Published in final edited form as:
Nature. 2015 February 26; 518(7540): 507–511. doi:10.1038/nature14174.

Crystal Structure of the V(D)J Recombinase RAG1-RAG2

Min-Sung Kim[1], **Mikalai Lapkouski**[1,¶], **Wei Yang**[1,*], and **Martin Gellert**[1,*]

[1]Laboratory of Molecular Biology, NIDDK, NIH, Bethesda, MD 20892, USA

Summary

V(D)J recombination in the vertebrate immune system generates a highly diverse population of immunoglobulins and T cell receptors by combinatorial joining of segments of coding DNA. The RAG1-RAG2 protein complex (RAG1/2) initiates this site-specific recombination by cutting DNA at specific sites flanking the coding segments. We report here the crystal structure of the RAG1/2 complex at 3.2Å resolution. The 230 kDa RAG1/2 heterotetramer is Y-shaped, with the N-terminal domains of the two RAG1 chains forming an intertwined stalk. Each RAG1/2 heterodimer composes one arm of the "Y" with the active site in the center and RAG2 at its tip. The RAG1/2 structure rationalizes more than 60 mutations identified in immunodeficient patients, as well as a large body of genetic and biochemical data. The architectural similarity between RAG1 and the hairpin-forming transposases Hermes and Tn5 suggests the evolutionary conservation of these DNA rearrangements.

Introduction

To combat the great range of possible infectious agents, the vertebrate immune system deploys a highly diverse population of immunoglobulins and T cell receptors. In many species this diversity is generated by V(D)J recombination [1]. By combinatorial joining of segments of coding sequence, V(D)J recombination is capable of assembling millions of different functional immunoglobulin and T cell receptor genes [1,2]. This recombination is initiated by DNA double strand breaks produced by the RAG1-RAG2 recombinase, at sites flanked by specific recombination signal sequences (RSS). The RSSs are of two types, with either 12 or 23 non-conserved nucleotides between conserved heptamer and nonamer modules; one RSS of each type is strictly required for recombination [2]. The two RSS varieties are partitioned so as to focus recombination on V to J, or V to D to J, joining. RAG1 and RAG2 are the only lymphoid-specific factors involved in V(D)J recombination [3,4], while the resulting hairpinned coding ends are processed by general repair factors of the non-homologous end-joining pathway [5,6]

[*]Co-corresponding authors: Wei Yang, Tel: (301) 402-4645, wei.yang@nih.gov. Martin Gellert, Tel: (301) 451-8168, gellert@helix.nih.gov.
[¶]Current address: Department of Cell and Molecular Biology, Karolinska Institute, 171, 77 Stockholm, Sweden and Centre for Structural Systems Biology, DESY, 22607 Hamburg, Germany

Author contributions

M-S.K. prepared SEC complexes and carried out crystallography. M.L. developed protocols for protein expression and initial crystallization. W.Y. and M.G. designed the project, and M-S.K., W.Y. and M.G. prepared the manuscript.

Atomic coordinates and structure factors have been deposited with the Protein Data Bank with accession code 4WWX.

The authors declare no competing financial interests.

Since the identification of the RAG1 and RAG2 genes [7,8], RSS-dependent DNA cleavage by purified RAG1/2 has been reconstituted [9]. RAG1 and RAG2, of 1040 and 527 residues, cooperate in all their known activities. The catalytic core, regulatory regions, active site residues, DNA-binding domains, two zinc-binding motifs, and some aspects of the interface of RAG1 and RAG2 have been characterized [3,4]. It was also found that RAG1/2 can function in vitro as a transposase [10,11], inserting RSS-terminated DNA into a second DNA molecule. Moreover a large number of human mutations in both RAG proteins that cause severe combined immunodeficiency (SCID) or a milder form known as Omenn syndrome have been identified [12,13].

Biochemical and functional studies have shown that portions of RAG1 and RAG2 can be deleted, and the "core" proteins, residues 384–1008 of RAG1 and 1–387 of RAG2, retain targeted cleavage activity in vitro and recombination activity (though not fully regulated) in cells[14–17]. An earlier low-resolution electron microscopic study of the core complex, containing two subunits each of RAG1 and RAG2 bound to a 12 and 23RSS DNA pair, revealed the overall shape and localization of RAG proteins [18]. Here we report the 3.2Å crystal structure of the RAG1/2 heterotetramer and its implications for V(D)J recombination.

SEC complex and structure determination

The catalytic cores of mouse RAG1 (384–1008aa) and RAG2 (1–387aa) with maltose binding protein (MBP) fused to their N-termini were expressed in HEK293T cells and readily purified (Methods). RAG1/2 was assembled with pre-cleaved 12RSS and 23RSS DNAs in the presence of HMGB1 to form a signal-end complex (SEC) [19], and the purified SEC after removal of the cleaved MBP tags and HMGB1 (Fig. 1a–b) was homogeneous and active in strand transfer (Extended Data Figure 1).

Crystals were grown over a period of 2–4 weeks (Methods). For phase determination, methionines in the RAG1/2 proteins were substituted by selenomethionine to a level of 40% (Methods). Single-wavelength anomalous diffraction (SAD) datasets of high redundancy were collected at the Se absorption peak from six crystals. Fifty-four of 58 selenium sites were located, together with two Zn^{2+} atoms (one in each RAG1). The electron density map using all SAD data, nominally at 3.7Å, was superior to that calculated using only the two best sets according to anomalous correlation coefficient [20] (Fig. 1c–d, Extended Data Fig. 1a). The heterotetramer of RAG1/2 recombinase was readily traceable (Extended Data Fig. 1b). Although 12RSS and 23RSS DNAs were included in the SEC complex and were also present in dissolved crystals (Extended Data Fig. 1c–d), DNA was not found in the electron density map. Only the four protein chains, with residues 391 to 1008 of RAG1 and 2–350 of RAG2, were modeled and refined to 3.2Å (Extended Data Table 1). The C-terminal 37 residues of RAG2 are disordered. In fact, RAG2 (1–351) forms active heterotetramers with RAG1 for RSS DNA cleavage in vitro (Extended Data Fig. 2), and supports V(D)J recombination in cells [21].

Architecture of RAG1/2

The RAG1/2 crystal structure is remarkably similar to the low-resolution model generated from two-dimensional averaging of negatively-stained EM images [18]. It is Y-shaped (125Åx150Åx90Å) with the RAG1 dimer forming the bulk and RAG2 situated at the tip of each arm (Fig. 2). There is an oval-shaped gap (40Åx60Å) separating the two RAG1/2 heterodimers (Fig. 2). RAG1 is elongated (100Å) and composed of seven structural modules (Fig. 3a). The N-terminal nonamer-binding domain (NBD), which superimposes well with the structure determined previously [22], forms a domain-swapped and intertwined dimer. The following dimerization and DNA binding domain (DDBD) is connected to the NBD by a flexible linker, and the last helix of the three-helix C-terminal domain (CTD) folds back to complete the DDBD (Fig. 3b), which may be why previous domain dissections failed to isolate this structural entity. Three conserved carboxylates in RAG1 (D600, D708 and E962) have been identified as essential for catalysis [23–25], but it was not clear how the catalytic domain would fold owing to their large separation in the amino-acid sequence. Our structure shows that RAG1 adopts an RNase H fold with an elongated central four-stranded β-sheet, similar to other DDE transposases (named after the three catalytic carboxylates, which are all within the same RAG1 subunit) (Fig. 3) [26–28]. Following the DDBD, the extended pre-RNase H (preR) and catalytic RNase H (RNH) domains further lengthen RAG1 to ~100Å. Each active site is located in the middle of a Y arm, and the two are about 45Å apart (Fig. 2). Two domains intervene between D708 and E962, ZnC2, which protrudes towards RAG2, and the highly helical ZnH2, which increases the third dimension of RAG1 from 25Å to 65Å and eventually brings E962 back to the catalytic center (Fig. 3a–c). Two regions for Zn^{2+} binding were previously identified [29,30]. The first (C727, C730) and second (H937 and H942), despite being far apart in the sequence, form one zinc-binding site (Fig. 3c–d) that juxtaposes the catalytic center, interface with RAG2 (ZnC2), and DNA binding (DDBD and ZnH2) domains.

RAG2 is folded into a six-bladed β-propeller, or Kelch-repeat structure, as predicted [31,32] (Extended Data Fig. 3). The first N-terminal β-strand belongs to the sixth 4-stranded β-blade and ties the first and last Kelch repeats together. Compared with other β-propeller structures, the six blades of the doughnut-shaped RAG2 are more distorted in planarity and spacing between blades (Extended Data Fig. 3). These distortions are most notable at the interface with RAG1 (Fig. 4a). As usual for β propeller proteins, one face of RAG2 has extended loops. Many loops, particularly those connecting adjacent blades or the middle two strands of each blade, are involved in interacting with the preR, RNH and ZnC2 domains of RAG1 (Fig. 4b–c). The interface between RAG1 and RAG2 is highly conserved from fish to humans, encompassing both polar and hydrophobic interactions and dovetailed with ridges and canyons (Fig. 4). Interestingly, RAG2 contacts RAG1 near the active site[33], including E607 and V615 (connected by a disordered loop) and E719 to V724 contacted by the long RAG2 loop 335–339 and residue R39, respectively (Fig. 2, 4c). It is likely that in the presence of DNA substrate, RAG2 assists in formation of the catalytic site and DNA cleavage.

SCID and Omenn syndrome mutations

Over 60 missense mutations leading to SCID or Omenn syndrome (OS), a milder type of immunodeficiency, due to defective DNA processing in V(D)J recombination, have been mapped to the catalytic cores of RAG1 and RAG2 [12,13]. The disease-involved residues are identical between human and mouse RAG proteins, except for M435 in human RAG1 being replaced by L432 in mouse. Residue numbers in mouse and human RAG1 differ by three (Extended Data Table 2). For consistency, all residues are numbered here according to the mouse protein that we are studying. The SCID and OS mutations fall into four classes. The first class of mutations clearly destabilizes the tertiary structure of RAG1/2. For example, mutations of the zinc-binding site, C727E and the adjacent L729F (Fig. 3d), would perturb the structure of the ZnC2, ZnH2 and adjoining domains. Three immunodeficiency mutations, W893R, Y909C and I953R, probably destabilize the hydrophobic core next to the zinc-binding site (Fig. 3d), reflecting the importance of this region in the overall structure of RAG1/2. A number of laboratory-generated mutations with loss of function also are likely to destabilize the structure of certain domains in RAG1, for example W893A.

The second class includes polar residues exposed to solvent and concentrated in two areas likely for DNA binding. Seven out of nine immunodeficiency mutations found in the DDBD and CTD domains are at a cluster of polar residues (Arg, Lys, Ser or Gln) (Fig. 3e). Systematic mutation of positively charged residues in RAG1 [34] also showed that K966, R969, R977 and H990 in CTD contribute to DNA binding and cleavage. In parallel, twelve SCID/OS mutations in the NBD domain (Fig. 3b, Extended Data Fig. 4), which overlap with many highly conserved non-polar and polar residues, are involved in structural integrity and sequence-specific binding to the nonamer DNA [22].

The third class of mutations is clustered around the active site (Fig. 3f). Some may alter the structure of the catalytic center (S598P, C599W, A619P, R696Q/W, G706D), and others may change its DNA-binding properties (E666G, R621C/H, R713W). It is not surprising that engineered mutations of conserved polar residues surrounding the active site (E597, E709, D792, H795 and E959) lead to defects in DNA cleavage [25,34].

The last class of SCID/OS mutations is located at the interface of RAG1 and RAG2. Four of six disease mutations in the RAG2 core are concentrated at the subunit interface (Fig. 4). G35V, R39G, and C41W are at the interface with the ZnC2 and RNH domains of RAG1. R229 forms salt bridges with D546 (preR) of RAG1, which has been accurately mapped to the interface with RAG2 [35]. G95R is at the base of a long loop that reaches to the RNH domain and sandwiches E666 (which itself is mutated in SCID patients) with G35 (Fig. 4a). The importance of the subunit interface is also evident in the disease mutations located on the RAG1 side (R556S, R558C/H, E666G and R773Q), and additional mutations identified in laboratories (E719Q and R773A) [25,34] (Extended Data Table 3)

A model of the RAG1/2—RSS DNA complex

Structures of five DDE transposases complexed with DNA have been reported[36–40]. They can be segregated into two groups based on the drastically different DNA orientations relative to the catalytic dimer (Extended Data Fig. 5, 6). Hermes (a member of the hAT

transposase family) and Tn5, both of which generate a hairpin intermediate during DNA processing, belong to the same group, and both contain an α-helix extending from one catalytic center (by contributing the last catalytic residue Glu) to contact the DNA bound to the other subunit (Extended Data Fig. 5). Very similar architectural features are found in the RAG1 dimer. After superimposition of the RNH domain of RAG1 and Hermes (Fig. 5a), not only does the 16bp DNA co-crystallized with Hermes fit into the active site of RAG1, but the second DNA of Hermes is close to the other RAG1 catalytic center (Fig. 5b). Remarkably, the two DNAs modeled in RAG1 are connected by the α-helix (964–975aa) that immediately follows E962 in the CTD, just as is found in Hermes and Tn5. The first 7 bp of the DNA may mimic the heptamer of each RSS (Fig. 5b). None of the aromatic residues previously suggested as functioning in DNA hairpin formation [41,42] is situated near the active site. G851 and N852 in the ZnH2 domain of RAG1 appear to replace W319 in the equivalent helical insertion domain of Hermes that stacks on the 3′-end DNA base (Fig. 5b). In corroboration of this model, mutations of N852 or residues 970–978 in the CTD are implicated in Omenn syndrome (Fig. 3e, 5b).

The rest of the RSS DNA can also be modeled based on the published crystal structure of an isolated NBD bound to a 12bp nonamer DNA [22] (Extended Data Fig. 4). From DDBD to CTD, the two halves of the RAG1/2 tetramer are rather symmetric (Fig. 5c), but the NBD domains are not related by the same dyad. Although perfectly symmetric internally, the intertwined NBDs are tilted relative to the rest of the protein. As a result, the two nonamer DNAs bound to the NBDs would be oriented differently to the two catalytic centers. One is very close to the 16 bp DNA modeled into the RNase H domain (Fig. 5c). Notably, the sum of the two DNA segments is about 28 bp, the total length of 12RSS DNA. In contrast, the other pair of DNA segments is separated by ~30Å, which may mimic the 23RSS with its additional 11 bp connecting the nonamer and heptamer (SI video). In this model, the nonamer and heptamer ends of each RSS interact with different RAG1 subunits in a trans configuration, as mutational study has suggested [27]. A sharp kink is unavoidable in each RSS DNA as modeled (Fig. 5c), and HMGB1 could stabilize such kinks to facilitate the gene rearrangement [43,44].

The surface of RAG1/2 traversed by the modeled DNAs is both highly positively charged and highly conserved (Extended Data Fig. 7). The only exception is the NBD, which is not as highly charged as the DDBD and ZnH2 regions. This may correlate with the sequence-specific recognition of the nonamer. Beyond the regions that are modeled to bind RSS DNAs, extensive surface areas along the rim of the Y arms from RAG2 to ZnH2 are positively charged and partially conserved (SI video). These areas could bind up to 20 bp of coding DNA flanking the RSS. Although 6 bp coding flanks can be slowly cleaved by RAG1/2, efficient cleavage requires more than 15 bp of flanking DNA [45]. Interactions of the coding flanks with the top of the RAG1/2 complex may explain why many mutations in the RAG1 and RAG2 interface impact DNA cleavage[33].

Concluding remarks

The structure of RAG1/2 reveals the architecture of the complex and the composition of its functional sites. It rationalizes the effects of many mutations associated with human immune

deficiencies. Evolutionarily, eukaryotic hAT transposases and RAG1/2 recombinase, which cleave DNA duplex in two steps and leave a hairpin on the flanking DNA, are thought to be rather different from the bacterial transposases that cleave DNA in three steps and form a hairpin intermediate on the recognition DNA (Fig. 5d). The similar enzyme-substrate association found in Hermes and Tn5 and their structural relationship with RAG1 has led us to propose that the two DNA recombination processes are identical in mechanism and configuration, differing only in the nucleophiles used at each step, a water molecule versus a 3′-OH.

METHODS

Protein expression and purification

Mouse core RAG1 (384–1008aa) and RAG2 (1–387aa) were cloned into the pLEXm-based [46] mammalian expression vector, modified with an N-terminal His_8 tag followed by MBP tag and a Prescission cleavage site (LEVLFQ/GP) (with "/" indicating the cleavage site). The N-terminal Met of RAG2 was mutated to Val (M1V) during cloning. To express the RAG1/2 complex in HEK293T cells, 500 μg of each of the RAG1 and RAG2 expression plasmids were mixed with 4 mg of polyethylenimine (Polysciences) in 35 mL of Hybridoma medium (Invitrogen) to transfect 1 L of HEK293T cells grown in suspension culture in Freestyle 293 medium (Invitrogen), supplemented with 1 % FBS when the cell density reached 1.5 million per milliliter. Four days after transfection, cells were harvested and stored at −80 °C. Cell paste (~8 g from 1L culture) was resuspended in 50 mL of lysis buffer containing 20 mM HEPES (pH 7.3), 1 M KCl, 1 mM Tris (2-carboxyethyl) phosphine (TCEP) (pH 7.0), 1 mM EDTA and protease inhibitor cocktail (Roche), and lysed by sonication. After centrifuging at 35,000 rpm for 1 hr, the clarified lysate was mixed with 5 mL of amylose resin (NEB), which was pre-equilibrated with the lysis buffer and incubated with rotation for 1 h. After pouring the resin into a column and washing thoroughly with 200 resin volumes of the lysis buffer, the RAG1/2 protein was eluted by an amylose elution buffer containing 20 mM HEPES (pH 7.3), 500 mM KCl, 40 mM maltose and 1 mM TCEP. The eluted protein, which consisted mainly of RAG1/2 heterotetramer[47] was concentrated and stored at −80 °C after adding glycerol to 20 % final concentration. In contrast to RAG1 and RAG2 core proteins expressed in insect cells, which required activity-based purification for structural studies [18], RAG1/2 expressed in human cells is highly active. Human HMGB1 (1–163aa) was prepared as reported previously [18,19].

To make the SEC complex, purified RAG1, RAG2, 12RSS and 23RSS DNAs were mixed at 2:2:1:1 molar ratio in the presence of HMGB1 with the buffer of 20 mM HEPES (pH 7.3), 150 mM KCl, 5 mM $CaCl_2$, 1 mM TCEP and HMGB1 and incubated for 1 hr at 37 °C. After removing the MBP tag by addition of PreScission protease (1:100 ratio of protease to substrate) overnight at 4 °C, further purification was performed using gel filtration (Superdex 200, GE Healthcare) in 20 mM HEPES (pH 7.3), 500 mM KCl, 5 mM $CaCl_2$ and 1 mM TCEP, which removed HMGB1 along with free DNA, the MBP tag, and PreScission protease. All purification steps were performed at 4 °C. HMGB1 could be retained with RAG1/2 and RSS DNA if KCl concentration was reduced to below 100 mM (Extended Data Figure 2), but we were unable to crystallize SEC completed with HMGB1. To prepare

selenomethionine (SeMet)-labeled RAG1/2 complex, HEK293T cells were transferred after transfection to methionine-free Freestyle 293 medium (Invitrogen) supplemented with 25 mg/L L-SeMet (Acros Organics) and 1% dialyzed FBS (Invitrogen). 3 days later, cells were collected and protein was purified in the same way as native protein. Mass spectrometry analysis of trypsin digested SeMet-labeled RAG1/2 peptides was performed at the Taplin Mass Spectrometry Facility (taplin.med.harvard.edu). It showed that about 40% of methionines were substituted by SeMet.

Crystallization and data collection

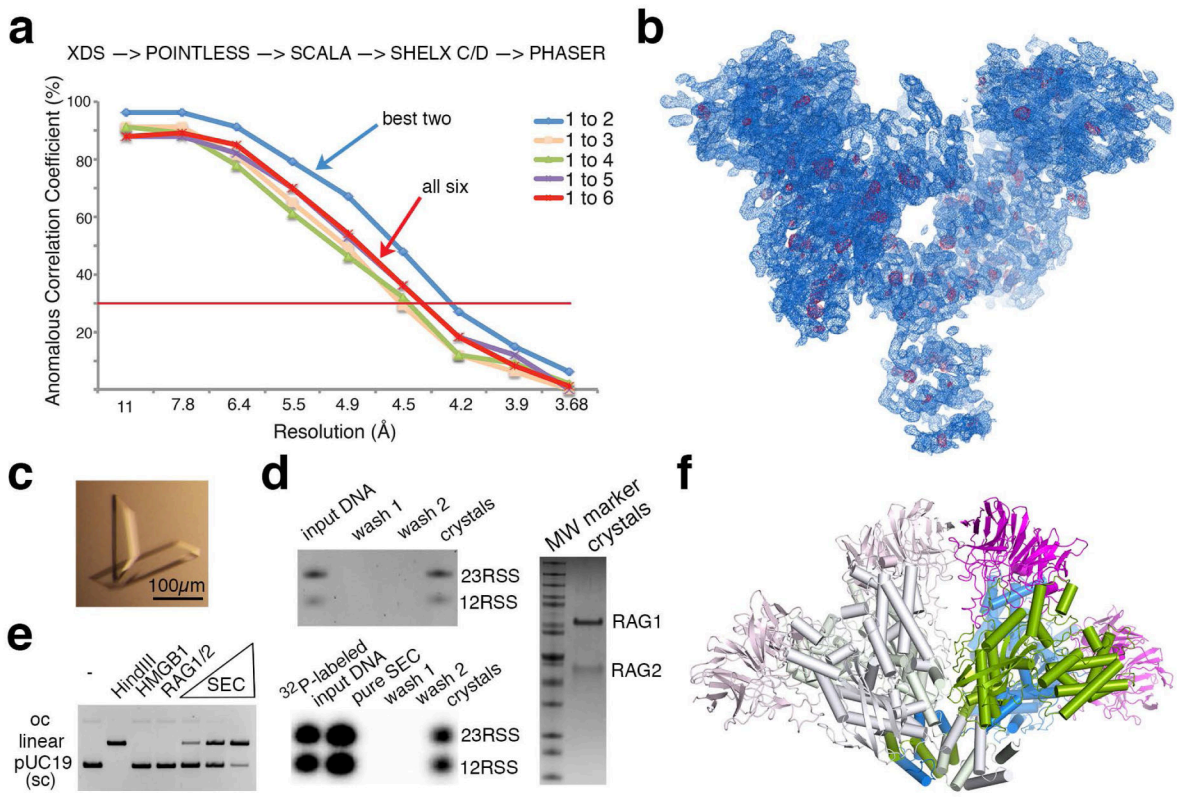
Crystals of the RAG1/2-DNA complexes were grown by the hanging-drop vapor diffusion method at 4°C over 3 weeks. Equal volumes of protein (~5 mg/ml) and reservoir solution containing 100 mM MES (pH 7.1), 10–15% PEG 3350, 200 mM tribasic ammonium citrate (pH 7.0) and 100 mM KCl were mixed in each droplet. Crystals were cryo-protected in reservoir solution supplemented with 25 % ethylene glycol and flash frozen in liquid nitrogen. We were able to crystallize tetrameric RAG1/2 alone as well as RAG1/2 with a single 12RSS or 23RSS, but these crystals were small, and none of them diffracted X-rays as well as the SEC complex. Crystals of SeMet-labeled RAG1/2 complex were grown under similar conditions. Native and SeMet-labeled complex both crystallized in the $C222_1$ space group with two RAG1 and two RAG2 (one RAG1/2 heterotetramer) in each asymmetric unit. Data were collected at 100°K for native and SeMet derivative crystals at beam lines 22ID and 23ID of the Advanced Photon Source (APS) at Argonne National Laboratory. All data were indexed, integrated, and scaled with the XDS package [48] (Extended Data Tables 1).

Structure determination and refinement

Phases were determined by the single-wavelength anomalous diffraction (SAD) method and multi-crystal averaging [20]. SAD data were collected from six SeMet substituted crystals with the best resolution of 3.7Å (Extended Data Table 1). Data were processed according to the published procedure [20] (Extended Data Fig. 1a). Selenium sites were identified using SHELXD [49] and refined with PHASER [50]. Out of the 58 highest anomalous peaks, 54 corresponded to selenium sites, and 2 were Zn^{2+} ions. Phases were improved by density modification using RESOLVE [51] and the overall figure-of-merit was 0.79. The RAG1/2 model was built manually in COOT [52]. Although the experimental electron density map contained breaks in the main chains, and side-chain definition was not perfect, the register of the polypeptide chains was readily determined based on the SeMet sites. This initial model was refined in Phenix [53] and manually improved using COOT. Secondary structure restraints and non-crystallographic two-fold symmetry averaging restraints were used throughout the refinement. The RAG1/2 structure was refined to 3.2Å with R_{work} and R_{free} of 20.6 % and 25.9 %, respectively (Extended Data Table 1). The quality of the structure was validated with MolProbity [54]. 90.7 % of residues are in the favored regions of the Ramachandran plot, 9.3 % in additional allowed regions, and no residue in the disallowed region. The final model contains amino acids 391–1008 of RAG1 and amino acids 2–350 of RAG2, and one Zn^{2+} ion in each RAG1. The N-terminal residues (391–404) are ordered in one RAG1 subunit. Due to poor electron densities, residues 608–616 of RAG1 and 82–88, 242–244, 254–255 and 334–337 of RAG2 were not included in the final model. Crystal

packing of two neighboring RAG1/2 tetramers appears to occlude one nonamer-binding site in each RAG1/2 complex (Extended Data Fig. 1e). No water molecules were added. All structure figures were prepared with PyMOL (www,pymol.org), and sequence conservation analysis was performed using ClustalW [55].

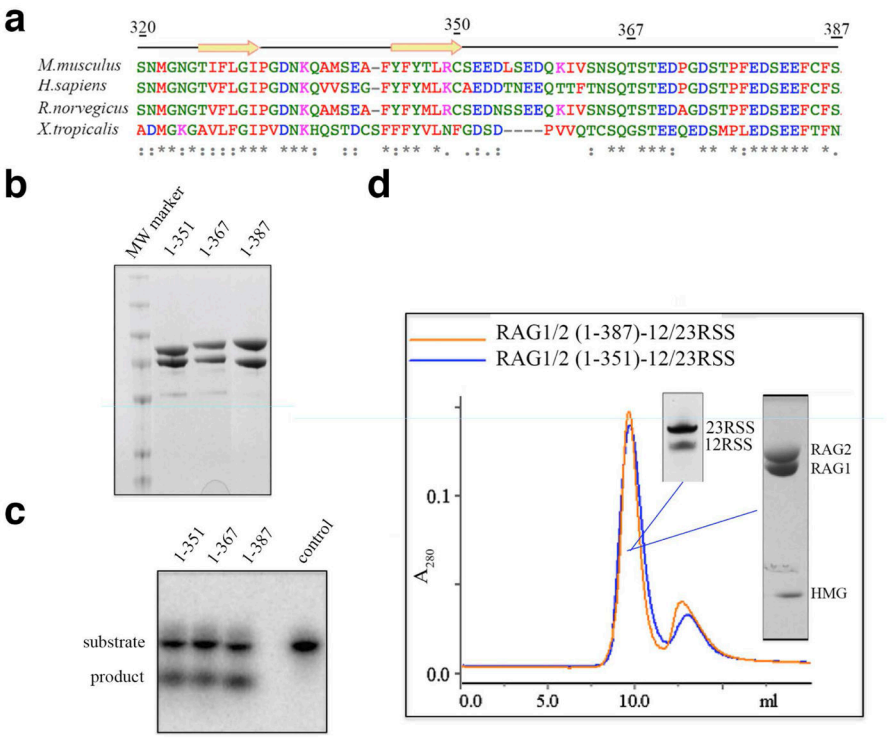
Extended Data



Extended Data Figure 1. Structure determination

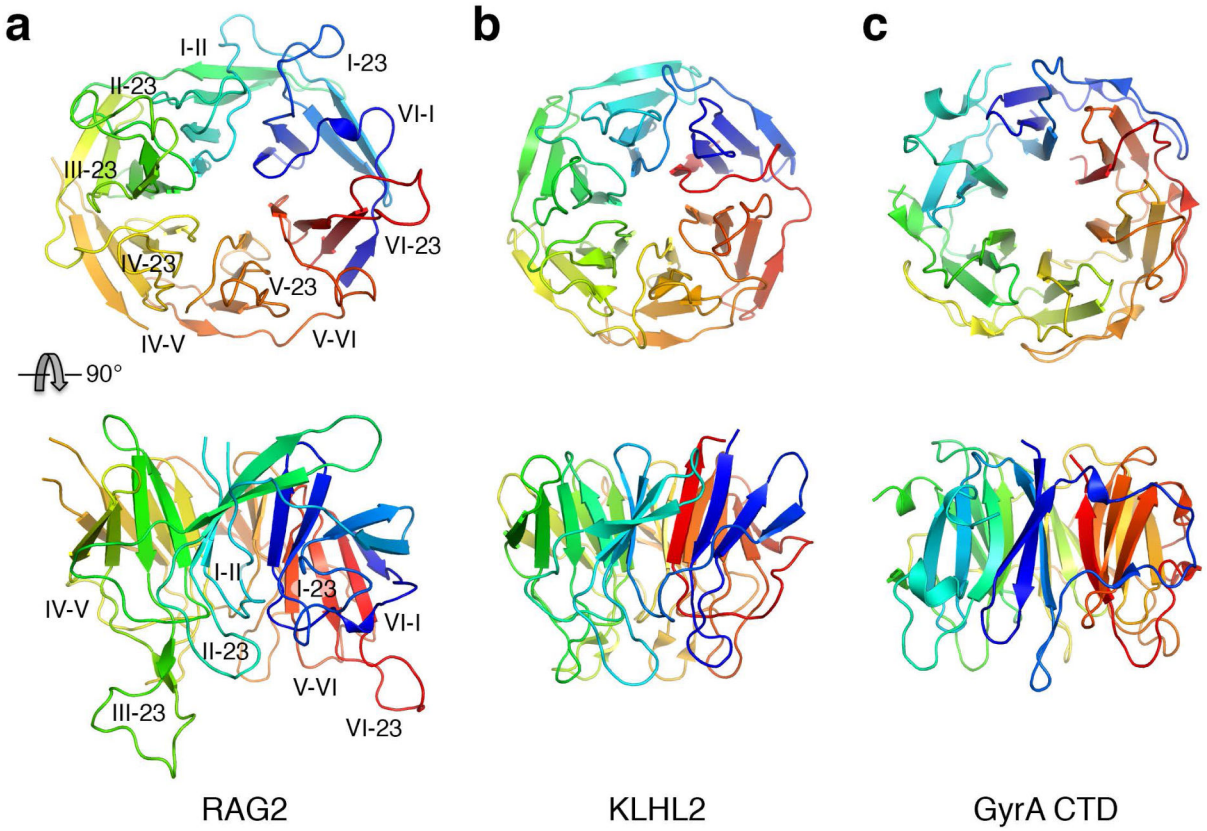
(**a**) The plot of correlation coefficient (CC) of anomalous signal versus resolution. The red line indicates the cutoff of CC=0.3. Merging data from the two best crystals produced a better CC than merging data from all six crystals. The data processing procedure is outlined above the plot [20]. (**b**) The SAD experimental map contoured at 1.3σ showed the content of an asymmetric unit. The anomalous peaks of selenium atoms are shown in red. (**c**) A typical crystal of RAG1/2. (**d**) The content of crystals was examined by protein and DNA denaturing gels after a thorough wash of the crystals and stained by Coomassie Blue and SYBR-Green. To confirm the 1:1 molar ratio of 12 and 23RSS DNA, ^{32}P-labeled input RSS DNAs and those in SEC complexes before and after crystallization are shown beneath the SYBR-Green stained DNA gel. (**e**) Transposition assay of the purified SEC (RAG1/2-12/23RSS DNA complex) used for crystallization. Supercoiled pUC19 (sc, with a small amount of open circle, oc) was the target; it was linearized by HindIII as a control. The SEC (0.25, 0.5 and 1.0 μM) was active in concerted transposition and thus linearizing pUC19. In contrast, RAG1/2 or HMGB1 (0..5μM) each alone was not active. (**f**) Crystal

packing of neighboring RAG1/2 complexes (shown in dark and light colors) occludes one nonamer-binding site in each heterotetramer of RAG1/2.



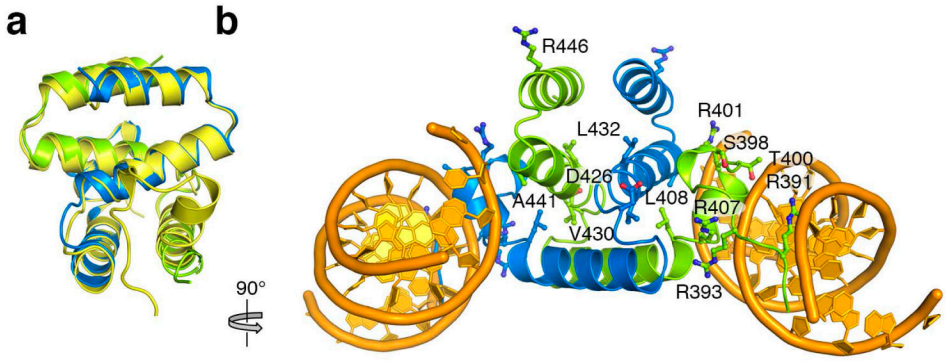
Extended Data Figure 2. RAG2 core fragment (1–351aa) is active

(**a**) Sequence alignment of RAG2 from mouse (320–387aa), human, rat and Xenopus with predicted secondary structures shown above. (**b**) Core RAG2 (1–387) and two further truncated RAG2 variants (1–351 and 1–367) were constructed with a non-cleavable N-terminal MBP tag and co-expressed with the tag-less core RAG1. The Coomassie Blue R-250 stained SDS gel shows the purified RAG1/RAG2 complexes. (**c**) Purified RAG1/2 complexes with truncated RAG2 variants are equally active in cleaving a ^{32}P-labeled 12RSS DNA (in the presence of a 23RSS and Mg^{2+}, as examined by TBE-Urea gel). (**d**) Elution profiles of RAG1/RAG2 (both long and short forms) complexed with DNA from Superdex-200 (S200) in a low salt buffer (50mM HEPES 7.0, 60 mM KCl, 1mM maltose and 2 mM DTT). Regardless of the length of RAG2, the major S200 eluant peak came out at the same time point and contained RAG1, RAG2 (1–351 or 1–387) and HMGB1 proteins, as shown in the SDS gel (right insert), as well as 12 and 23RSS oligonucleotides, as confirmed by a TBE-Urea gel stained by SYBR-Green (left insert).



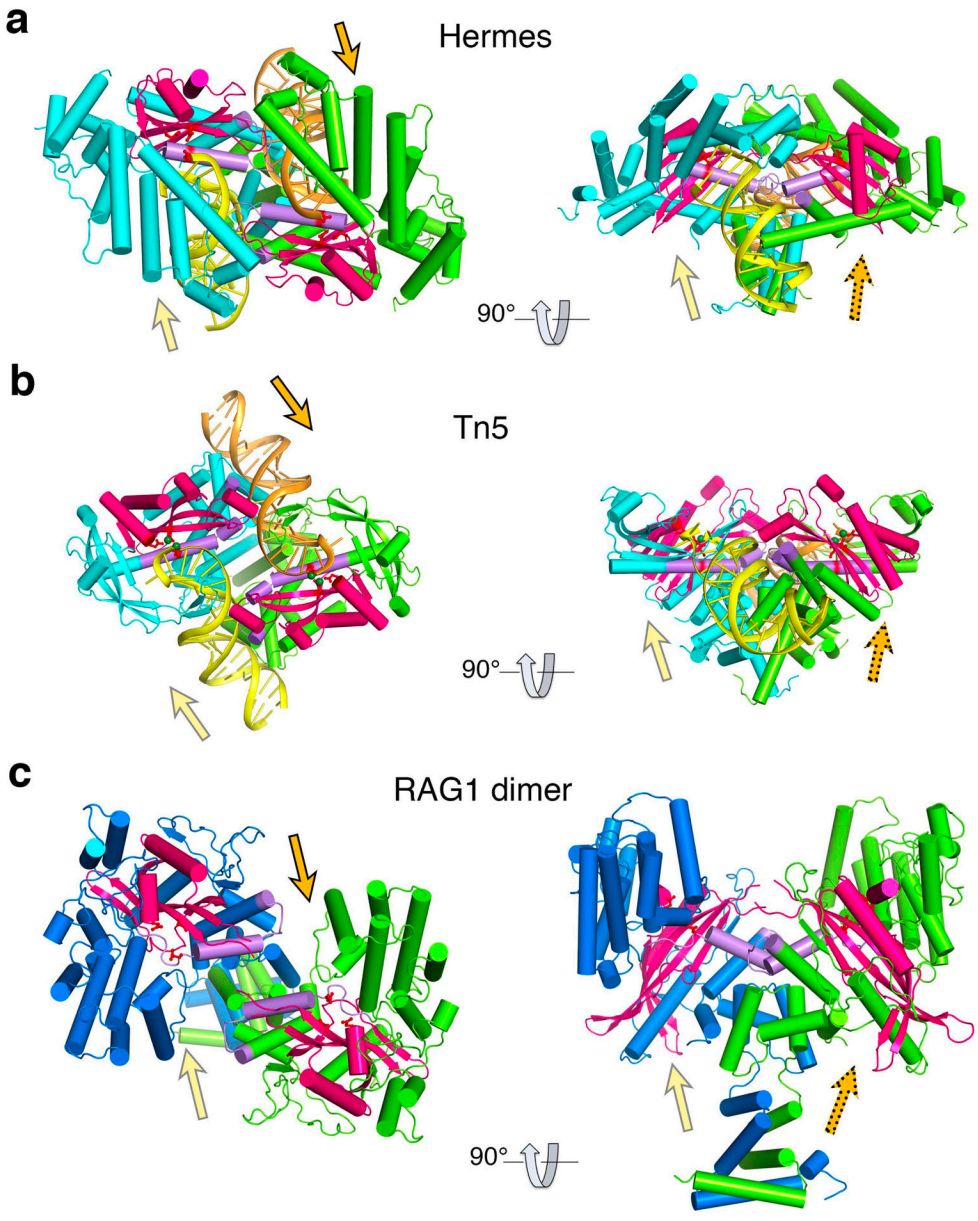
Extended Data Figure 3. Comparison of RAG2 with β-propeller and β-pinwheel structures

KLHL2 (PDB: 4CHB) [56] is selected to represent the β-propeller proteins, and the C-terminal domain (CTD) of GyrA (PDB: 1SUU) [57] is selected to represent the β-pinwheel structures. After superposition, RAG2, KLHL2 and GyrA are shown side-by-side individually in two orthogonal views. Each structure is colored from N- to C-terminus in blue to red rainbow colors. The loops in RAG2 that interact with RAG1 are labeled. The six β blades are named by Roman numerals, I to VI from N to C terminus; four β strands in each blade are named by Arabic numerals, 1 to 4.



Extended Data Figure 4. Comparison of RAG1 and NBD-DNA complex

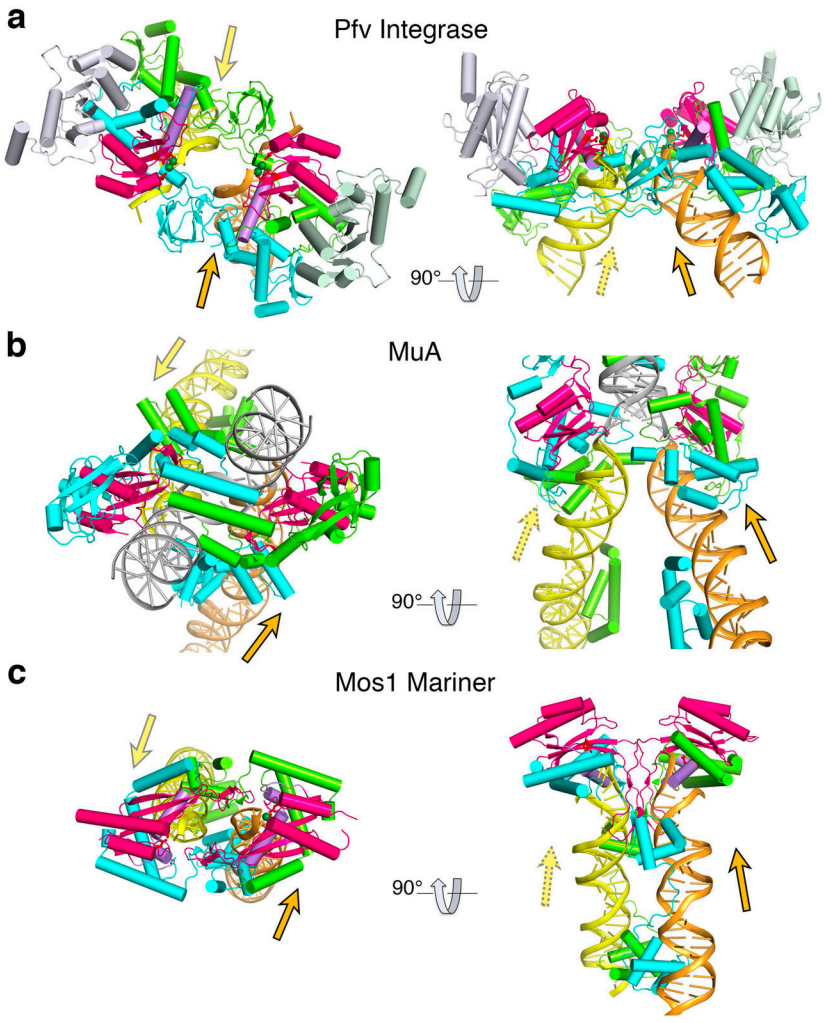
(**a**) The NBD in the RAG1/2 core complex (blue and green) superimposes well with the published structure of the NBD-DNA complex (PDB: 3GNA, protein colored yellow) [22]. (**b**) The twelve SCID/OS mutations in the NBD domain are mapped onto the crystal structure of the NBD-DNA complex. Six SCID/OS (R391 to R407) mutations are located on a positively charged surface patch that interacts with the nonamer; five remaining SCID/OS mutations (L408 to A441) appear to affect the structural integrity of the NBD, and R446 may interact with the spacer DNA in each RSS.



Extended Data Figure 5. Transposases that form a hairpin intermediate

(**a**) Hermes (PDB: 4D1Q)[37], (**b**) bacterial Tn5 (PDB: 1MUS)[40], and (**c**) RAG1 dimers are shown as ribbon diagrams in two orthogonal views, with the dyad perpendicular to the

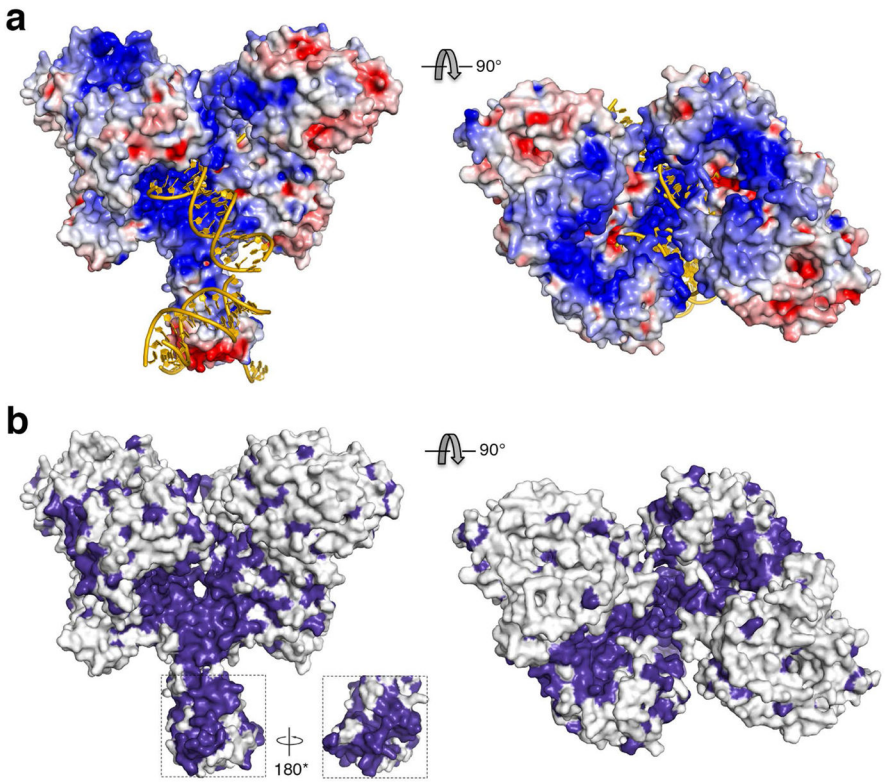
viewing plane (left) or in the plane (right). Each dimer consists of a cyan and a green subunit. The catalytic RNH domains are highlighted in pink, and the conserved catalytic residues are shown as red ball-and-sticks. The catalytic divalent metal ions are shown as green spheres if present. The DNAs, colored in yellow (cleaved by the cyan subunit) and orange (cleaved by the green subunit), have similar orientations in the Hermes and Tn5 complexes (as indicated by the arrows). Arrows with dashed outlines indicate that the DNAs are in the back of the viewing plane. Notably, the pair of RNH domains is oriented similarly in all three cases. The predicted orientations of DNAs bound to RAG1 are indicated by the yellow and orange arrows, and the α-helices connected to the third catalytic carboxylates (shown in light purple) probably bridge two DNAs in RAG1 recombinase as in Hermes and Tn5.



Extended Data Figure 6. Transposases that do not form a hairpin intermediate

(**a**) Retroviral integrase from Prototype foamy virus (Pfv, PDB: 3OS0)[36], (**b**) bacterial MuA transposase (PDB: 4FCY)[38], and (**c**) eukaryotic Mos1 mariner transposase (PDB: 3HOT)[39] are shown in comparable views and same representations as Hermes, Tn5 and RAG1/2 in Extended Data Figure 5. Each catalytic dimer consists of a cyan and a green subunit. Two

accessory subunits in Pfv are shown in light blue and green, and two accessory subunits of the MuA structure are omitted for clarity. The catalytic RNH domains are highlighted in pink. The DNAs, colored in yellow (cleaved by the cyan subunit) and orange (cleaved by the green subunit), have similar orientations (within 30°) as indicated by the arrowheads, but each differs more than 90° from the corresponding DNA in Hermes or Tn5 transposase. The gray DNA in the MuA complex represents the target of transposition. Among these three recombinases, the α-helix that follows the third catalytic carboxylate (colored in light purple) does not cross over to interact with a second DNA.



Extended Data Figure 7. Surface potential and conservation of RAG1/2 complex

(**a**) Orthogonal views of the electrostatic potential surface of the RAG1/2 structure. Blue indicates positive charges, and red negative. (**b**) Orthogonal views of the molecular surface of RAG1/2 with absolutely conserved residues highlighted in deep purple. The NBD is well conserved. The views with dyad in the plane here are related to the image shown in Fig. 5c by ~50° rotation around the dyad.

Extended Data Table 1

Statistics of native and SeMet SAD data collection and structure refinement.
Asterisk (*) indicates that Data in the highest resolution shell is shown in parenthesis.

	Native	Crystal #1	Crystal #2	Crystal #3	Crystal #4	Crystal #5
Space group	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$

	Native	Crystal #1	Crystal #2	Crystal #3	Crystal #4	Crystal #5
Cell dimensions						
a, *b*, *c* (Å)	168.8, 180.1, 200.2	168.7, 179.0, 199.3	168.5, 179.2, 200.3	169.1, 180.1, 200.8	168.7, 180.3, 202.3	169.4, 179.3, 1
α, *β*, *γ* (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Absorption (Se)		*Peak*	*Peak*	*Peak*	*Peak*	*Peak*
Wavelength (Å)	1.0000	0.97918	0.97918	0.97918	0.97918	0.97913
Resolution * (Å)	50-3.2 (3.31 – 3.2)	50.0-3.8 (3.94-3.8)	50.0-3.7 (3.83-3.7)	50.0-4.0 (4.14-4.0)	50.0-3.9 (4.04-3.9)	50.0-3.8 (3.94-
R_{merge}*	0.105 (0.58)	0.151 (0.778)	0.149 (0.831)	0.199 (0.859)	0.194 (0.796)	0.174 (0.873)
$I/\sigma I$*	12.75 (2.23)	17.7 (4.6)	17.5 (3.9)	14.4 (4.1)	14.2 (4.5)	16.8 (4.5)
Completeness * (%)	98.82 (99.9)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)
Redundancy *	7.1 (3.4)	15.0 (15.3)	15.0 (15.3)	14.8 (15.2)	14.9 (15.3)	15.0 (15.3)
Refinement						
Resolution (Å)	50 - 3.2					
No. reflections	49907					
R_{work}/ R_{free}	0.206 / 0.259					
No. atoms						
Protein	14976					
Ligand/ion (Zn^{2+})	2					
Water	0					
B-factors						
Protein	106.4					
Ligand/ion	85.0					
Water	-					
R.m.s deviations						
Bond lengths (Å)	0.007					
Bond angles (°)	1.114					

Extended Data Table 2

Missense mutations of RAG1 and RAG2 identified in human SCID/OS patients.

All SCID/OS mutations here are listed in reference #12 and #13 except for three, for which references are given in the table (#58–60). Red residues are the mutations made in mouse [4,34].

Human mutation	Mouse residue	Predicted Structural Effects
RAG1		
R394W/Q	R391A	nonamer binding
R396L/H/C	R393A	nonamer binding
S401P	S398	nonamer binding
T403P	T400	nonamer binding
R404W/A/Q	R401A	nonamer binding

Human mutation	Mouse residue	Predicted Structural Effects
R410Q/W	R407A	nonamer binding
L411P	L408	Structural integrity of NBD
D429G	D426	Structural integrity of NBD
V433M	V430	Structural integrity of NBD
M435V	L432	Structural integrity of NBD
A444V	A441	Structural integrity of NBD
R449K	R446	Probably DNA binding (spacer)
L454Q	L451	Structural integrity of RAG1 dimer
R474S/H/C	R471A	Structure & DNA binding in DDBD
S480G [58]	S477	Structure & DNA binding in DDBD
L506F	L503	Structural integrity of DDBD
R507W	R504	Solvent exposed, DNA binding?
G516A	G513	Structural integrity of RAG1
W522C	W519	Structural integrity of preR
D539V	D536	Exposed, near RAG1/2 interface
R559S	R556	At the edge of RAG1/2 interface
R561H/C	R558	RAG1/2 interface (T169 of RAG2)
A565D	A562	Structural integrity of preR
S601P	S598	Structural integrity of active site
C602W	C599	Structural integrity of active site
H612R	H609L	Disordered, near RAG2
A622P	A619	Structural integrity of RNH
R624C/H	R621A	Active site, adjacent to D600
E669G	E666	RAG1/2 interface
R699Q/W	R696	Structural stability of RNH
G709D	G706	Structural integrity, active site
R716W	R713A	Structural integrity of RAG12
G720C [59]	G717	RAG1/2 interface
E722K	E719K	RAG1/2 interface
C730E/F	C727	Structural integrity of ZnC2
L732F/P	L729	Structural integrity of ZnC2
R737H	R734A	Possibly DNA binding (coding end)
H753L	H750A	Structural integrity of RAG1/2
R764P	R761	At the edge of RAG1/2 interface
E770K	E767	RAG1/2 interface
R776Q	R773A	RAG1/2 interface

Human mutation	Mouse residue	Predicted Structural Effects
R778Q/G/W	R775A	Structural integrity of RAG1/2
P786L	P783	RAG1/2 interface
I794T [60]	I791	Structural integrity of ZnH2
R841W/Q	R838	DNA binding (near heptamer)
N855I	N852	Interacts with 3′ end of RSS
L885R	L882	Structural integrity of ZnH2
W896R	W893A	Structural integrity of ZnH2
Y912C	Y909	Structural integrity of ZnH2
I956T	I953R	Structural integrity of ZnH2
R973H/C	R970	Heptamer binding (intra-subunit)
F974L	F971	Structural integrity of CTD
R975W/Q	R972	Structural integrity of CTD
Q981R/P	Q978	Heptamer binding (inter-subunit)
K992E/R	K989	Probable DNA binding
M1006V	M1003	Domain interface of CTD-DDBD
RAG2		
G35V	G35	RAG1/2 interface (E666 of RAG1)
R39G	R39A	RAG1/2 interface (E719, R773)
C41W	C41	Structure, and RAG1/2 interface
G95R	G95	Structure integrity of RAG2
R229E/Q/W	R229	RAG1/2 interface (D546 of RAG1)
M285R	M285	Partially exposed, maybe structure

Extended Data Table 3

Mouse RAG1/2 mutations presented in reference #4.
Mutations that correspond to (SCID/OS mutations in human are highlighted in grey. R795A listed in reference #4 (highlighted in red) should be H795A.

Mutations	Location and potential functional roles
RAG1 mutations: RAG2 binding +, all other −	
K405A/H406A/R407A	Nonamer binding
R748A/H750A	R748 is near ZnC2 and ZnH2 and structurally important. H750 may stabilize the ZnC2 structure.
R773A/R775A	R773 is at RAG1/2 interface between E719 (RAG1) and R39 (RAG2). R775 is exposed to solvent.
H937A/K938A	H937 coordinates the zinc, and K938 forms a salt-bridge with E709 next to the catalytic D708.
H942A	Zinc coordination

Mutations	Location and potential functional roles
R969A/R970A	Next to CTD in the positive groove for DNA binding
RAG1 mutations: RSS binding +, Nicking and Hairpinning −	
K596A	H-bond to the carbonyl of A957 and A955, stabilizes the W956 conformation in the apo-structure
R621A/H	Next to D600 in the active site
R713A	H-bond to “O” of E719, Y725 and I726 (near RAG1/2 interface)
E719K	RAG1/2 interface
R734A	Solvent exposed, but could bind coding-end DNA
W760A	RAG1/2 interface
H795A	In the active site, next to D708
W956A	W956 is near E962 but facing exterior and separate from the active site by the protein backbone.
RAG1 mutations: Nicking +, Hairpinning −	
K608A	Disordered
H609L	Disordered
R855A/K856A	K856 is oriented toward the CTD/DDBD, R855 is solvent exposed
R890A	Near D797 carboxylate, structure integrity, near R855
W893A	Structural integrity of ZnH2
K980A	On the CTD charged surface, potential for heptamer binding
RAG1 mutations: Joining negative or defective	
R401A/R402A	Probable nonamer binding
E423Q	Forming a salt bridge with R407 that probably binds nonamer
R440A	Probable nonamer binding
E547Q	RAG1/2 interface
S723A/C	Adjacent to E719 (OS) and close to R39 of RAG2
RAG1 mutations: Gain-of-function (12RSS processing)	
E649A	Solvent exposed and adjacent to N961 and S963 near E962.
RAG2 mutations: RAG1 binding+, all other −	
K119A	Solvent exposed and part of the positive top rim of the “Y” shaped RAG1/2 complex
K283A/R	Near M285R (OS), stabilizing the loop of 306–315
RAG2 mutations: RSS binding+. Nicking and Hairpinning −	
K38A/R39A	RAG1/2 interface, possible coding-end binding
RAG2 mutations: Joining negative or defective	
K34A	Near OS mutation G35V, RAG1/2 interface (adjacent to R73 and K97)
K56A/K58A	Solvent exposed, may interact with coding DNA flank
R73A	RAG1/2 interface (adjacent to K34 and K97)
H94A	H94 is a part of the RAG2 structural core and important for the correct folding
K97A	Near OS mutation G95R at RAG1/2 interface (K34 and R73)

Mutations	Location and potential functional roles
K119R	Exposed to solvent and adjacent to K56 and K58
R167A	Near RAG1 interface (N525 of RAG1), forming positive-p stack with W172 (near P674 of RAG1)
361 frame shift	without PHD and regulatory domains

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank G. Grundy and S. Ramon-Maiques for the pioneering work that made this study possible and D. Leahy for editing the manuscript. The research was supported by the intramural research program of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health.

References

1. Sakano H, Huppi K, Heinrich G, Tonegawa S. Sequences at the somatic recombination sites of immunoglobulin light-chain genes. Nature. 1979; 280:288–294. [PubMed: 111144]
2. Lewis SM. The mechanism of V(D)J joining: lessons from molecular, immunological, and comparative analyses. Adv Immunol. 1994; 56:27–150. [PubMed: 8073949]
3. Gellert M. V(D)J recombination: RAG proteins, repair factors, and regulation. Annu Rev Biochem. 2002; 71:101–132. [pii]. 10.1146/annurev.biochem.71.090501.150203090501.150203 [PubMed: 12045092]
4. Schatz DG, Swanson PC. V(D)J recombination: mechanisms of initiation. Annu Rev Genet. 2011; 45:167–202.10.1146/annurev-genet-110410-132552 [PubMed: 21854230]
5. Deriano L, Roth DB. Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. Annu Rev Genet. 2013; 47:433–455.10.1146/annurev-genet-110711-155540 [PubMed: 24050180]
6. Boboila C, Alt FW, Schwer B. Classical and alternative end-joining pathways for repair of lymphocyte-specific and general DNA double-strand breaks. Adv Immunol. 2012; 116:1–49. B978-0-12-394300-2.00001-6 [pii]. 10.1016/B978-0-12-394300-2.00001-6 [PubMed: 23063072]
7. Oettinger MA, Schatz DG, Gorka C, Baltimore D. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science. 1990; 248:1517–1523. [PubMed: 2360047]
8. Schatz DG, Oettinger MA, Baltimore D. The V(D)J recombination activating gene, RAG-1. Cell. 1989; 59:1035–1048. 0092-8674(89)90760-5 [pii]. [PubMed: 2598259]
9. McBlane JF, et al. Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. Cell. 1995; 83:387–395. [PubMed: 8521468]
10. Agrawal A, Eastman QM, Schatz DG. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. Nature. 1998; 394:744–751. [PubMed: 9723614]
11. Hiom K, Melek M, Gellert M. DNA transposition by the RAG1 and RAG2 proteins: A possible source of oncogenic translocations. Cell. 1998; 94:463–470. [PubMed: 9727489]
12. Lee YN, et al. A systematic analysis of recombination activity and genotype-phenotype correlation in human recombination-activating gene 1 deficiency. J Allergy Clin Immunol. 2014; 133:1099–1108. S0091-6749(13)01558-3 [pii]. 10.1016/j.jaci.2013.10.007 [PubMed: 24290284]
13. Piirila H, Valiaho J, Vihinen M. Immunodeficiency mutation databases (IDbases). Hum Mutat. 2006; 27:1200–1208.10.1002/humu.20405 [PubMed: 17004234]
14. Cuomo CA, Oettinger MA. Analysis of regions of RAG-2 important for V(D)J recombination. Nucleic Acids Res. 1994; 22:1810–1814. [PubMed: 8208604]

15. Sadofsky MJ, Hesse JE, Gellert M. Definition of a core region of RAG-2 that is functional in V(D)J recombination. Nucleic Acids Res. 1994; 22:1805–1809. [PubMed: 8208603]
16. Sadofsky MJ, Hesse JE, McBlane JF, Gellert M. Expression and V(D)J recombination activity of mutated RAG-1 proteins. Nucleic Acids Res. 1993; 21:5644–5650. [PubMed: 8284210]
17. Silver DP, Spanopoulou E, Mulligan RC, Baltimore D. Dispensable sequence motifs in the RAG-1 and RAG-2 genes for plasmid V(D)J recombination. Proc Natl Acad Sci USA. 1993; 90:6100–6104. [PubMed: 8327489]
18. Grundy GJ, et al. Initial stages of V(D)J recombination: the organization of RAG1/2 and RSS DNA in the postcleavage complex. Mol Cell. 2009; 35:217–227. S1097-2765(09)00460-2 [pii]. 10.1016/j.molcel.2009.06.022 [PubMed: 19647518]
19. Jones JM, Gellert M. Intermediates in V(D)J recombination: a stable RAG1/2 complex sequesters cleaved RSS ends. Proc Natl Acad Sci U S A. 2001; 98:12926–12931. [pii]. 10.1073/pnas.221471198221471198 [PubMed: 11606753]
20. Liu Q, Zhang Z, Hendrickson WA. Multi-crystal anomalous diffraction for low-resolution macromolecular phasing. Acta crystallographica. Section D, Biological crystallography. 2011; 67:45–59.10.1107/S0907444910046573
21. Gigi V, et al. RAG2 mutants alter DSB repair pathway choice in vivo and illuminate the nature of ‘alternative NHEJ’. Nucleic Acids Res. 2014; 42:6352–6364. gku295 [pii. 10.1093/nar/gku295 [PubMed: 24753404]
22. Yin FF, et al. Structure of the RAG1 nonamer binding domain with DNA reveals a dimer that mediates DNA synapsis. Nat Struct Mol Biol. 2009; 16:499–508. nsmb.1593 [pii]. 10.1038/nsmb.1593 [PubMed: 19396172]
23. Fugmann SD, Villey IJ, Ptaszek LM, Schatz DG. Identification of two catalytic residues in RAG1 that define a single active site within the RAG1/RAG2 protein complex. Mol Cell. 2000; 5:97–107. [PubMed: 10678172]
24. Kim DR, Dai Y, Mundy CL, Yang W, Oettinger MA. Mutations of acidic residues in RAG1 define the active site of the V(D)J recombinase. Genes Dev. 1999; 13:3070–3080. [PubMed: 10601033]
25. Landree MA, Wibbenmeyer JA, Roth DB. 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. 1999; 13:3059–3069. [PubMed: 10601032]
26. Nesmelova IV, Hackett PB. DDE transposases: Structural similarity and diversity. Advanced drug delivery reviews. 2010; 62:1187–1195.10.1016/j.addr.2010.06.006 [PubMed: 20615441]
27. Swanson PC. The DDE motif in RAG-1 is contributed in trans to a single active site that catalyzes the nicking and transesterification steps of V(D)J recombination. Mol Cell Biol. 2001; 21:449–458.10.1128/MCB.21.2.449-458.2001 [PubMed: 11134333]
28. Yang W, Steitz TA. Recombining the structures of HIV integrase, RuvC and RNase H. Structure. 1995; 3:131–134. [PubMed: 7735828]
29. Gwyn LM, Peak MM, De P, Rahman NS, Rodgers KK. A zinc site in the C-terminal domain of RAG1 is essential for DNA cleavage activity. J Mol Biol. 2009; 390:863–878. S0022-2836(09)00670-6 [pii]. 10.1016/j.jmb.2009.05.076 [PubMed: 19500590]
30. Rodgers KK, et al. A zinc-binding domain involved in the dimerization of RAG1. J Mol Biol. 1996; 260:70–84. S0022-2836(96)90382-4 [pii]. 10.1006/jmbi.1996.0382 [PubMed: 8676393]
31. Aravind L, Koonin EV. Gleaning non-trivial structural, functional and evolutionary information about proteins by iterative database searches. J Mol Biol. 1999; 287:1023–1040. [PubMed: 10222208]
32. Callebaut I, Mornon JP. 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. 1998; 54:880–891. [PubMed: 9760994]
33. Swanson PC, Desiderio S. RAG-2 promotes heptamer occupancy by RAG-1 in the assembly of a V(D)J initiation complex. Mol Cell Biol. 1999; 19:3674–3683. [PubMed: 10207091]
34. Huye LE, Purugganan MM, Jiang MM, Roth DB. 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. 2002; 22:3460–3473. [PubMed: 11971977]

35. Ko JE, Kim CW, Kim DR. Amino acid residues in RAG1 responsible for the interaction with RAG2 during the V(D)J recombination process. J Biol Chem. 2004; 279:7715–7720. M311471200 [pii]. 10.1074/jbc.M311471200 [PubMed: 14670978]
36. Hare S, Gupta SS, Valkov E, Engelman A, Cherepanov P. Retroviral intasome assembly and inhibition of DNA strand transfer. Nature. 2010; 464:232–236.10.1038/nature08784 [PubMed: 20118915]
37. Hickman AB, et al. Structural basis of hAT transposon end recognition by Hermes, an octameric DNA transposase from Musca domestica. Cell. 2014; 158:353–367.10.1016/j.cell.2014.05.037 [PubMed: 25036632]
38. Montano SP, Pigli YZ, Rice PA. The mu transpososome structure sheds light on DDE recombinase evolution. Nature. 2012; 491:413–417.10.1038/nature11602 [PubMed: 23135398]
39. Richardson JM, Colloms SD, Finnegan DJ, Walkinshaw MD. Molecular architecture of the Mos1 paired-end complex: the structural basis of DNA transposition in a eukaryote. Cell. 2009; 138:1096–1108.10.1016/j.cell.2009.07.012 [PubMed: 19766564]
40. Steiniger-White M, Rayment I, Reznikoff WS. Structure/function insights into Tn5 transposition. Curr Opin Struct Biol. 2004; 14:50–57.10.1016/j.sbi.2004.01.008 [PubMed: 15102449]
41. Grundy GJ, Hesse JE, Gellert M. Requirements for DNA hairpin formation by RAG1/2. Proc Natl Acad Sci U S A. 2007; 104:3078–3083. 0611293104 [pii]. 10.1073/pnas.0611293104 [PubMed: 17307873]
42. Lu CP, Sandoval H, Brandt VL, Rice PA, Roth DB. Amino acid residues in Rag1 crucial for DNA hairpin formation. Nat Struct Mol Biol. 2006; 13:1010–1015. nsmb1154 [pii]. 10.1038/nsmb1154 [PubMed: 17028591]
43. Aidinis V, et al. The RAG1 homeodomain recruits HMG1 and HMG2 to facilitate recombination signal sequence binding and to enhance the intrinsic DNA-bending activity of RAG1-RAG2. Mol Cell Biol. 1999; 19:6532–6542. [PubMed: 10490593]
44. van Gent DC, Hiom K, Paull TT, Gellert M. Stimulation of V(D)J cleavage by high mobility group proteins. EMBO J. 1997; 16:2665–2670.10.1093/emboj/16.10.2665 [PubMed: 9184213]
45. Kim DR, Oettinger MA. Functional analysis of coordinated cleavage in V(D)J recombination. Mol Cell Biol. 1998; 18:4679–4688. [PubMed: 9671478]
46. Aricescu AR, Lu W, Jones EY. A time- and cost-efficient system for high-level protein production in mammalian cells. Acta crystallographica. Section D, Biological crystallography. 2006; 62:1243–1250.10.1107/S0907444906029799
47. Swanson PC. A RAG-1/RAG-2 tetramer supports 12/23-regulated synapsis, cleavage, and transposition of V(D)J recombination signals. Mol Cell Biol. 2002; 22:7790–7801. [PubMed: 12391148]
48. Kabsch W. Xds. Acta Crystallogr D Biol Crystallogr. 2010; 66:125–132. S0907444909047337 [pii]. 10.1107/S0907444909047337 [PubMed: 20124692]
49. Sheldrick GM. A short history of SHELX. Acta Crystallogr A. 2008; 64:112–122. S0108767307043930 [pii]. 10.1107/S0108767307043930 [PubMed: 18156677]
50. McCoy AJ, et al. Phaser crystallographic software. J Appl Crystallogr. 2007; 40:658–674.10.1107/S0021889807021206 [PubMed: 19461840]
51. Terwilliger TC. Maximum-likelihood density modification. Acta Crystallogr D Biol Crystallogr. 2000; 56:965–972. S0907444900005072 [pii]. [PubMed: 10944333]
52. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010; 66:486–501. S0907444910007493 [pii]. 10.1107/S0907444910007493 [PubMed: 20383002]
53. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta crystallographica. Section D, Biological crystallography. 2010; 66:213–221.10.1107/S0907444909052925
54. Chen VB, et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr. 2010; 66:12–21. S0907444909042073 [pii. 10.1107/S0907444909042073 [PubMed: 20057044]
55. Chenna R, et al. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. 2003; 31:3497–3500. [PubMed: 12824352]

56. Schumacher FR, Sorrell FJ, Alessi DR, Bullock AN, Kurz T. Structural and biochemical characterization of the KLHL3-WNK kinase interaction important in blood pressure regulation. The Biochemical journal. 2014; 460:237–246.10.1042/BJ20140153 [PubMed: 24641320]

57. Corbett KD, Shultzaberger RK, Berger JM. The C-terminal domain of DNA gyrase A adopts a DNA-bending beta-pinwheel fold. Proceedings of the National Academy of Sciences of the United States of America. 2004; 101:7293–7298.10.1073/pnas.0401595101 [PubMed: 15123801]

58. Schuetz C, et al. Lesson from hypomorphic recombination-activating gene (RAG) mutations: Why asymptomatic siblings should also be tested. The Journal of allergy and clinical immunology. 2014; 133:1211–1215.10.1016/j.jaci.2013.10.021 [PubMed: 24331380]

59. Corneo B, et al. Identical mutations in RAG1 or RAG2 genes leading to defective V(D)J recombinase activity can cause either T-B-severe combined immune deficiency or Omenn syndrome. Blood. 2001; 97:2772–2776. [PubMed: 11313270]

60. Dhingra N, et al. Severe combined immunodeficiency caused by a new homozygous RAG1 mutation with progressive encephalopathy. Hematol Oncol Stem Cell Ther. 2014; 7:44–49. S1658-3876(13)00093-9 [pii]. 10.1016/j.hemonc.2013.11.001 [PubMed: 24333136]

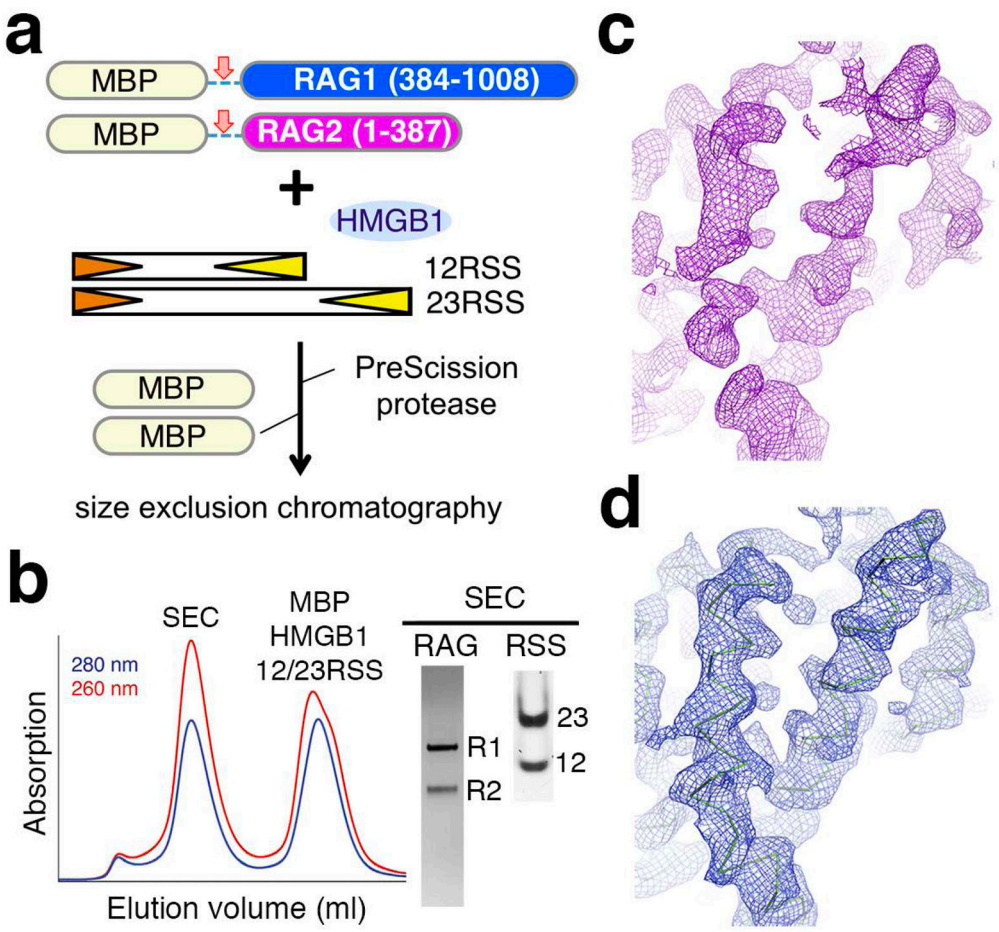


Fig. 1.

Structure determination of RAG1/2 recombinase. (**a**) Procedure of assembling SEC from purified RAG1/2, RSS DNAs and HMGB1. The MBP tags were cleaved off by PreScission protease after SEC formation. (**b**) The RAG1/2 – DNA complex was purified away from MBP tags, free DNA, and HMGB1 on a Superdex-200 column. The eluted peak contains RAG1/2 protein and RSS DNAs, as shown in the protein and DNA denaturing gels stained by Coomassie Blue and SYBR-Green, respectively. (**c, d**) The experimental electron density map calculated from merging (**c**) two best SAD datasets or (**d**) all six SAD datasets.

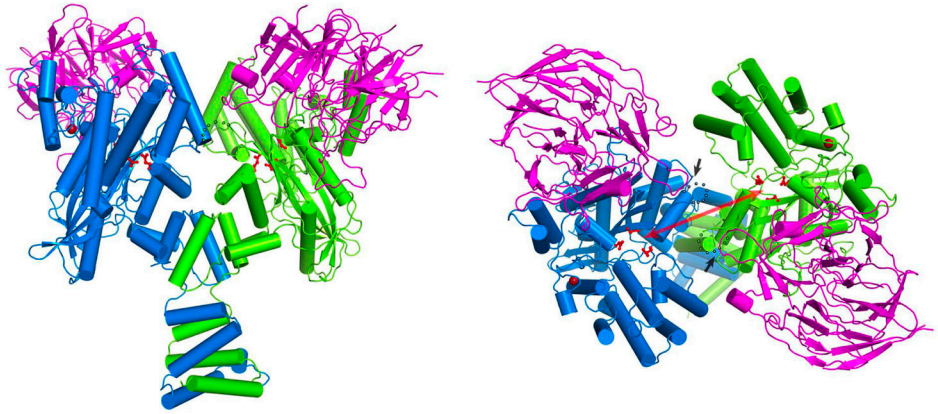


Fig. 2.

Crystal structure of RAG1/2. (**a**) The front and (**b**) top view of the RAG1/2 heterotetramer. The two RAG1 chains are shown in blue and green ribbon diagrams, and both RAG2 subunits are shown in magenta. The active sites are highlighted by the three carboxylates shown as red sticks. The zinc ions are shown as dark red spheres. The distance between the two active sites is ~45Å (marked by the red double arrowheads). The disordered loop of residues 608–614 in RAG1 near the RAG1/2 interface is marked by a dotted line and grey arrowhead.

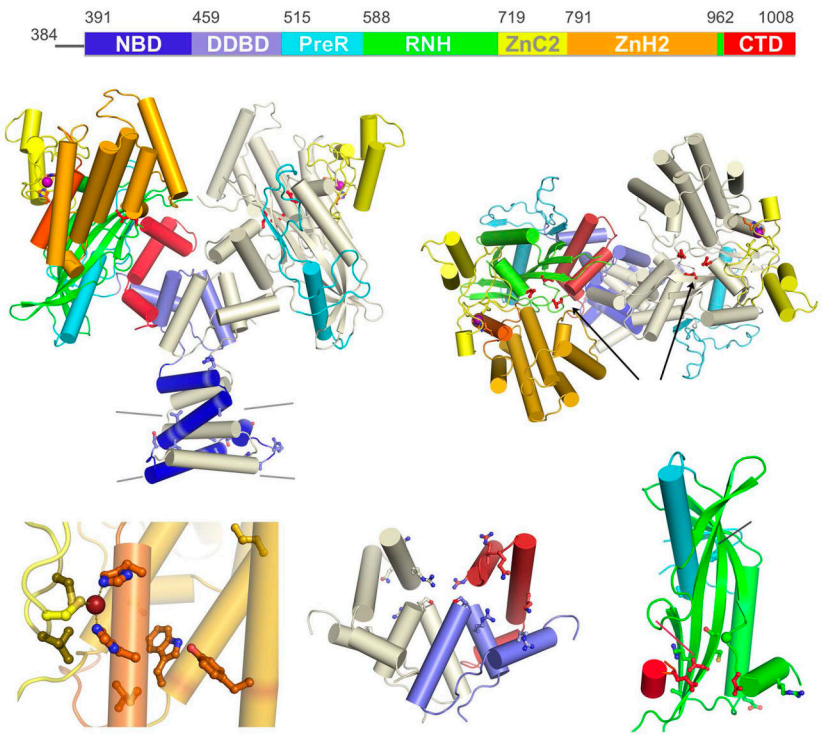


Fig. 3.

The RAG1 structure. (**a**) A diagram of RAG1 domains with boundaries indicated by residue numbers. (**b**) Cartoons of RAG1 dimer. One subunit is color-coded as in (a), and the other in silver except for the preR and ZnC2 domains. (**c**) An orthogonal view of (b). (**d**) The Zn^{2+} coordination by two Cys (of ZnC2) and two His residues (of ZnH2). (**e**) The DDBD and CTD domains. (**f**) The RNH and a portion of preR domain with the catalytic carboxylates shown in red sticks. In panels b, d e, f, SCID/OS mutations are shown as colored sticks with black labels.

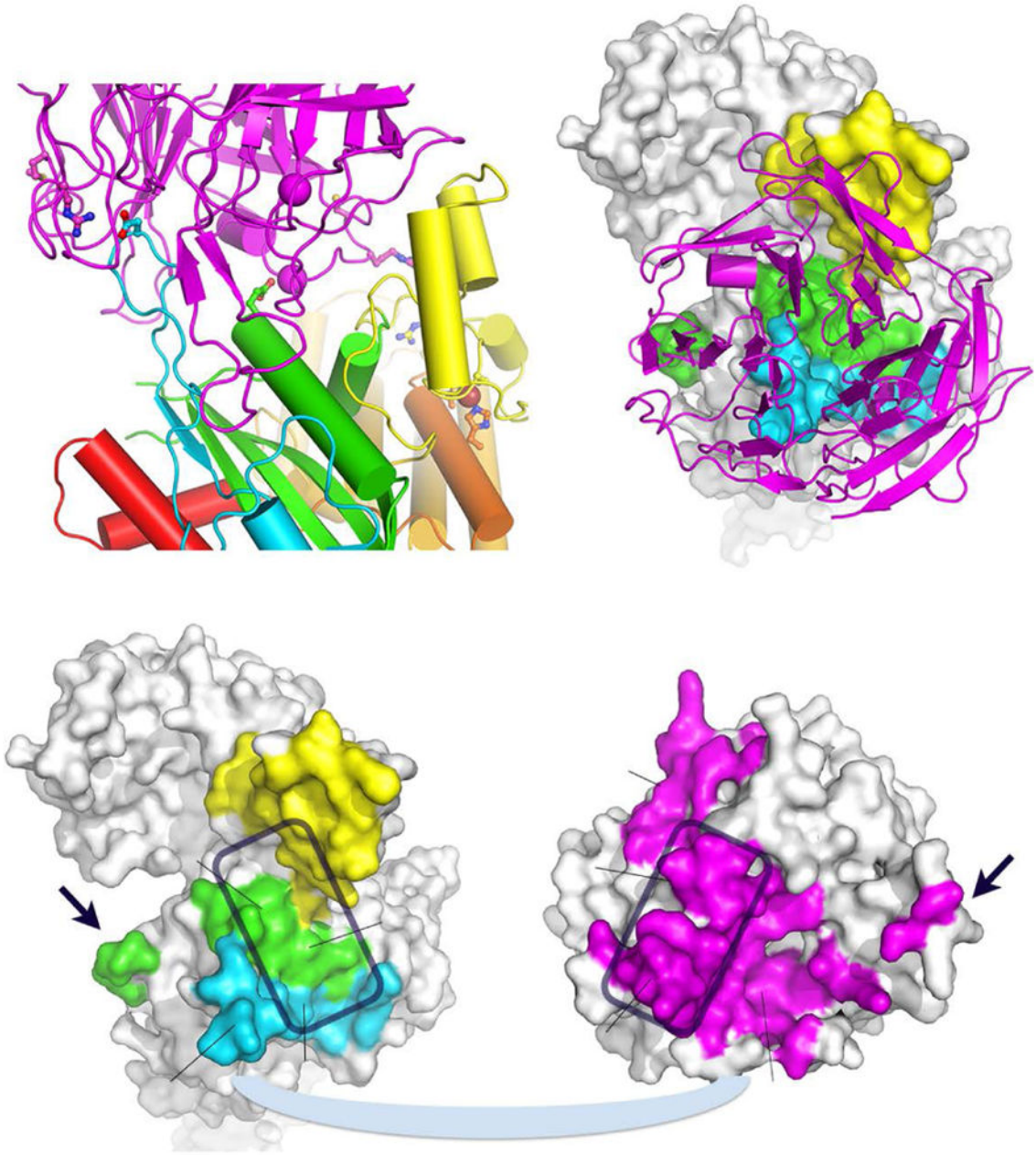


Fig. 4.

The interface between RAG1 and RAG2. (**a**) One side of the doughnut-shaped RAG2 interacts with the preR, RNH and ZnC2 domains of RAG1 (color-coded as in Fig. 3). The SCID/OS mutations in RAG2 are shown and labeled in black. (**b**) An orthogonal view of the RAG1/2 interface. The Kelch repeats of RAG2 are labeled I to VI. (**c**) Based on (b), the interface of RAG1 and RAG2 is shown as an open book. The regions at the interface are indicated according to the color code, and SCID/OS mutations are labeled in black. The mirrored arrowheads and boxes indicate matching surfaces.

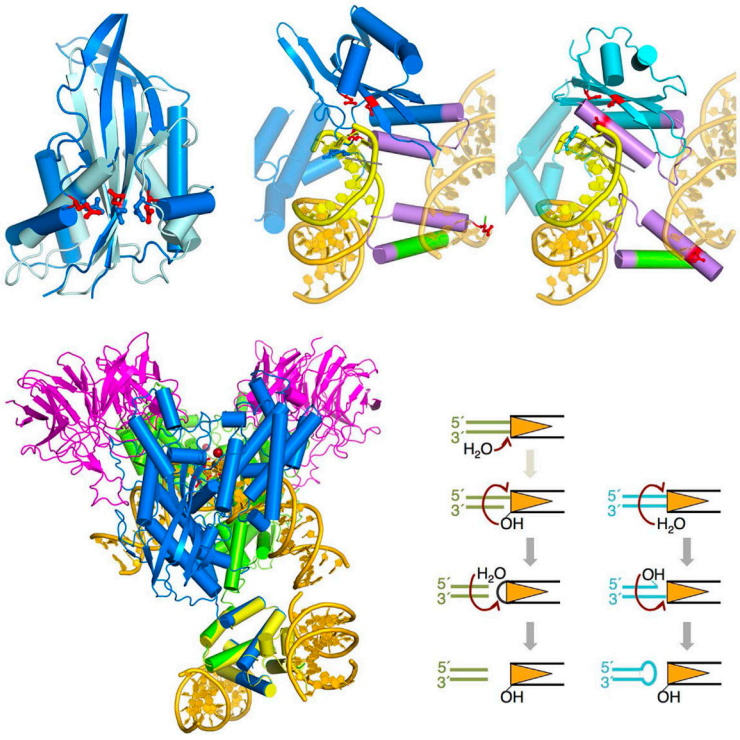


Fig. 5.

A RAG1/2-DNA model. (**a**) Superposition of the RNH domains of RAG1 (red active site) and Hermes (PDB: 4D1Q [37]). (**b**) The superposition places the 16bp DNA of the Hermes-DNA complex in the RAG1 active site. The DNA is colored yellow (the first 7 bp) and gold. The second DNA is shown without additional manipulation. The α-helices bridging the two DNAs are colored lilac. (**c**) The RAG1/2-RSS DNA model resulting from superposition with the Hermes and the NBD–DNA complex (12bp, PDB: 3GNA) [22]. (**d**) DNA cleavage by hairpin-forming bacterial and eukaryotic transposases. The recognition sequences are represented by orange triangles.