Series

Volume 642
THE SARCOMERE AND SKELETAL MUSCLE DISEASE
Edited by Nigel G. Laing

Volume 643
TAURINE 7
Edited by Junichi Azuma

Volume 644
TROPOMYOSIN
Edited by Peter Gunning

Volume 645
OXYGEN TRANSPORT TO TISSUE XXX
Edited by Per Liss, Peter Hansell, Duane F. Bruley, and David K. Harrison

Volume 646
EARLY NUTRITION PROGRAMMING AND HEALTH OUTCOMES IN LATER LIFE
Edited by Berthold Koletzko, Tamás Desci, Denes Molnár, and Anne De la Hunty

Volume 647
THERAPEUTIC TARGETS OF THE TNF SUPERFAMILY
Edited by Iqbal Grewal

Volume 648
ARTERIAL AND ALLIED CHEMORECEPTORS
Edited by Constancio Gonzalez, Colin A. Nurse, and Chris Peers

Volume 649
MOLECULAR MECHANISMS OF SPONDYLOARTHROPATHIES
Edited by Carlos López-Larrea, and Roberto Díaz-Peña

Volume 650
V(D)J RECOMBINATION
Edited by Pierre Ferrier

V(D)J Recombination

Edited by

Pierre Ferrier, MD, PhD

Centre d'Immunologie de Marseille-Luminy, Université d'Aix Marseille, Marseille, France

Springer Science+Business Media, LLC

Landes Bioscience

Springer Science+Business Media, LLC
Landes Bioscience

Printed in the USA.

Springer Science+Business Media, LLC, 233 Spring Street, New York, New York 10013, USA
http://www.springer.com

Please address all inquiries to the publishers:
Landes Bioscience, 1002 West Avenue, Austin, Texas 78701, USA
Phone: 512/ 637 6050; FAX: 512/ 637 6079
http://www.landesbioscience.com

The chapters in this book are available in the Madame Curie Bioscience Database.
http://www.landesbioscience.com/curie

V(D)J Recomination, edited by Pierre Ferrier. Landes Bioscience / Springer Science+Business Media, LLC dual imprint / Springer series: Advances in Experimental Medicine and Biology.

ISBN: 978-1-4419-0295-5

Library of Congress Cataloging-in-Publication Data

V(D)J recombination / edited by Pierre Ferrier.
p. ; cm. -- (Advances in experimental medicine and biology ; v. 650)
Includes bibliographical references and index.
ISBN 978-1-4419-0295-5
1. Genetic recombination. 2. DNA-binding proteins. I. Ferrier, Pierre, 1951- II. Series: Advances in experimental medicine and biology ; v. 650.
[DNLM: 1. Recombination, Genetic. 2. VDJ Recombinases. 3. DNA-Binding Proteins. 4. Recombinant Proteins. W1 AD559 v.650 2009 / QU 475 V393 2009]
QH443.V35 2009
571.9'648--dc22
2009011105

PREFACE

V(D)J recombination: for the community of immunologists and developmental biologists, the molecular route by which B and T lymphocytes acquire their unique function of affording adaptive immunity. Yet, for many—from experienced scientists to trainees—it represents a (rather too) sophisticated process whose true insight is excessively demanding. However, when not simply considered as a private ground for a few aficionados, it can be seen as a way of understanding how mature lymphocytes carry on their basic functions. For the group of aficionados—which includes this editor—it is an elegant paradigm featuring many fascinating evolutionary achievements of which the biological world alone has the secret. These include a subtle biochemical principle most likely hijacked some 470 million years ago from an ancestral gene invader and since then cleverly adapted by jawed vertebrates to precisely cleave and rearrange their antigen receptor (Ig and TCR) loci. This invader would itself have assigned the services of the nonhomologous end joining (NHEJ) DNA repair machinery as well as various DNA polymerases or transferases to work in concert with developmental clues in lymphoid cell lineages to generate an immune repertoire and efficient host surveillance while avoiding autoimmunity.

Recently, important new refinements in these systems have emerged, continuing to challenge our knowledge and beliefs. These are just the topics covered by the senior authors—all established leaders in this field—and their colleagues, whilst writing the various chapters in *V(D)J Recombination*. They lead us through the latest findings concerning the biochemical properties of the V(D)J recombinase (Swanson), its buried and potentially harmful transposase and translocase activities (Oettinger; Roth), the increasing importance of NHEJ, whose dysfunction causes severe forms of immune deficiencies (de Villartay), and the numerous facets in the control of gene rearrangement via non-coding RNA transcription and exquisitely regulated changes in chromosomal structure (Corcoran; Feeney; Jouvin-Marche; Krangel; Oltz and Spicuglia).

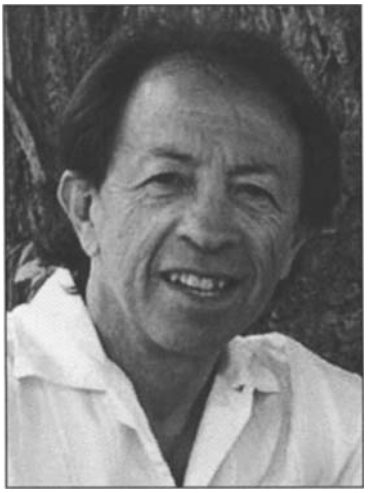
Burning progress on regulatory aspects has included the large-scale dynamics and nuclear compartmentalization of Ig and TCR loci (Singh), the anticipated—but difficult to ascertain—role of dedicated transcription factors (Zhang), the relationships between structural properties of the recombination core apparatus and its cell cycle phase-dependant accumulation/degradation or connection to the chromatin

template (Desiderio), the evolution of these regulatory aspects throughout the phylogeny (Hsu), and how abnormalities in the recombination apparatus/process can contribute to lymphoid malignancies (Macintyre).

Overall, *V(D)J Recombination* represents a tour over this, in all respects, vital process and I would like to greatly acknowledge the efforts of these eminent colleagues for concisely describing its so many aspects. We believe that every advance in this field contributes to strengthening knowledge of fundamental importance both academically and clinically. Together, we hope that the result is an attractive book which will captivate its readers and encourage some to pursue further digging in this seemingly inexhaustible mine of biological resources.

Pierre Ferrier, MD, PhD

ABOUT THE EDITOR...



PIERRE FERRIER is a Principal Investigator and Research Director at the Centre d'Immunologie de Marseille-Luminy (CIML), France. He has also worked as a Director of Marseille-Nice Genopole, a local consortium of more than twenty laboratories aimed at developing high-throughput research techniques in genomics. Main research interests include the analysis of the molecular mechanisms responsible for the control of gene expression and recombination programs during hematopoietic cell development and pathogenesis. He is a member of several national and international scientific organizations including the Institut National de la Santé et de la Recherche Médicale (Inserm), the Agence Nationale de la Recherche (ANR), the Association pour la Recherche sur le Cancer (ARC), the Human Frontier Science Program Organization (HFSPO), and the Université Virtuelle Médicale de Monaco (UVMM). Pierre Ferrier received his academic degrees from Montpellier (MD) and Marseille (PhD) Universities, France. He was a post-doctoral fellow (1986-90) in the laboratory of Prof. F.W. Alt at the Columbia University College of Physicians and Surgeons, New York, NY, USA.

PARTICIPANTS

Iratxe Abarrategui
Centre for Epigenetics
Biotech Research and Innovation Centre
Copenhagen
Denmark

Vahid Asnafi
INSERM EMI U0210
Hopital Necker-Enfants Malades
AP-HP
University Paris-Descartes
Paris
France

Kheïra Beldjord
INSERM EMI U0210
Hopital Necker-Enfants Malades
AP-HP
University Paris-Descartes
Paris
France

Daniel J. Bolland
Laboratory of Chromatin and Gene Expression
Babraham Institute
Babraham Research Campus
Cambridge
UK

Marie Bonnet
Centre d'Immunologie de Marseille-Luminy
Université d'Aix Marseille
CNRS, UMR 6102
Inserm, U 631
Marseille
France

Vicky L. Brandt
Department of Pathology and Program in Molecular Pathogenesis
The Helen L. and Martin S. Kimmel Center for Biology and Medicine
Skirball Institute for Biomolecular Medicine
New York University School of Medicine
New York, New York
USA

Robin Milley Cobb
Department of Microbiology and Immunology
Vanderbilt University
Nashville, Tennessee
USA

Anne E. Corcoran
Laboratory of Chromatin and Gene Expression
Babraham Institute
Babraham Research Campus
Cambridge
UK

Saïda Dadi
Centre d'Immunologie
de Marseille-Luminy
Université d'Aix Marseille
Marseille, France
CNRS, UMR 6102
Inserm, U 631
Marseille
France

Stephen Desiderio
Department of Molecular Biology
and Genetics
Institute for Cell Engineering
The Johns Hopkins University
School of Medicine
Baltimore, Maryland
USA

Jean-Pierre de Villartay
INSERM, U768, Unité Développement
Normal
et Pathologique du Système Immunitaire
and
Faculté de Médecine René Descartes
Université Paris-Descartes
and
AP-HP, Hôpital Necker Enfants Malades
Service d'Immunologie et
d'Hématologie Pédiatrique
Paris
France

Ann J. Feeney
The Scripps Research Institute
Department of Immunology
La Jolla, California
USA

Pierre Ferrier
Centre d'Immunologie
de Marseille-Luminy
Université d'Aix Marseille
CNRS, UMR 6102
Inserm, U 631
Marseille
France

Patrizia Fuschiotti
Department of Immunology
University of Pittsburgh
School of Medicine
Pittsburgh, Pennsylvania
USA

Ellen Hsu
Department of Physiology
and Pharmacology
State University of New York
Health Science Center at Brooklyn
Brooklyn, New York
USA

Kristen Johnson
Department of Molecular Genetics
and Cell Biology
Howard Hughes Medical Institute
The University of Chicago
Chicago, Illinois
USA

Mary Elizabeth Jones
Department of Immunology
Duke University Medical Center
Durham, North Carolina
USA

Evelyne Jouvin-Marche
INSERM, U823
Faculté de Médecine
Institut d'Oncologie/Developpement
Albert Bonniot et Institut Français
du Sang
Université Joseph Fourier-Grenoble I
Grenoble
France

Michael S. Krangel
Department of Immunology
Duke University Medical Center
Durham, North Carolina
USA

Sushil Kumar
Department of Medical Microbiology
and Immunology,
Creighton University Medical Center
Omaha, Nebraska
USA

Sandrine Le Noir
INSERM EMI U0210
Hôpital Necker-Enfants Malades
AP-HP
Université Paris-Descartes
Paris
France

Yun Liu
Department of Molecular Biology
and Genetics
Institute for Cell Engineering
The Johns Hopkins University
School of Medicine
Baltimore, Maryland
USA

Elizabeth A. Macintyre
INSERM EMI U0210
Hôpital Necker-Enfants Malades
AP-HP
Université Paris-Descartes
Paris
France

Patrice Noël Marche
INSERM, U823
Faculté de Médecine
Institut d'Oncologie/Developpement
Albert Bonniot et Institut Français
du Sang
Université Joseph Fourier-Grenoble I
Grenoble
France

Adam G.W. Matthews
Department of Molecular Biology
Massachusetts General Hospital
and
Department of Genetics
Harvard Medical School
Boston, Massachusetts
USA

Marjorie A. Oettinger
Department of Molecular Biology
Massachusetts General Hospital
and
Department of Genetics
Harvard Medical School
Boston, Massachusetts
USA

Eugene M. Oltz
Department of Microbiology
and Immunology
Vanderbilt University
Nashville, Tennessee
USA

Prafulla Raval
Department of Medical Microbiology
and Immunology
Creighton University Medical Center
Omaha, Nebraska
USA

Karen L. Reddy
Department of Molecular Genetics
and Cell Biology
Howard Hughes Medical Institute
The University of Chicago
Chicago, Illinois
USA

David B. Roth
Department of Pathology and Program
in Molecular Pathogenesis
The Helen L. and Martin S. Kimmel
Center for Biology and Medicine
Skirball Institute for Biomolecular
Medicine
New York University School
of Medicine
New York, New York
USA

Harinder Singh
Department of Molecular Genetics
and Cell Biology
Howard Hughes Medical Institute
The University of Chicago
Chicago, Illinois
USA

Salvatore Spicuglia
Centre d'Immunologie
de Marseille-Luminy
Université d'Aix Marseille
CNRS, UMR 6102
Inserm, U 631
Marseille
France

Patrick C. Swanson
Department of Medical Microbiology
and Immunology
Creighton University Medical Center
Omaha, Nebraska
USA

Lance R. Thomas
Department of Microbiology
and Immunology
Vanderbilt University
Nashville, Tennessee
USA

Andrew L. Wood
Laboratory of Chromatin and Gene
Expression
Babraham Institute
Babraham Research Campus
Cambridge
UK

Li Zhang
Department of Molecular Biology
and Genetics
Institute for Cell Engineering
The Johns Hopkins University
School of Medicine
Baltimore, Maryland
USA

Yuan Zhuang
Department of Immunology
Duke University Medical Center
Durham, North Carolina
USA

CONTENTS

CHAPTER 1

Early Steps of V(D)J Rearrangement: Insights from Biochemical Studies of RAG-RSS Complexes

Patrick C. Swanson,* Sushil Kumar and Prafulla Raval

Abstract

V(D)J recombination is initiated by the synapsis and cleavage of a complementary (12/23) pair of recombination signal sequences (RSSs) by the RAG1 and RAG2 proteins. Our understanding of these processes has been greatly aided by the development of in vitro biochemical assays of RAG binding and cleavage activity. Accumulating evidence suggests that synaptic complex assembly occurs in a step-wise manner and that the RAG proteins catalyze RSS cleavage by mechanisms similar to those used by bacterial transposases. In this chapter we will review the molecular mechanisms of RAG synaptic complex assembly and 12/23-regulated RSS cleavage, focusing on recent advances that shed new light on these processes.

Introduction

The antigen-binding variable domains of immunoglobulins and T-cell receptors exhibit great structural diversity that mostly originates from a site-specific DNA rearrangement process, called V(D)J recombination, that assembles the exons encoding the variable domains of these proteins from germline variable (V), diversity (D) and joining (J) gene segments during lymphocyte development.[1] Adjacent to each gene segment lies a recombination signal sequence (RSS); each RSS contains a conserved heptamer and nonamer motif (consensus heptamer: 5'-CACAGTG-3'; consensus nonamer: 5'-ACAAAAACC-3') separated by "spacer" DNA, normally 12 base pairs (bp) or 23 bp long (12-RSS and 23-RSS, respectively), which displays some sequence preferences proximal to the heptamer[2] but is otherwise not well conserved. V(D)J recombination is generally directed between two gene segments with different RSSs, a restriction termed the 12/23 rule that serves to facilitate productive receptor gene assembly.

The biochemistry of V(D)J recombination can be conceptually divided into a cleavage phase and a joining phase (Fig. 1). To initiate the cleavage phase, two lymphoid cell-specific proteins encoded by recombination activating gene-1 and -2 (RAG1 and RAG2, respectively[3,4]), possibly assisted by high mobility group proteins of the HMG-box family (HMGB1 and HMGB2, called HMGB1/2 henceforth; discussed further below), bring two different gene segments into close proximity through interactions with the adjoining 12- and 23-RSS (forming a "synaptic" complex) and then catalyze a DNA double-strand break (DSB) at each RSS between the heptamer and the coding segment.[5,6] RAG-mediated cleavage produces two types of DNA ends: blunt and 5'-phosphorylated signal ends containing the RSS and coding ends covalently sealed as DNA hairpins.[7,8] These reaction intermediates originate from a two-step cleavage mechanism in which

*Corresponding Author: Patrick C. Swanson—Department of Medical Microbiology and Immunology, Creighton University Medical Center, Omaha, Nebraska 68178, USA. Email: pswanson@creighton.edu

V(D)J Recombination, edited by Pierre Ferrier.

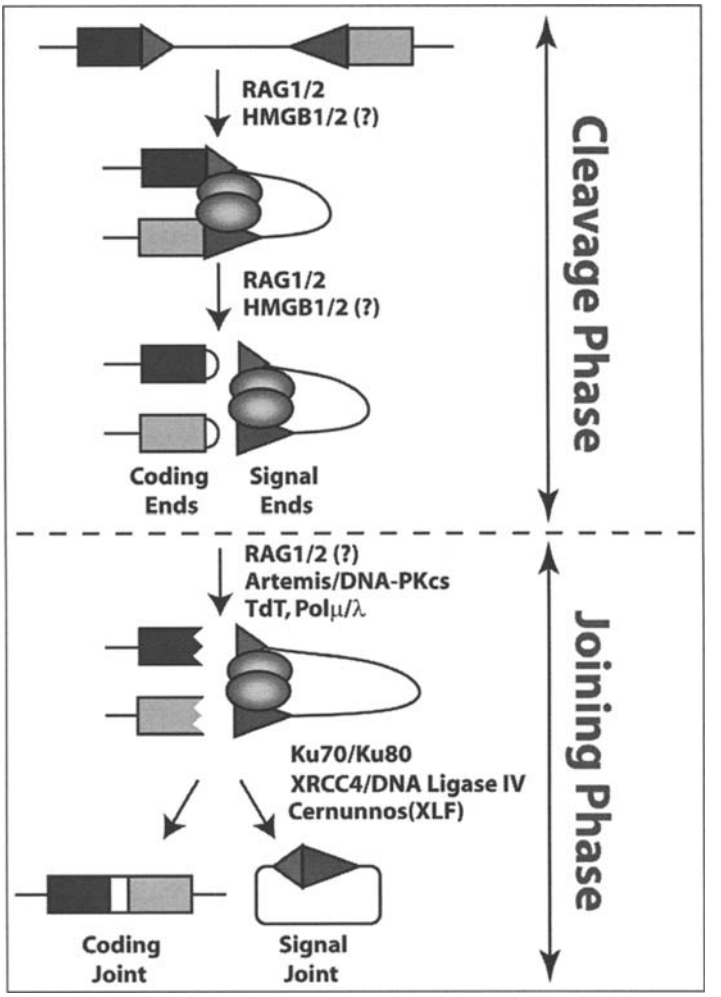


Figure 1. Overview of V(D)J recombination (adapted from Fugmann et al[6]). In the cleavage phase of V(D)J recombination, coding segments (filled rectangles), flanked by a 12-RSS or 23-RSS (small or large triangles, respectively) are assembled into a synaptic complex by the RAG proteins, possibly assisted by HMGB1/2 (filled ovals). Coupled cleavage by the RAG proteins yields blunt signal ends and coding ends sealed as DNA hairpins. In the joining phase of V(D)J rearrangement, sealed coding ends are resolved by an Artemis/DNA-PKcs complex and may be further processed by TdT (if present) and DNA polymerases μ and/or λ (Polμ/λ). Processed coding ends are joined to create imprecise coding joints that may have gained palindromic (P) or nontemplated (N) nucleotides through asymmetric hairpin opening or TdT-mediated addition, respectively, or lost nucleotides through end processing reactions (open rectangle). Signal ends are joined to create signal joints that are typically precise. Alternative, less frequent joining events, such as open-shut and hybrid joints are not shown for simplicity. Signal and coding joint formation is mediated by the NHEJ pathway, which includes Ku70, Ku80, XRCC4, DNA Ligase IV and Cernunnos (XLF). Although the processing and joining reactions are shown as sequential processes, these steps may be integrated and iterative for joining of incompatible coding ends, involving single-strand ligation, processing of the unligated strand by Artemis/DNA-PKcs and DNA polymerases and eventual ligation of the second strand resulting in repaired double-stranded DNA.[101]

the RAG proteins first nick the RSS at the 5' end of the heptamer and then use the resulting 3'-OH to catalyze a direct transesterification reaction on the opposing phosphodiester bond.[9] In the joining phase, the two signal ends are typically ligated precisely, forming a signal joint, and the coding ends are subjected to reactions that resolve the hairpins and then process and connect the DNA ends to form coding joints. As a result, coding joints often show evidence of nucleotide gain or loss at the coding ends. Infrequently, alternative outcomes of V(D)J recombination are observed in which one gene segment is joined to the RSS of another gene segment ("hybrid joint") or is separated and rejoined to the same RSS ("open-shut joint").[10,11] Efficient signal and coding joint formation requires a competent nonhomologous end-joining (NHEJ) repair pathway, including Ku70, Ku80, XRCC4, DNA Ligase IV and XLF/Cernunnos.[12-14] Coding joint formation requires two additional factors not strictly essential for joining signal ends, Artemis and DNA-PKcs, which together function as a structure-specific endonuclease responsible for opening the DNA hairpins on coding ends.[15] Asymmetric hairpin opening can give rise to palindromic (P) nucleotides being inserted in coding joints. Terminal deoxynucleotidyl transferase (TdT) and DNA polymerases μ and/or λ (Pol μ/λ) can further diversify these junctional sequences by catalyzing addition of nontemplated (N) nucleotides to coding ends (TdT) and processing incompatible DNA ends to facilitate end-joining (Pol μ/λ).[16,17] A detailed consideration of the proteins involved in the processing and repair of V(D)J recombination intermediates is beyond the scope of this review, but has been discussed elsewhere.[18-20]

Here we review and discuss the molecular mechanisms of V(D)J recombination, focusing on the cleavage phase of this process and emphasizing new insights. Readers are referred to previous reviews for more detailed discussion of early studies of RAG protein biochemistry, including the establishment of cell-free assays of V(D)J cleavage and joining[5,6] and the identification and characterization of the various structural domains of the RAG proteins.[21]

Assembly and Organization of Single Site and Synaptic RAG-RSS Complexes

Cell-free assays of V(D)J cleavage established using truncated, catalytically active "core" forms of RAG1 (full-length 1040 a.a.; core residues 384-1008) and RAG2 (full-length 517 a.a.; core residues 1-387) demonstrated that the RAG1/2 complex is both necessary and sufficient to mediate RSS cleavage[9] and that RAG cleavage activity exhibits metal ion-dependence: Mn^{2+} supports RAG-mediated cleavage of a single RSS, whereas Mg^{2+} is required for coupled cleavage of RSS pairs abiding by the 12/23 rule.[22,23] In natural progression, later studies identified and characterized discrete RAG-RSS complexes with increasing complexity, with early work focused on RAG complexes assembled on a single RSS and later work analyzing higher-order RAG synaptic complexes. Most of this work has been reviewed and discussed elsewhere.[5,6,24] Therefore, only salient features will be highlighted here.

Core RAG1 contains three structurally distinct regions:[21] an amino-terminal nonamer binding domain (NBD, residues 389-442) that interacts with the RSS nonamer,[25,26] a central domain (residues 528-760) that recognizes the heptamer and exhibits single-strand DNA binding activity and a C-terminal domain (residues 761-979) that binds double-stranded DNA nonspecifically and cooperatively. Core RAG1 alone exists in solution primarily as a stable dimer[27-29] and binds an isolated RSS with moderate affinity (Kd ~41 nM)[28] as a dimer[27,28,30] (although higher-order aggregates are detectable at elevated RAG1 concentrations and conditions of low ionic strength[31]) whereas RAG2 is predominantly monomeric in solution[29] and shows little, if any DNA binding activity.[25,26,32-34] RAG1 and RAG2 interact with one another in the absence of DNA[27,29,35] and together bind a single RSS with greater specificity than RAG1 alone.[32,33,36] Purified core RAG1/2 proteins variably assemble one[29,32,33] or two[34,37] major protein-DNA complexes detectable using an electrophoretic mobility shift assay (EMSA). The relative abundance of these complexes, now generally called SC1 and SC2 (for "single RSS complex"), depends partly on how the RAG proteins are expressed and purified:[37,38] in our laboratory, individually expressed and purified RAG proteins tend to assemble only SC1, coexpressed RAG proteins purified under high salt conditions form

more SC1 than SC2 and coexpressed RAG proteins purified using milder conditions predominantly assemble SC2. Both complexes possess similar intrinsic cleavage activity,[34,37] but differ in RAG protein stoichiometry. Swanson reported that both complexes contain a RAG1 dimer, but incorporate either one (SC1) or two (SC2) RAG2 molecules.[37] Mundy et al reported comparable results for RAG2 in these complexes, but presented evidence suggesting SC1 and SC2 contain three or more RAG1 subunits.[34] Possible explanations for this apparent discrepancy have been discussed previously[24] and will not be revisited here, but we note that recent data reported by De et al provides corroborating evidence supporting the contention that RAG1 exists as a dimer in an SC (RAG2 stoichiometry was not determined).[39] The tetrameric RAG1/RAG2 configuration reported for SC2 is also consistent with data published by Bailin et al.[29]

Mutagenesis studies[40-42] revealed that RAG1 contains three carboxylate residues (asp-600, asp-708 and glu-962) critical for catalysis that resemble a "DDE motif" found in many transposases and integrases.[43] Similar to the Tn5 transposase,[44,45] biochemical studies established that a single RAG1 subunit contributes all three carboxlate residues to single active site which mediates sequential nicking and hairpin formation steps of the cleavage reaction[46,47] and that these reactions are catalyzed in trans; that is, by the subunit of the RAG1 heterodimer not bound to the nonamer of the RSS being cleaved.[47]

While the RAG proteins themselves are sufficient for assembling SC1 and SC2, HMGB1/2 proteins are known to facilitate RAG-mediated binding and cleavage of an isolated 23-RSS, but not a 12-RSS, in vitro.[48] The RAG proteins also require the presence of HMGB1/2 to efficiently assemble a complex containing a complementary (12/23) pair of RSSs ("paired complex" or PC) and mediate coupled cleavage at both RSSs adhering to the 12/23 rule in vitro.[48,49] Whether HMGB1/2 also assist the RAG proteins during V(D)J recombination in vivo has not been formally established nor entirely ruled out,[50] since HMGB1/2 exhibit functional redundancy in RAG binding and cleavage assays.[51] The HMGB1/2 proteins are nonhistone chromosomal DNA binding proteins known to promote DNA bending and facilitate assembly of nucleoprotein complexes;[52] HMGB1 further functions as an alarmin to signal cellular damage in response to inflammatory processes.[53] HMGB1/2 proteins contain tandem homologous HMG-box domains (called A and B) attached to a basic linker and an acidic tail. HMGB1/2 interacts with the NBD of RAG1 in the absence of DNA and enhances the intrinsic DNA bending activity of the RAG proteins.[54] The integration of HMGB1/2 into RAG-RSS complexes can often be detected as a supershift by EMSA.[51,55] Recent structure-function studies conducted in our laboratory[56,57] suggest that both HMG-box domains must be competent to bend DNA and physically linked together in either orientation (AB or BA) to stimulate RAG-mediated 23-RSS cleavage in the presence of Mg^{2+}. Interestingly, single HMG-box domains can be integrated into 23-RSS-RAG complexes,[56-58] but cannot stimulate 23-RSS cleavage unless Mn^{2+} replaces Mg^{2+} in the reaction,[57,58] or 12-RSS partner is added to promote synapsis.[57] These results suggest the two HMG-box domains have separable but potentially redundant roles in stimulating RAG binding and cleavage activity in vitro and that synapsis promotes a conformational change that bypasses the need for one of these domains. HMGB1 lacking the acidic tail stimulates RAG binding and cleavage activity at lower concentrations than full-length HMGB1, but promotes aggregation of RAG-RSS complexes.[56,58] Moreover, loss of the acidic tail enables HMGB1 mutants that otherwise fail to support RAG-mediated synapsis to stimulate PC formation.[56] These data suggest the acidic tail helps maintain the correct oligomerization state of RAG synaptic complexes. The acidic tail is also known to facilitate HMGB1-mediated nucleosome repositioning,[59,60] which may help promote RSS accessibility in nucleosomal DNA.[61-63]

Synaptic complex assembly is thought to proceed via initial formation of SC2 followed by capture of an appropriate partner RSS to form a PC. This "capture model" of assembly was suggested initially by biochemical experiments demonstrating that SC2 can be driven to form the PC by adding appropriate partner RSS[34] and the observation that RAG cleavage activity is greater when synaptic complexes are assembled in step-wise fashion by adding free 23-RSS to a 12-RSS-RAG complex (or vice versa) than when they are assembled by mixing preformed 12-RSS-RAG and

23-RSS-RAG complexes together.[64] This model has gained in vivo experimental support from a recent study by Curry et al[65] showing that nicks can be detected at endogenous 12-RSSs, but not at 23-RSSs, in lymphoid cells. These findings lead the authors to propose a model in which RAG proteins bind and nick a 12-RSS first, then capture and nick a 23-RSS and, in rapid succession, finally cleave both RSSs. This model is consistent with previous biochemical studies showing that nicking can occur on an RSS in the absence of synapsis,[66,67] but nicking at one RSS is required for efficient cleavage of its partner.[22,66] The capture model is also consistent with data this laboratory and others have published showing that the complement of RAG proteins is the same between a RAG complex bound to a single RSS (as SC2) and the PC.[34,37] Interestingly, these studies show that molecules of RAG2, but not RAG1, freely re-assort during PC assembly.[34,37] Work from this laboratory suggests that the PC contains two molecules each of RAG1 and RAG2 and that this heterotetramer configuration remains the same through the cleavage steps of V(D)J recombination.[37] Another study reported the same stoichiometry for RAG2 in the PC,[34] but others conclude the PC contains three or more RAG1 subunits.[34,46] Possible scenarios to explain these discordant results have been discussed elsewhere.[24]

How are the RSSs arranged in the synaptic complex? Early observations that the efficiency of in vitro coupled cleavage[22] and in vivo V(D)J rearrangement[68] is more sensitive to shortening of the intersignal distance when the RSSs are positioned in an inversional configuration than when they are positioned in a deletional configuration argued that the RSSs are aligned in a parallel, rather than anti-parallel orientation in the synaptic complex. To test this possibility more directly, Cibutaru et al recently measured levels of fluorescence resonance energy transfer (FRET) in RAG synaptic complexes assembled under various conditions on 12- and 23-RSS oligonucleotide substrates labeled with FAM and TAMRA in different configurations.[69] Significant FRET was detected only when the following three conditions were met: (i) the fluorophores were placed on different RSSs (but not the same RSS); (ii) the two RSSs contained different length spacers (i.e., abiding by the 12/23 rule); and, (iii) synaptic complexes were assembled in binding reactions containing Mg^{2+} and the full complement of RAG1/2 and HMGB1/2 proteins. Interestingly, FRET was observed in synaptic complexes regardless of which end of a given RSS was labeled; the only apparent requirement was that the two fluorophores were placed on different RSSs (12 and 23). These data suggest that the distance between the ends of the two bound RSSs in the synaptic complex are approximately the same. Given this constraint and limitations on the maximal distance between fluorophores to observe FRET, the authors propose the two RSSs likely adopt a bent and crossed configuration in the PC.[69]

Insights into RAG-Mediated RSS Recognition and Cleavage Mechanisms

Interactions between the RAG proteins and DNA have been investigated using a variety of approaches and the insights from these studies have greatly improved our understanding of how the RAG proteins recognize and cleave their RSS targets. Much of the early work has been extensively reviewed,[5,6,24] so it will not be covered in depth here. Chemical and DNase I protection and modification interference footprinting assays performed on RAG complexes assembled on a single RSS suggest RAG1 primarily interacts with the nonamer and adjacent spacer sequence, whereas RSS contacts in complexes containing both RAG proteins are overlapping, but more expansive, extending from the nonamer, through the spacer and into the 3′ end of the heptamer, with a bias of phosphate contacts toward one face of the DNA helix.[32,70,71] Photo cross-linking studies suggest RAG1 mediates most of the contact with the RSS, with RAG2-RSS interactions more localized to the junction of the heptamer and coding segment.[27,36,72,73] Integration of HMGB1/2 into 23-RSS-RAG complexes enables detection of heptamer-spacer contacts resembling those observed in 12-RSS-RAG complexes that are not otherwise visualized in 23-RSS complexes containing RAG1/2 alone,[51,55] suggesting HMGB1 stabilizes RAG association with the heptamer in these complexes. Ethylation interference footprinting suggests HMGB1/2 contacts the 23-RSS proximal to the nonamer, expanding the footprint of the RAG proteins in this region.[51] Although RAG contacts at the junction of the heptamer and coding sequence are not readily detected in RAG

complexes assembled on a single RSS, this region is protected from DNase I cleavage in synaptic complexes.[74] Nagawa et al showed that synaptic complexes assembled with nicked RSS substrates show slight expansion of the DNase I footprint relative to precleavage synaptic complexes (from ~12 nt to ~16 nt), suggesting that RAG-mediated nicking causes more intimate and stable RAG association with the coding sequence.[75] Pull-down assays showing that nicked RSS substrates are more readily incorporated into synaptic complexes than intact substrates support this contention. Interestingly, two different joining-deficient RAG1 mutants (S723C[76] and K118/9A[77]) were shown to exhibit poor protection of the heptamer-coding junction, leading to speculation that the joining defect is caused by poor coding end retention in the postcleavage synaptic complex.[75] However, close inspection of the mutant RAG1 footprinting patterns in precleavage complexes also reveals that these mutants exhibit less protection of spacer and nonamer sequences compared to wild-type RAG1. This observation argues that these mutations cause a global defect in RAG-RSS complex stability, but can also be interpreted to suggest that the RAG proteins require stable contact with the coding sequence in order to maintain strong interactions with the RSS (or vice versa) in precleavage complexes.

Direct and interference footprinting experiments suggest RAG-RSS complex formation is accompanied by structural distortions in the spacer region and near the site of DNA cleavage.[32,51,70,71] Studies showing that the RAG proteins mediate RSS bending, which is augmented by HMGB1/2,[54] plausibly explain spacer hypersensivity to chemical and enzymatic probes in RAG-RSS complexes. Structural distortions near the cleavage site are likely attributed to base unpairing mediated by the RAG proteins to promote hairpin formation, which is suggested by observations that RAG-mediated RSS cleavage is facilitated by incorporating base-pair mismatches[78,79] or abasic sites[80] at the coding flank. Clues to how these structural distortions may be induced and stabilized are suggested by structural studies of the related Tn5 transposase, which, like the V(D)J recombinase, catalyzes DNA hairpin formation (except that hairpins are formed at the transposon end, which is equivalent to the signal end in V(D)J recombination).[81] Analysis of a Tn5 postcleavage synaptic complex reveals that the transposase promotes extrusion of a thymine from the DNA helix, stabilizing the "flipped base" via stacking interactions with an aromatic tryptophan residue (trp-298).[44] Recent studies indicate a similar mechanism is operative in V(D)J recombination. Two lines of evidence suggest the terminal nucleotide on the bottom strand of the coding flank (C1b, see Fig. 2 inset) is stabilized in an extrahelical configuration by the RAG proteins. First, when thymine is incorporated into the RSS at position C1b, this base exhibits hypersensitivity to permanganate modification under conditions favoring RAG-RSS synaptic complex formation.[82] Second, base removal at C1b potentiates hairpin formation.[80] Both outcomes are consistent with comparable studies of the flipped T2 thymine in the Tn5 transposon end.[83,84] One notable contrast between the two recombination systems is that although the base subjected to flipping in the RSS coding flank and the Tn5 transposon end are both located opposite the nicking site within the hairpin-forming sequence, they are offset from one another by one nucleotide: in the RSS, the base is at the terminus of the sequence; in the Tn5 transposon end, it occupies the penultimate position.

When does base-flipping occur during RSS cleavage? Base-flipping appears to occur after nicking, rather than upon RAG binding to the RSS, as permanganate hypersensitivity is not observed in RAG synaptic complexes assembled on intact substrates.[82] Interestingly, permanganate interference assays reveal that intact substrates bearing oxidized thymine at C1b and S2b are selectively bound by the RAG complex relative to unmodified substrates, with the latter modification being much preferred over the former.[32,51] If the RAG proteins stabilize base-flipping at C1b during the hairpin-forming step, why is prior modification of S2b selected over C1b in interference assays? Since base-flipping is most evident in synaptic complexes assembled on nicked substrates,[82] one possibility is that a conformational change in the RAG complex occurs after synapsis or nicking that alters the position of thymine binding pocket relative to the cleavage site. Thus, an oxidized extrahelical thymine at S2b may be preferentially accommodated over C1b in the binding pocket of a RAG complex bound to an intact RSS. Alternatively, modified S2b may be selected because C1b is more easily flipped if the oxidized base at S2b is already displaced from the DNA helix.