# Expression and V(D)J recombination activity of mutated RAG-1 proteins

Moshe J.Sadofsky, Joanne E.Hesse, J.Fraser McBlane and Martin Gellert\* Laboratory of Molecular Biology, NIDDK, NIH, Bethesda MD 20892, USA

Received September 2, 1993; Revised and Accepted October 31, 1993

# ABSTRACT

The products of the RAG-1 and RAG-2 genes ([1], [2]) are essential for the recombination of the DNA encoding the antigen receptors of the developing immune system. Little is known of the specific role these genes play. We have explored the sequences encoding mouse RAG-1 by deleting large parts of the gene and by introducing local sequence changes. We find that a RAG-1 gene with 40% of the coding region deleted still retains its recombination function. In addition, a series of small deletions within the strongly conserved remaining 60% of the coding region was tested. Nine out of ten of these prove unable to provide RAG-1 activity, but one is quite active. Certain peptide sequences were also specifically targeted for mutagenesis. The RAG-1 protein generated from this expression system is transported to the nucleus and is degraded with a 15 minute half-life. The fate of the proteins made by the deletion mutants were also assessed. Transport of RAG-1 protein to the nucleus was found even with the most extensive deletions studied. The functionality of the deleted proteins is discussed with relation to an alignment of RAG-1 sequences from five animal species.

# INTRODUCTION

The DNA recombination process that assembles the genes encoding the antigen receptors of the immune system ('V(D)J recombination') has been extensively characterized at the level of the DNA reactants and products (reviewed in [3]), but much less is known about the trans-acting factors that directly participate in the reaction. Two candidate genes, named RAG-1 and RAG-2, were isolated by virtue of their ability to activate V(D)J recombination in (normally inactive) fibroblasts (reviewed in [4]). The essential and specific role that these genes play in lymphoid V(D)J recombination is convincingly demonstrated by the behavior of knockout mice ([5], [6]). Disruption of either the RAG-1 or RAG-2 gene completely eliminates V(D)J recombination, without any other apparent effect. However, to date there is no direct evidence that either of these genes acts directly in the recombination reaction.

Tests of the functionality of RAG-1 and RAG-2 are most conveniently carried out in transfected fibroblasts ([1]), with the help of a recombination assay using extrachromosomal substrates that was previously developed in this laboratory ([7]). A combination of plasmid DNAs can be delivered to recipient cells to provide the V(D)J reaction substrate as well as sources of RAG-1 and RAG-2. In fibroblasts, mRNA derived from expression of the endogenous RAG loci is not detectable ([2]), so that V(D)J recombination is entirely dependent on RAG expression from the exogenous plasmids. It is therefore possible to test the activity of mutated versions of the RAG proteins by transfecting altered RAG genes into fibroblasts and using the extrachromosomal assay to measure the resulting V(D)J recombination. In this report we use this approach to assess the effects of various RAG-1 mutations on activity. Identifying regions of the protein that are essential for activity may provide insight into the function of the protein and help to develop a biochemical characterization of the mechanism of this recombination reaction. Related results on N- and C-terminal deletions and some complementary site specific mutations have recently been reported ([8]).

# MATERIALS AND METHODS

# Plasmids

cDNA was synthesized by reverse transcription of total RNA obtained from cell line 22D6 ([9]), an Abelson leukemia virus transformed pre-B cell line. RAG-1 and RAG-2 cDNAs were then amplified by PCR, using specific primers designed also to add restriction enzyme recognition sites adjacent to the largest open reading frame for subsequent cloning. Expressed versions of the genes were made by subcloning into the shuttle vector pCDM8 ([10]) which provides eukaryotic signals for transcription and mRNA processing. The functional constructs used in this report are designated pJH548 (RAG-1) and pJH549 (RAG-2). Plasmid pMS106 was constructed by fill-in of the MluI site in the RAG-1 coding region, followed by insertion by blunt end ligation of an oligonucleotide with translation termination codons in all three reading frames (New England BioLabs XbaI linker). This results in the substitution of a valine at residue 1010, and truncation of the remaining 30 residues in the protein product. Plasmid pMS108 was similarly constructed by insertion of the

<sup>\*</sup> To whom correspondence should be addressed

same linker at the blunt ended Bst 1107 site. This truncates the encoded protein immediately following tyrosine 994.

The vector for constructs encoding the epitope tagged proteins was first modified to remove the MluI site in pCDM8 by fill-in and blunt end ligation. In plasmid pMS119A, the sequence at the 3' end of the RAG-1 coding region between MluI and NotI (created during the cloning) was replaced with oligonucleotide A (CGCGTCCGAGCAAAAGCTCATTTCTGAAGAGGAC-TTG) followed by B (CGCGCGTTATAAG). This encodes one copy of a human c-myc epitope and recreates the flanking restriction enzyme recognition sites. Plasmid pMS119C contains three tandem copies of oligonucleotide A followed by one copy of B. Plasmid pMS122 is otherwise identical to pMS119C but for two point mutations introduced by PCR, resulting in specific mutations of tyrosines 994 and 998 to phenylalanines. Sitespecific mutagenesis to create plasmid pMS124 was performed by PCR. The region surrounding the mutations was subsequently subcloned into the parent and sequenced to assure that no spurious mutations had been inadvertently created. All the remaining deletion mutants were made by PCR amplification using specific oligonucleotide primers. Plasmid pCATE3 was synthesized by subcloning the CAT coding region from pSV2-cat ([11]) into pCDM8. Subsequent PCR amplification was performed with primers designed to add a MluI site adjacent to the 3' end of the coding region. The epitope-coding portion of pMS119C was subcloned at this site. All plasmids used for transfection were column purified (Qiagen).

#### **Cell culture**

The A-MuLV transformed pre-B cell line 1.8 ([9]) was cultured in RPMI 1640 (ICN) plus 10% fetal bovine serum, penicillin, streptomycin, and 2-mercaptoethanol (50  $\mu$ M) and incubated at 37°C in a 7% CO<sub>2</sub> atmosphere. Fibroblast lines NIH3T3 and 3TGR ([12]) were grown in DMEM (Gibco/BRL), 10% calf serum, penicillin, and streptomycin in a 5% CO<sub>2</sub> atmosphere. Monoclone MYC 1-9E10.2 ([13]) was obtained from ATCC and cultured in serum-free medium (QBSF-55, Quality Biological).



Figure 1. Graphic representation of the mouse RAG-1 gene with location of several mutations. The full 1040 amino acid sequence is drawn, with the minimal core that showed recombinational activity striped. A cysteine and histidine-rich region is indicated in the figure. Also marked is the proposed homology to the yeast gene HPR1, with the putative active site boxed and partial homology (dashed lines) extending in both directions. Below is shown the region subcloned as construct pMS119, with the epitope tag marked with vertical dashes. The associated point mutants pMS122 and pMS124 are shown within pMS119. At the bottom of the figure is the construct pMS127 with epitope tag marked with vertical dashes. The derived constructs with small deletions are indicated below.

## **Protein expression**

Plasmids were transfected into cell line 1.8 by electroporation (Bio-Rad Gene Pulser) with  $2 \times 10^7$  cells and  $5-30 \mu g$  plasmid DNA in 0.25 ml of growth medium, in 0.2 cm width cuvettes at 200 V and 960  $\mu$ F. Protein expression was enhanced by culturing the cells in 5 mM sodium butyrate following transfection. Protein labeling experiments were performed by starving cells for 30 minutes in methionine-free medium, followed by a 30 minute pulse with 100 to 500  $\mu$ Ci of 35S methionine (NEN) per sample. Timing of the pulse and duration of the chase varied between experiments, as described in the text. Protein was harvested by extraction with RIPA buffer (150 mM NaCl, 50 mM Tris pH 8.0, 1.0% Triton X-100, 0.5% Na deoxycholate, 0.1% SDS) from intact cells or from cells fractionated into cytoplasmic and nuclear pools by prior extraction with cytoplasmic lysis buffer (60 mM NaCl, 10 mM Tris pH 7.5, 3 mM MgCl<sub>2</sub>, 30% glycerol, 0.5% Triton X-100). The protease inhibitors Aprotinin, PMSF, Leupeptin, and Pepstatin were added to both extractions. Immunoprecipitation was performed using the monoclonal anti-myc epitope antibody and recombinant protein G-agarose (Gibco/BRL).

#### Extrachromosomal substrate assay

Calcium phosphate-mediated transfection of fibroblasts was performed according to the manufacturer's instructions (Pharmacia CellPhect kit). Typical plasmid quantities in the 3.5 ml mix for a 60 mm dish were: 6  $\mu$ g of pJH200 ([7]), 2.1  $\mu$ g of RAG-1 expression plasmid and 2.5  $\mu$ g of RAG-2 expression plasmid. Quantities were adjusted to compensate for changes in molecular weight owing to deletions. Recovery and processing of recombinant plasmids was performed by Hirt extraction ([14]). In some experiments, as indicated, cells were incubated in the presence of 5mM sodium butyrate for the 40 to 48 hours following transfection.

Recombination of the substrate plasmid pJH200 leads to the expression of chloramphenicol acetyltransferase, and thus renders bacteria containing the rearranged DNA resistant to chloramphenicol. This plasmid also confers ampicillin resistance to the host DH5 $\alpha$ , while the expression constructs cannot. Replication of the plasmid in the eukaryotic cell removes the prokaryotic DNA methylation pattern, and makes the replicated DNA resistant to the restriction enzyme DpnI. Digestion with DpnI therefore eliminates the background of substrate molecules that failed to enter the eukaryotic cell, and allows a measurement of recombination frequency. In each experiment, an expression plasmid containing one of the RAG-1 variants was cotransfected with the RAG-2 expression plasmid (pJH549) and pJH200. Plasmids were recovered after 40-48 hours of incubation, digested with DpnI, and selected in bacteria for ampicillin and chloramphenicol resistance. Colonies that acquired chloramphenicol resistance were further characterized by colony lift hybridization to an oligonucleotide that would anneal under stringent conditions only to a perfect signal junction ([2]). The number of resulting positive colonies, when compared to the number of colonies obtained from selection on plates containing ampicillin alone, allowed the calculation of a relative level of recombination.

#### **Computer analysis**

The multiple sequence alignment was assembled using the Pileup program of the GCG sequence analysis software package and subsequently modified manually for display purposes.

#### **RESULTS AND DISCUSSION**

#### **Carboxy-terminal alterations of RAG-1**

Figure 1 shows the mouse RAG-1 gene and many of the mutations discussed in this report. Table 1 shows the RAG-1 sequences contained in each expression plasmid and the associated recombination activity.

A series of RAG-1 expression constructs were prepared that modify the carboxy terminus of the protein. Recombination activity was assayed by cotransfection of these RAG-1 expression plasmids with a RAG-2 expression plasmid (pJH549) and the test substrate pJH200, a plasmid which retains a signal joint upon recombination. In experiments without butyrate induction, the unmodified RAG-1 expression plasmid, pJH548, gave 0.4% recombination. Truncation of the C-terminal 31 residues to amino acid 1009 (plasmid pMS106) had little measurable effect. However, further truncation to residue 994 (plasmid pMS108) eliminated recombination activity. This result and that of another carboxy-terminal alteration (pMS122) will be discussed later.

Plasmid pMS119A removes the C-terminal 32 amino acids, and adds 14, in which are contained the 10 residues (EQKLISEEDL) that constitute the specific epitope recognized by the monoclonal antibody MYC 1-9E10.2 ([13]). Plasmid pMS119C contains three tandem copies of the epitope tag. Both plasmids supported recombination at levels comparable to the unmodified control, demonstrating that the epitope tag does not interfere with RAG-1 activity in this assay.

We note that treatment of the transfected fibroblasts with sodium butyrate increases the recombination frequency by a factor of 5 to 10. This may reflect increased transfection efficiency ([15]), and/or a specific induction of the CMV promoter contained in the expression plasmids ([16]), and possibly other effects.

#### Amino-terminal deletions of RAG-1

A series of constructs progressively truncating the coding region of the mouse RAG-1 gene from its 5' end was generated starting from plasmid pMS119C. Each construct was designed to initiate translation at a methionine codon in the context of a NcoI restriction enzyme recognition sequence (CCATGG), which also serves as a good eukaryotic translation initiation sequence ([17]). Plasmid pMS126 starts with methionine and alanine and continues with cysteine 332. Similarly plasmid pMS127 deletes residues 2-383, and continues with valine 384. Plasmid pMS128 deletes residues 2-437 and continues with alanine 438. These constructs were tested for function in the extrachromosomal substrate assay and the results are presented in Table 1. V(D)J recombination was observed with pMS126 and pMS127, but not with pMS128.

The amino-terminal deletions of pMS126 and pMS127 remove entirely a cysteine and histidine-rich region which has been noted ([1]), on the level of primary sequence, to show homology to the zinc-finger DNA-binding domain of the glucocorticoid receptor. These constructs evidently function in the absence of this region, at levels approaching the natural protein. A similar behavior has been reported for a related series of mutants ([8]). A contrasting result was obtained from a construct which specifically mutates three residues of the same region. Plasmid

Table 1. RAG-1 expression plasmids and recombination activities

Plasmid	RAG-1 sequence	tag	(+) butyrate # screened	% Rec	<ul><li>(-) butyrate</li><li># screened</li></ul>	% Rec
pJH548 1 – 1040			7500K	0.41 (15)	820K	3.5 (2)
pMS106	1-1009,V		430K	0.85 (3)		
pMS108	1-994		4000K	< 0.0001 (5)		
pMS119A	1-1008	1	75K	0.59		
pMS119C	1-1008	3	80K	0.46		
pMS122	1-1008, Y994F, Y998F3		270K	0.57 (2)		
pMS124	1 - 1008, C293S,					
-	H307L, C313S	3			630K	0.05 (4)
pMS126	MA, 332-1008	3			100K	0.43
pMS127	M, 384-1008	3			100K	2.2
pMS128	M, 438-1008	3			120K	< 0.001
pMS127B	M, $384 - 1008 + AH_9$	3			35K	2.2
pMS129	pMS127B, ADKEEG 419 VD				19K	< 0.005
pMS130	" , AEKVLL 506 VD				58K	< 0.002
pMS131	", VDEYPV 545 VD				1 <b>9K</b>	< 0.005
pMS132	", SEKLGS 606 VD				50K	0.64
pMS133	", AEREAM 677 VD				36K	< 0.003
pMS134	", LEASQN 735 VD				25K	< 0.004
pMS135	", IETVPS 785 VD				20K	< 0.005
pMS136	", QETVDA 860 VD				20K	< 0.005
pMS137	", AELLST 917 VD				26K	< 0.004
pMS138	", SEGNES 958 VD				25K	< 0.004

The plasmids encode the mouse RAG-1 amino acid sequences listed with alterations given in one letter code. The copy number of the carboxy terminal epitope tag, where present, (see text) is indicated under 'tag'. The percent recombination is the average of duplicates performed within each experiment, and reflects true signal junction positive recombinants as tested by oligonucleotide hybridization. The number of separate experimental repetitions is shown in parenthesis when greater than one. The (-) butyrate experiments were performed using NIH3T3 fibroblasts or the derivative 3TGR cell line. The (+) butyrate experiments were performed using 3TGR exclusively. Also listed for each mutant is the approximate number of recovered pIH200-derived plasmids screened for recombination (K represents thousands). 'H9' is a sequence of nine histidines. Specific mutations are listed such that the original sequence, to the left of the number, is replaced by the sequence to the right. For example, in pMS129, the six residues ADKEEG starting at 419 are replaced by VD.



Figure 2. Protein stably assessed by immunoprecipitation of epitope-tagged RAG-1 and CAT proteins. Cells were pulse labeled with <sup>35</sup>S methionine and chased for the indicated time (minutes). The RAG-1 and CATE3 protein bands are marked with arrows. Positions of size markers in kD are indicated adjacent to the photographs. A. Total cellular extracts from 1.8 cells transfected with pMS119C and pCATE3. B. Cytoplasmic and nuclear fractions of 1.8 cells transfected with deletion mutant pMS127B and pCATE3. C. Graph of degradation time course. RAG-1 signal was quantitated and normalized to CAT signal at each time point.

pMS124 contains three point mutations in the background of pMS119C which replace cysteine 293 with serine, histidine 307 with leucine, and cysteine 313 with serine. In the extrachromosomal substrate assay, correct signal joint recombinants were detectable at a level distinctly above background, but only about 1% that of the parent plasmid.

# **Internal deletions**

The combination of C-terminal and N-terminal deletions defines a core, 60% of the original length of RAG-1, that could not be further truncated from the outside without losing function. Ten additional small deletions were individually constructed and tested to probe the susceptibility of this region to local alteration. In each case a unique Sall restriction enzyme site (encoding the amino acids valine, aspartic acid) was added to the sequence replacing two residues chosen so that the changes would be fairly conservative. Each insertion was followed by a deletion of sequence encoding four amino acids from the parent. The parent plasmid for this series was pMS127B, derived from pMS127 as described above, with the addition of an additional ten amino acid motif (alanine, histidine9) prior to the epitope tag at the Cterminus. The results of the extrachromosomal substrate assays performed using these constructs are presented in Table 1. The parent construct gave recombinants at the level of the unmodified



Figure 3. Immunoprecipitation of epitope tagged RAG-1 and CAT proteins from cytoplasmic and nuclear fractions. Cells were cotransfected with pCATE3 and one RAG-1 expression plasmid. Lane numbers refer to plasmids pMS124, pMS131 and pMS133 respectively. The letters A and B indicate the positions of the full length and deletion RAG-1 proteins. CAT indicates the CATE3 protein. Positions of size markers in kD are indicated adjacent to the photographs.

protein. Nine of the remaining constructs do not generate detectable recombinants, but one plasmid, pMS132, yielded recombinants at roughly 30% of the level of the natural gene. In this construction serine 606 and glutamic acid 607 are replaced with valine and aspartic acid, and residues 608-611 are deleted.

#### Nuclear transport and degradation of RAG-1 protein

The combination of epitope tag and high affinity monoclonal antibody made it possible to directly explore the expression of the RAG-1 protein. The same expression plasmids used for the functional studies were used to produce protein in detectable amounts. We chose to deliver the expression plasmids to the pre-B cell line 1.8, which is intrinsically active in V(D)J recombination and therefore likely to process the RAG-1 protein in a manner most reflecting its normal environment. Metabolic labeling with <sup>35</sup>S-methionine, followed by immunoprecipitation, provides a sensitive assay for detecting the protein. In these experiments the CAT protein (expressed from the cotransfected plasmid pCATE3) was modified to carry the same reiterated epitope tag at its carboxy terminus and served as an internal control for variations in the efficiency of transfection and harvest. The amount of RAG-1 protein could be normalized to the level of CAT protein co-immunoprecipitated by the same anti-epitope antibody. Since the CAT protein is localized in the cytoplasm, it also served as a marker for the subcellular fractionation. The CAT protein itself is stable over a period of hours (not shown).

We find that the RAG-1 protein produced in this way is largely soluble. Pulse-chase experiments were performed with two constructs, pMS119C and pMS127B. In both cases, the RAG-1 protein was chased from cytoplasmic to nuclear pools and rapidly degraded. The half life was estimated to be 15 minutes for the pMS119C protein and 18 minutes for the pMS127B protein. Representative autoradiograms are shown in Figure 2A and B, and a graphic representation of the decay kinetics is shown in

# 5648 Nucleic Acids Research, 1993, Vol. 21, No. 24

humr1 rabr1 musr1 chkr1 xlr1 Cons.	a fpp g v ctp g a lps s itsrmd meva MA-STL-	f e pov tkm LSSAPDEIQH	h d ya ft y P-IKFSEWKF	k pq klr KLFRVRSFEK	t a k.df s tt n n s a a d y e sdksgcin.kdg geva r s eet gsev ynseget y -PEE-QKEK88-EGKP-1	a d ad a d as y pe pg ns stdknit h deevpr e l ypk ty ed lslgsap s LEQSP-VL-XGQKP	vp 11 a evap af v i1 ral 1 i1 kdfmgn t tnfk qqsek s TQPK-H P	dn 9 d 9 a 9 qale dvn 9 nvvdnhet 9 KFSKKFH-D 1	4 3 3 3 4 100
humr1 rabr1 musr1 chkr1 xlr1 Cons.	e arg i g ar i g ss v m iq nta l rleeea i -KDKA-HQ	r f n kq tviqq ANLRHLCRIC	rae lte rsgs vtcyk amqq GMSFK-D-HM	th s RRYPVHGPVD	g lg r s qg a qs f rv de lw se bdi rr v e -KTLLRKKEKATSWPDL :	t ri k rg t r s kt r t q IAKVFRIDVK ADVD91HPTE	r i t n q FCHNCWSIMH R	sa 1 ga 1 sshsq 1 nt 1 i niss 1 KFSPCEV 2	194 193 193 193 194 194
humr1 rabr1 musr1 chkr1 xlr1 Cons.	v kv v s l gav YFPRN-TMEN	t a t g san ev t san yv HPH-PSCDIC	n r. y r. f h. h ps. v heskpw krk -TARGLKR	sl l rnh sqp hg sap l phkm KQPWVQLS	dq qaq ra dq qaq rv nhd rkt rviiaer vn gi nqv r rgpefvkksktssgns : KKLKTVLARR-RK-R-	s d ka te ma vselks kn et knr igwknm afn gmkdsk QARI-SK-VM K-I-NCSKIH	y dnn vl y LSTKLLAVDF P	e 2 2 2 2 2 2 3 4 2 2 3 4 1 2 2 3 4 1 2 2 3 4 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	293 292 290 289 294 300
humr1 rabr1 musr1 chkr1 xlr1 Cons.	lv s Qicshiladp	n v v tr Q VETSCKHLFC	v t sir kyil RICILRCLKV	r q r s w MGSYCPSC-Y (pMS126) r	SV c vt d tv y v a PCFPTDLESPVKSFLNILNS : maC	m k i s p m qd sir vp d ilbg g ll tvsg d i g s L-VKCPA-EC NEEVSLEKYN	i v qfn mkd ik tkg HH-SSHKE 8 (pMS12	if 3 rdtf 3 tl 3 l lynp 3 y va 3 KEVHINK 4 77) mV	391 390 388 389 394 400
humr1 rabr1 musr1 chkr1 xlr1 Cons.	t Gorpronlls	LTRRACKHRL	l i ● r RELK-QVKAF	e i e Adkeeggdvk	m a m k k SVCLTLFLLALRANEHRQA (pMS128) mA	k e a h DELRAIMQGR GSGLQPAVCL	i Airvetfisc s	GYHKMYRTV :	491 490 488 489 494 500
humr1 rabr1 musr1 chkr1 xlr1 Cons.	V t Kaitgrqifq	t PLHALRNA <b>B</b> K	a a i t VLLPGYHPFE	k t r t WQPPLKNVSS	s s r a e n e pli r nq l e -TDVGIIDALSGL-SSVDDY	i t k S a k PVDTIAKRFR YDSALVSALM	e ) 1 ) DMEEDILEGN F	e takn ta vs RSQDLDDYLN (	591 590 588 589 594 600
humr1 rabr1 musr1 chkr1 xlr1 Cons.	C GPFTVVVK <b>ES</b>	t CDGMGDVSEK	V 1 Hosopavper	i k k r n AVRFSFTVM-	as a og a og p a deneriri v syphkngp ri t ITI-H-SQNVKVPEE-KPMS	ELCCKPLCIM LADESDHETI	TAILSPLIAR P	m ( m ( t ( ta 1 ( REAMKSSEL- '	691 690 688 689 694 700
humr1 rabr1 musr1 chkr1 xlr1 Cons.	l p i LEMOGILRTP	s Kpijrotgid	eklvrevegl	t 1 EASGSVIICT	t a n ICDATRLEASQNLVFHSITR	q i c q m b Shaenlerie VWRSNPihes	d pd Veelrdrvkg v	1 VSAKPFIETV	791 790 788 789 794 800
humri rabri musri chkri xlri Cons.	PSIDALHCDI	t r r GNAABFYKIF	DTEIGEAIKN M J	n n dt 1st k P-ASKEERKR	r l k m n i r R WQATLDKHLRKKMNLKPIMR	d i g d s s mngnfarkim tretveavcj	k kc V C Qa K LIPSEERHEA 1	e k t LRELMDLYLK	891 890 888 889 894 900
humri rabri musri chkri xlri Cons.	MEPVWR88CP	1 1 AKECPESLCQ	y b Tspnsqrfae	LLSTKPKYRY	<b>EGKITNYFHKTLAHVPEIIE</b>	rdgsigamas Egnesonkli	? RRFRKMMARQ :	f f Skcyemedvl	991 990 988 989 994 1000
humr1 rabr1 musr1 chkr1 xlr1 Cons.	KHIMLY <b>T</b> SKY	kt II LQKPMKAHMA A	t . vn . rsq aid. ng vdl LK-SGFTMW-	pqas sqvs e sket pddg s p dnpd eqr . LGDPLG	1043 Y 1042 1040 P 1 n V 1041 s m 1e 1045 B IEDBLESQDSMEF 1051				

Figure 4. The amino acid sequence of RAG-1 of five species (human [1], rabbit [20], mouse [1], chicken [21], and Xenopus, [22]) are aligned above that of a consensus sequence. Dots in the five sequences are spaces introduced to maximize alignment. Dashes in the consensus sequence are positions where no consensus was obtained. The locations of the start points of three of the deletion constructs are indicated beneath the consensus. The asterisk below mouse position 1009 shows the site of epitope addition in some of the constructs.

Figure 2C. In the figure, the total cellular content of the deleted form of RAG-1 from plasmid pMS127B was summed from the two gels which separately analyzed the cytoplasmic and nuclear pools obtained from the same labeled cell sample. In the case of pMS127B, when separate nuclear and cytoplasmic fractions are analyzed, it is apparent that the cytoplasmic pool decreases more rapidly than the nuclear, as protein is transported to the nucleus. Steady state levels in the nucleus are obtained within the 30 minute labeling period, demonstrating that the transport process is quite rapid. The epitope-tagged RAG-1 proteins of all the other deletion mutants were similarly examined for production and subcellular localization by pulse labeling, with no chase. In all cases, soluble protein was produced and showed specific partitioning to the nuclear fraction (figure 3A and B show a sample of the data). The relative abundance in cytoplasmic and nuclear fractions of each of the proteins tested was analyzed by phosphorimage analysis of the autoradiograms. While some variation between mutants was evident in the overall abundance and in the fraction partitioning to the nucleus, the amounts were generally within a factor of two, and the differences are not considered to be significant in this analysis (data not shown).

Only the triple point mutation construct pMS124 gives a significantly different result. While soluble protein is produced and distributes similarly to the other versions of the protein, the radioactive incorporation is only 12% of that of pMS127B. When corrected for different methionine content, this indicates a tenfold difference in molar levels of the two proteins. This reduction may be one factor in explaining the low recombination activity observed with pMS124 (70-fold lower than its parent pMS119C). However, these mutations may also lead to a disruption of the RAG-1 protein structure, or of its assembly into higher-order complexes, and thus lead to a more severe defect than deletion of this region. Silver *et al.* ([8]) similarly noticed a sharp drop in the amount of protein detectable by immunoblot analysis when point mutations were introduced into the cysteine and histidine-rich region.

## Site-specific alteration of the proposed topoisomerase homology region

Wang et al. ([18]) described a homology in amino acid sequence between the yeast gene HPR1 and the portion of RAG-1 from residue 472 onward. HPR1, in turn, had some homology to topoisomerases, particularly in the neighborhood of the tyrosine which corresponded to mouse RAG-1 residue 998, which was suggested as a potential topoisomerase active site ([18]). However, site specific mutation of this residue (as well as tyrosine 994) in plasmid pMS122 did not interfere with the ability of RA-G-1 to function in recombination. A similar result has been obtained in two other studies ([19], [8]); our test differs only in that the two specific amino acid replacements occur in the context of the carboxy-terminal deletion already introduced in the precursor plasmid pMS119C. While these results effectively rule out the participation of either of these tyrosines in forming a topoisomerase-like covalent bond to DNA, the surrounding region does appear to be essential for RAG-1 function, because a deletion (pMS108) from the C-terminus to residue 995 is not tolerated. Whatever the relationship of RAG-1 to HPR1 may be, it does not seem to involve a shared topoisomerase function.

Because one trivial explanation for the absence of recombination activity in pMS108 could be a failure at the level of gene expression, the level of RAG-1 specific RNA was checked. A blot of polyA + RNA obtained from cells transfected with constructs pJH548, pMS106 or pMS108 detected transcripts of the predicted size from all three plasmids in comparable amounts (data not shown). Therefore it is most likely that the deletion of pMS108 interferes with the function of RAG-1 at the protein level.

#### Correlation of mutations with a multiple sequence alignment

The end points for the C-terminal and N-terminal deletions studied in this report were selected in part by considering the alignment of the predicted RAG-1 translation products of five animal species (references in Figure 4). Figure 4 shows a consensus sequence, together with the individual differences. Inspection of the figure reveals regions with frequent amino acid variation, including occasional insertions, and other regions of striking sequence conservation. We note that the highly conserved region starting around position 384 of the mouse sequence correlates well with the protein's functional core as determined by the recombination assay, because plasmid pMS127, which encodes mouse RAG-1 sequence from amino acid 384 to 1010, is fully active. Further deletions from the amino terminal, as represented by pMS128 or pMS129, are not compatible with function. Deletion of thirteen amino acid residues from the C-terminal border of the conserved region through residue 995, as demonstrated by plasmid pMS108, also renders this construct nonfunctional. Within the defined core, deletions seem much less tolerated. Only one of the ten short deletions internal to this core supported recombination.

This study extends previous work in demonstrating that the amino terminal 383 amino acids can be deleted (plasmid pMS127B) without appreciably altering the activity of RAG-1 in extrachromosomal recombination. Furthermore, this deletion protein is transported to the nuclear fraction and exhibits a degradation rate similar to that of the almost full-length pMS119C. The functional deletions reported here remove from the RAG-1 protein all of the structures that have been proposed as significant on the basis of homology. These results do not yet allow a decision as to whether RAG-1 is an indirect activator or a direct participant in V(D)J recombination.

These experiments have tested the recombinational proficiency of RAG-1 mutants only in the context of forming signal joints in an artificial substrate. It is possible that recombination of the antigen receptor loci in their natural setting could require other elements of RAG-1 structure.

# ACKNOWLEDGEMENTS

We are grateful to David Schatz for providing us with the cell line 3TGR, in which many of these experiments were performed. We also thank the members of the Laboratory of Molecular Biology, NIDDK for encouragement and conviviality in the work place.

# REFERENCES

- 1. Schatz, D.G., Oettinger, M.A. and Baltimore, D. (1989) The V(D)J recombination activating gene, RAG-1. *Cell*, 59, 1035-1048.
- 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.
- Gellert, M. (1992) Molecular analysis of V(D)J recombination. Ann. Rev. Genet., 22, 425-446.
- Schatz, D.G., Oettinger, M.A. and Schlissel, M.S. (1989) V(D)J recombination: molecular biology and regulation. *Ann. Rev. Immunol.*, 10, 359-383.
- 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.
- Shinkai, Y., Rathbun, G., Lam, K.-P., Oltz, E.M., Stewart, V., Mendelsohn, M., Charron, J., Datta, M., Young, F., Stall, A.M., Alt, F.W. (1992) RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell*, 68, 855-867.
- Hesse, J.E., Lieber, M.R., Gellert, M. and Mizuuchi, K. (1987) Extrachromosomal DNA substrates in pre-B cells undergo inversion or deletion at immunoglobulin V-(D)-J joining signals. *Cell*, 49, 775-783.

- Silver, D.P., Spanopolou, E., Mulligan, R.C. and Baltimore, D. (1993) Dispensable sequence motifs in the RAG-1 and RAG-2 genes for plasmid V(D)J recombination. *Proc. Natl. Acad. Sci. U.S.A.*, 90, 6100-6104.
- Alt, F. W., Yancopoulos, G. D., Blackwell, T. K., Wood, C., Thomas, E., Boss, M., Coffman, R., Rosenberg, N., Tonegawa, S., Baltimore, D. (1984) Ordered rearrangement of immunoglobulin heavy chain variable region segments. *EMBO J.*, 3, 1209-1219.
- Seed, B. (1987) An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2. *Nature*, 329, 840-842.
  Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982) Recombinant
- Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982) Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.*, 2, 1044-1051.
- Schatz, D.G. and Baltimore, D. (1988) Stable expression of immunoglobulin gene V(D)J recombinase activity by gene transfer into 3T3 fibroblasts. *Cell*, 53, 107-115.
- Evan, G.I., Lewis, G.K., Ramsay, G. and Bishop, J.M. (1985) Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. *Mol. Cell. Biol.*, 5, 3610-6.
- Hirt, B. (1967) Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol., 26, 365-369.
- Goldstein, S., Fordis, C.M. and Howard, B.H. (1989) Enhanced transfection efficiency and improved cell survival after electroporation of G2/Msynchronized cells and treatment with sodium butyrate. *Nucleic Acids Res.*, 17, 3959-3971.
- Wilkinson, G.W. and Akrigg, A. (1992) Constitutive and enhanced expression from the CMV major IE promoter in a defective adenovirus vector. *Nucleic* Acids Res., 20, 2233-2239.
- 17. Kozak, M. (1991) Structural features in eukaryotic mRNAs that modulate the initiation of translation. J. Biol. Chem., 266, 19867-19870.
- Wang, J.C., Caron, P.R. and Kim, R.A. (1990) The role of DNA topoisomerases in recombination and genome stability: a double-edged sword? *Cell*, 62, 403-406.
- Kallenbach, S., Brinkmann, T. and Rougeon, F. (1993) RAG-1-a topoisomerase? Int Immunol., 5, 231-232.
- Fuschiotti, P., Harindranath, N., Mage, R.G., McCormack, W.T., Dhanarajan, P. and Roux, K.H. (1993) Recombination activating genes -1 and -2 of the rabbit: cloning and characterization of germline and expressed genes. *Molecular Immunol.*, 30, 1021-1032.
  Carlson, L.M., Oettinger, M.A., Schatz, D.G., Masteller, E.L., Hurley,
- Carlson, L.M., Oettinger, M.A., Schatz, D.G., Masteller, E.L., Hurley, E.A., McCormack, W.T.Baltimore, D. and Thompson, C.B. (1991) Selective expression of RAG-2 in chicken B cells undergoing immunoglobulin gene conversion. *Cell*, 64, 201–208.
- Greenhalgh, P.H., Olesen, C.E.M. and Steiner, L.A. (1993) J. Immunol., In press.