15 Oct. 2016

Final Scientific Report covering all the results obtained from 1st of July 2012 to its concluding stage on 15th Oct. 2016.

The present report summarizes the main results obtained throughout the entire project duration from the 1st July 2012 to the final stage 15th October 2016 for the project **PN-II-ID-PCCE-2011-2-0024,** entitled**: "**V(D)J recombination targeted *in cis* by transcription induced DNA supercoiling. " funded from PNII Romanian National Research funds via UEFISCDI and lead by Project Director Dr. Mihai Ciubotaru.

Project Summary

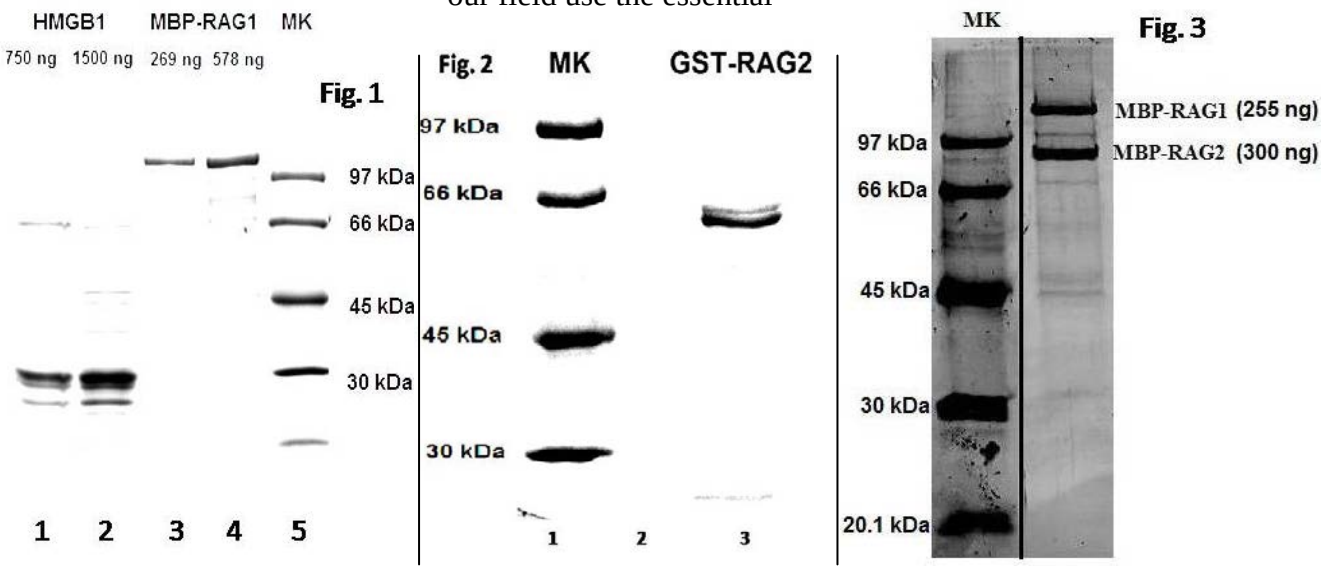
Somatic recombination assembles and diversifies the antigen receptor genes, of B and T cells. RAG recombinase binds two specific DNA sites (RSSs) to initiate recombination, a process called paired complex (PC) or synapsis formation. Synapsis is followed by cleavage and then by repair of the double stranded breaks which physiologically occur only between gene segments located intrachromosomally. Pairing preserves the integrity of the genome, preventing aberrant interchromosomal joinings (translocations) implicated in human B and T cell lymphomas. Despite its significance the mechanism that physiologically directs somatic recombination for sites located *in cis* with respect to DNA is entirely unknown. Prior to recombination, dechromatinization, and germline transcription of the variable gene segments of the immunoglobulin (Ig) or T cell receptor (TCR) loci lead to a substantial dynamic change in the torsional tension and supercoiling of the adjacent DNA (*2-7*). This work studies the configuration of RSSs in the PC, the role of transcription and its induced DNA supercoiling in RSS site orientation as active factors in controlling RAG intrachromosomal (in cis) sites pairing. To investigate the configuration of RSSs in the PC we developed a FRET based assay with which we show that both 12 and 23RSSs are arched in the PC in a U shape configuration which similarly positions in both RSS types the extreme bends at the junctions of RSS consensus heptamer/nonamer with the spacers. Using the same FRET assay we prove considerable 12RSS conformation differences to occur in the PC versus its signal complex with RAG, configurations which may account for 12/23 rule and its implications in hairpin formation. To supercoil the DNA subjected to recombination and to test both *in vitro* in the test tube as well as intracellularly the targeting mechanism we used a T7 phage heterologous controllable transcription unit juxtaposed in the proximity of two reporter 12RSSs with respect to a constitutive 23RSS. Although regardless of their deletional or inversional orientation intracellularly transcription enhances recombination from both reporter 12RSSs, the effects are substantially more pronounced for the deletional orientation and for the distal 12RSS(with respect to its 23RSS partner). We first studied the effects of site orientation on RAG catalysis by *in vitro* coupled cleavage on planar short constrained DNA circles and show that spatial phase RSS orientation is important displaying a sinusoidal activity correlation. Using longer DNA circles containing RSS sites of defined topology and orientation, we show that the rate and efficiency of RAG synapsis is affected both by their configuration and type of supercoiling two factors that we suggest favor in vivo RAG preferential *in cis* recombination. Moreover, we bring strong evidence in support that local transcription directly facilitates recombination, a novel and unexpected finding which should neither be associated with its subsidiary role in enhancing RAG site accessibility nor with its locally induced supercoiling. To study and contrast RAG *in trans* effects we developed a LacI DNA tethering artificial system which facilitates RAG induced targeting *in trans* with respect to

DNA and with it we reinvestigated in this context the effects of transcription. Lac I presence enhances dramatically in trans RAG coupled cleavage but unlike the *in cis* situation local transcription has only a moderate to weak enhancing catalytic effect. The knowledge of RAG preferred RSS configuration, orientation and topology allowed us to develop this LacI in trans DNA tethering system an unprecedented experimental setup with which we can induce intracellularly RAG targeted chromosomal translocations.
-The paragraphs denoting the objectives, the methods, the experiments addressing them and their interpretation keep the same initial labels with those used in the original project proposal.

Project Objectives:
Obj1-D2A. Identify the mechanism by which DNA supercoiling activates RAG catalysis.

(objective D2A is addressed in the experimental section of the proposal at subsection D3A)

D3A1. Purification of RAG1, RAG2 and HMGB1/2 proteins. Almost all *in vitro* studies in our field use the essential



"core" domains of RAG1 (384-1008 from a total of 1040aa)(1) and RAG2(region 1-383 from 527aa) of murine origin, which are catalytically more active, more soluble and easier to purify than full length proteins (3). To test the RAG activation mechanism by in vitro methods we use the following proteins, whose encoding vectors have been built in our previous work and are expressed and purified as follows:

a) MBP-cRAG1(384-1008)H6 the catalytic domain of RAG1 expressed in E. coli (fused at its N terminus with Maltose Binding Protein (MBP) and a poli 6xHis at C terminus). A similar catalytic mutant MBP-cRAG1D708A catalytic site mutant unable to nick and cleave DNA, was purified for our ligation assays. The purification took place in two steps, first on an amylose affinity column(elution with maltose) followed by a subsequent Nickel affinity column(elution with Imidazole)(wells 3 and 4, 102kDa, SDS-PAGE in Fig1).

b) HMGB1 (full length High mobility Group protein class B -1)fused at its C terminus with a poli 6xHis tail is expressed in E. coli, purified on Nickel (wells 1 and 2, 29KDa, Fig1).

We also expressed and purified a HMGB2 ΔCHis6 (1-185of 210aa) with its last acidic 25 aa from the C terminus deleted, a protein with higher DNA unwinding properties than full length HMGB1(data not shown), as described by Zwilling et al. 1995(5).

c) The catalytic domain of RAG2 GST-cRAG2 (1-387), fused at its N terminus with Glutathione-S-transferase(GST) is expressed in human HEK 293T fibroblasts following transient transfection with a pEBG vector (with pEBV Epstein Barr Virus promoter). Purification was done by affinity on a GST-Sepharose column, eluting with Glutathione (well 3, 62kDa,Fig2). Proteins described at a), b) si c) are used to study the role played by each individual component in the process of synaptic complex assembly. However, the catalytic activity of the reconstituted complex of MBP-RAG1 (384-1080) &GST-RAG2(1-387) is slightly lower than that of coexpressed RAGs (see d).

d) MBP-cRAG1(384-1040) and MBP-cRAG2 (1-387) coexpressed in in human HEK 293T fibroblasts following transient transfection with pcDNA vectors(RAG1, 120kDa and RAG2, 87kDa, Fig. 3). This mix of RAG1 andRAG2 is optimal for testing the RAG catalytic activity.

All purified proteins were tested for their activity in RAG coupled cleavage, and all subsequent in vitro experiments shown below were performed with them.

Ob1-D3A2a *In cis* RAG synapsis kinetic effects tested by IS264 facilitated ligation.

In our project proposal we presented preliminary data showing how in a facilitated ligation assay RAG increases the efficiency of circularization of a 431bp linear IS264 DNA(from pJH290 construct) with a pair of 12/23-RSSs (Fig. 4A pg.24 of the original proposal). Upon ligation in the presence of HMGB2 this 431bp linear DNA generates 3 topoisomeric circles a) relaxed circle RlxC1, b) slightly unwound UnwC2 and c) unwound UnwC3. D708ARAG1 and RAG2 and HMGB2 presence in the ligation mix(complete RAG facilitated ligation) enhances almost 5 folds the production of negatively supercoiled circles especially of the topoisomer UnwC3. To further investigate the nature of these topoisomers the DNA circles were gel purified and after appropriate treatment were imaged by Electron microscopy (EM), in collaboration with Dr. Jack Griffith's group from University of North Carolina at Chapel Hill. Fig. 4 shows their EM images. The

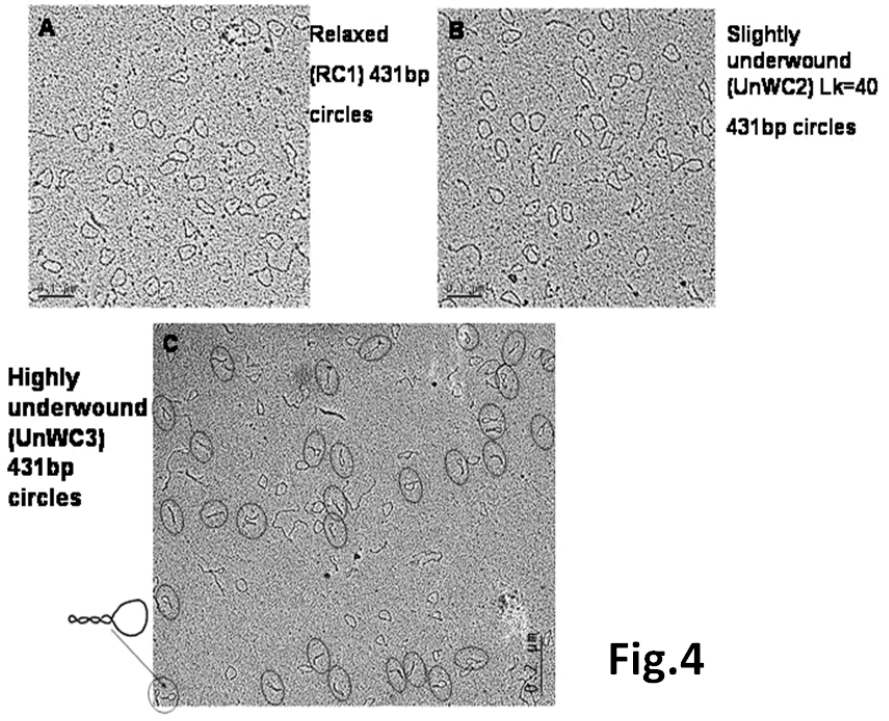
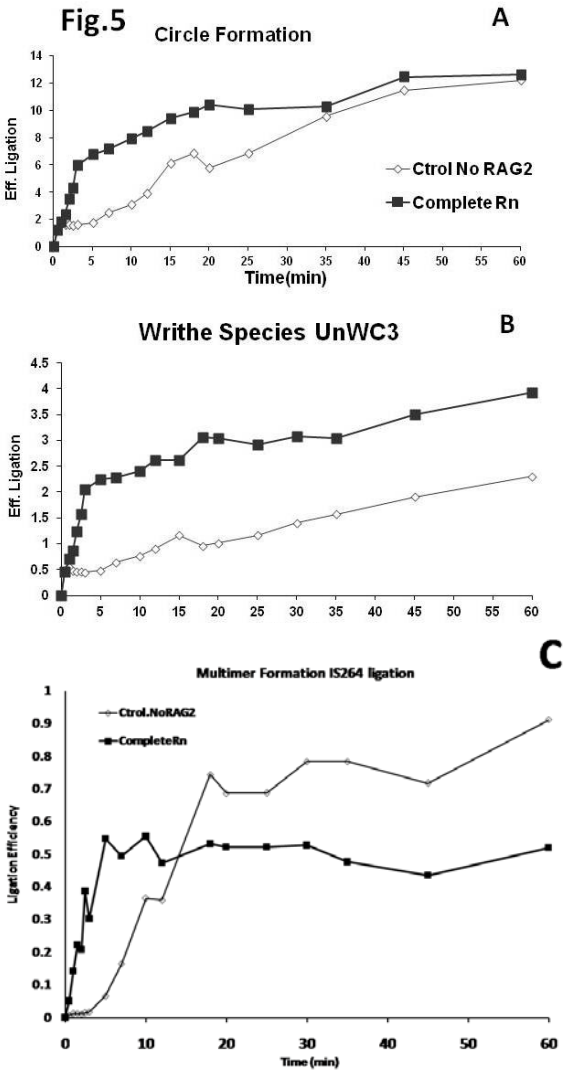


Fig.4

species denoted RlxC1 and UnwC2 (Fig. 4A and B) show circles with their circular contours contained in a single plane without self crossings. Based on these images and on the results of the topoisomerase treatment experiments (data not shown) we concluded RlxC1 and UnwC2 to be two circular topoisomers, one the relaxed circle of Lk = 41(Lk=linking number) and the other one a circle underwound by one full helical turn Lk = 40 (in absence of writhe, Lk =Tw (twist number) = N/10.5 coincides with the ratio of the total number of bases N (431bp) divided by 10.5bp/turn). Unlike these last two types of circles the UnwC3 species shown in Fig. 4C displays multiple self crossings (intertwinings which give rise to Writhe-included in ovals in Fig.4C) with most of the individual species described as a "handle pan" and some as plectonemic rods. The "handle pan" species drawn adjacently has a handle represented by a plectonemic intertwined portion of the DNA with a thick rod-shape aspect while the rest of the DNA describes a circular (or semicircular) path -"the pan". Because RAG cleaves considerably better UnwC3 than either the linear or the other two topoisomeric circular DNA (see Fig4B at pg25 original proposal, preliminary data) we attribute the enhanced catalytic activity of this type of substrate to its plectonemic nature. We further addressed the kinetics of a RAG facilitated ligation focusing especially on the advent and accumulation of UnwC3 product. Time course ligations were performed in parallel on the IS264 linear DNA in one set the reactions being facilitated by a complete RAG mix (HMGB2ΔC, MBPcR1D708A and GST-RAG2), and in a control set the mixtures being set under similar conditions but in absence of GST-RAG2(no RAG2 control). The products were separated on native PAGE gels(similarly with reaction described in Fig. 4A pg.24 of the original proposal) and labeled identified species quantified with PhosphorImager. Fig. 5A



displays as Ligation Efficiency (Lig. Eff. = Lig. Product/Total linear DNA substrate) versus time the cumulative results for all circles formed in 60 mins., whereas Fig. 5B, depicts only the formation of UnwC3. In each case both curves are presented, the complete RAG reactions(dark squares) and that of the control reactions without RAG2(no RAG2 -open diamonds). RAG enhances considerably the rate and efficiency of circle formation in the first 5 minutes of the reactions whereas its effects plateau in the next twenty minutes. The described effects are most dramatic for UnwC3 product who clearly depends upon RAG presence even after 60 min. of ligation. These results strongly suggest that RAG just does not passively bind better to already writhed DNA, stabilizing its formation during ligation(a case in which the UnwC3 should have been similarly generated in RAG complete and control reactions at least in the early reaction times), but instead, show that RAG actively induces writhe. Besides pairing the consensus 12/23-RSSs, RAG synapsis formation enhances adjacent flanking DNA intertwining. This was quite an unexpected result for us. Hence, negative intertwining is facilitated by RAG mediated synapsis of sites located in deletional orientation and at its turn these substrates with plectonemic negative supercoiling enhance the catalytic activity of the recombinase. As shown in Fig.4A of our original proposal (pg.24) ligation of IS264 construct besides generating monomeric circles also forms linear dimers and trimers of the DNA, products which assumes *in trans* RAG synapsis. In the IS264 kinetics experiment presented we also quantified the minute amount of multimers generated by RAG synapsis *in trans* and these data are presented as time-course in Fig. 5C. One should note first the reduced levels of ligation efficiencies displayed by this plot(RAG mix saturation levels are almost 20 times lower than those of monomer circle formation and five fold lower than those of UnwC3 topoisomer). However, in the early stages of ligation RAG enhances synapsis *in trans* but interestingly after 13 min. the presence of RAG-HMGB in the ligation mix has an inhibitory effect, most likely due to the preferential supercoiled substrate being used in the monomeric cyclization reaction. It is exactly this process that we account for in RAG preference for *cis* vs *trans* synapsis.

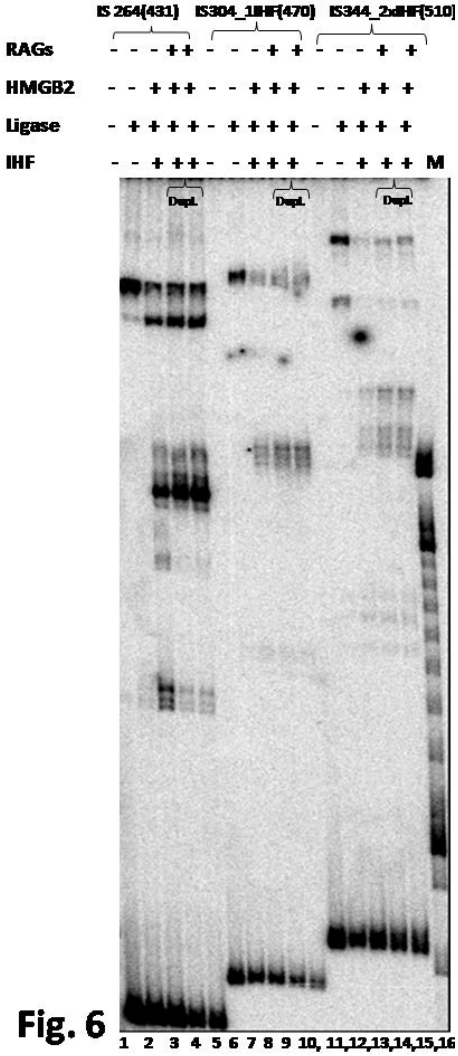


Fig. 6

In the original proposal at D3A2 section pg. 27 we described a strategy to preform Wr = +/-1 DNA constructs using the ability of the bacterial protein integration host factor IHF to bend the DNA. Starting with pJH290 vector we made first a construct where an 40bp IHF binding element was inserted in the arm 5' flanking 12-RSS (called IS304_1IHF) and one in the middle of the intervening sequence IS264 separating the two 12/23RSSs(see Fig.5 pg. 27original proposal) denoted IS344_2xIHF. Purified IHF factor was provided as a generous gift to our lab. by Dr. Nigel Grindley from Yale University. We performed facilitated ligation experiments with radiolabeled DNAs, in which a linear amplified region of 470bp IS304_1IHF respectively 510bp

IS344_2xIHF, as well as our original 431bp IS264 were subjected to ligation in the presence or absence of IHF, T4 DNA ligase, HMGB2ΔC, MBPcR1D708A and GST-RAG2. After deproteinization the products of the ligation reactions were resolved on a native PAGE whose autoradiogram is shown in Fig.6. In this figure for each construct Lanes 1, 6 and 11 show just the unligated linear DNAs, lanes 2, 7 and 12 show ligations in the presence of just T4DNA ligase, lanes 3,8 and 13 ligations in presence of HMGB2 and IHF(excess saturating conc.)-writhing mix, lanes 4 and 5, 9 and 10, and 14 and 15(duplicates), ligations in presence of full complement of RAG mix(HMGB2ΔC, MBPcR1D708A and GST-RAG2) with equimolar concentrations of IHF with respect to DNA. Whereas IHF does not interfere with the RAG and HMGB2 effects described above on IS264 construct (has no IHF binding elements) on both DNAs containing one or two such elements, it dramatically inhibited ligation circularization(compare lanes 7 with 8 for IS304_1IHF and lanes 12 and 13 for IS344_2xIHF). Adding RAG to these writhing conditions did not improved the circularization efficiency. To our discontent we concluded that IHF binding to its cognate DNA elements whereas bending dramatically the DNA also makes it rather rigid and sets out of register the two arms that need to anneal and get in the proximity . Thus prevents T4 DNA ligase to perform cyclic DNA closure. Because IHF binding instead of facilitating the formation of a writhed configuration prevents DNA circularization we had to adopt another alternative strategy from the one discussed in our original proposal (original proposal section D3A2 pg. 27).

Ob1-D3A1b, D3A2. Testing the supercoiling effect on RAG mediated catalysis; evidence for a preferred configuration.

The kinetics of IS264 circularization shows that early RAG synaptic complex formation precedes the onset of a writhed circular configuration, strongly suggesting that RAG and HMGB proteins induce a preferred topological configuration closely resembling the final one adopted in the UnwC3 topoisomer(thus, favoring our H3 hypothesis "The fitted supercoiled RSS model" pg26 in our original proposal). We needed to further test how RAG catalytic DNA cleavage is influenced by a topological constrained configuration, which sets the two RSSs in a certain fixed orientation. The rationale of the experiments described below is to constrain by length (short), torsion (supercoiled) and phase rotational orientation a pair of 12/23RSS, in small circles. RAG catalysing DNA cleavage on them we may test the validity of our H1,H2 and H3 hypotheses. It is known that DNA circles with their contour length smaller than double the persistence length of DNA(a)(a=50nm at 25^{o} C for dsDNA, equivalent with a length of 145bp) cannot adopt writhe, thus being confined in a single plane. For this, instead of IHF constrained circular DNAs we had to prepare small topoisomeric circles<300bp length which varied the intersignal distance IS

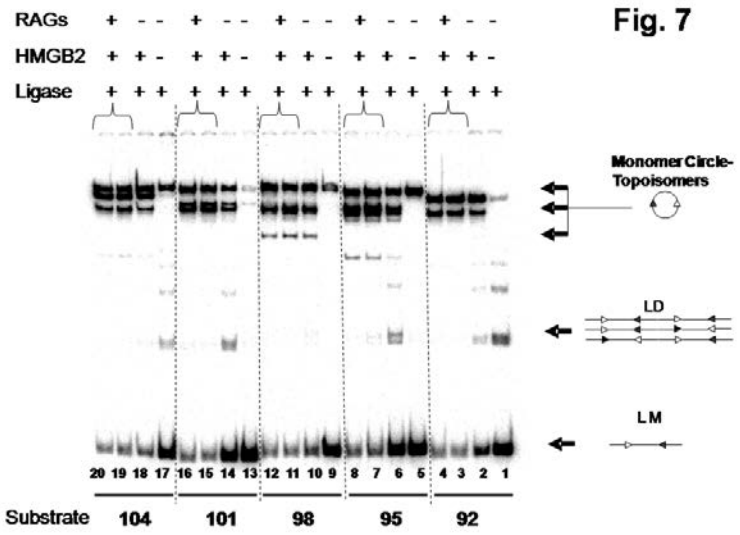


Fig. 7

between the 12 and 23-RSS's nonamers (in the deletional orientation they point inwards towards the intervening IS) in increasing increments of 3bp as follows: IS-77bp, 80bp, 83bp, 86bp, 89bp, 92bp, 95bp, 98bp, 101bp and 104bp. For each of these constructs the 2 arms emerging from the heptamers (heptamers are at oposite ends of an RSS from the nonamer) were set equal with 50bp each and ending into a XbaI site made overhang for appropriate annealing during ligation. We have previously studied (Ciubotaru et. al 2007)(1) the effect of T4 DNA ligase cyclization of these linear DNAs in the presence of a complete RAG mix (HMGB2ΔC, MBPcR1D708A and GST-RAG2) shown for IS92 to IS104 in Fig. 7. RAG synapsis enhances the effect of circularization of these DNAs and in conjunction with HMGB2 induces slow (Slw-relaxed) and fast mobility (F-overwound) topoisomeric species. T4 DNA ligase ring closure assays showed RAG preferentially synapses the two nonamers at a helical orientation of about 220^{o} in IS80 and IS101 with a cyclic helical phase dependence. This new strategy with small topoisomeric circles of various degree of supercoiling has the advantage, that could test the effects of nicking in the release of supercoiled energy without any need for an external nickase activity(Strategy D3A1 page 26 in the original proposal), simply comparing RAG catalytic efficiency on relaxed, and overwound topoisomers of the same construct .

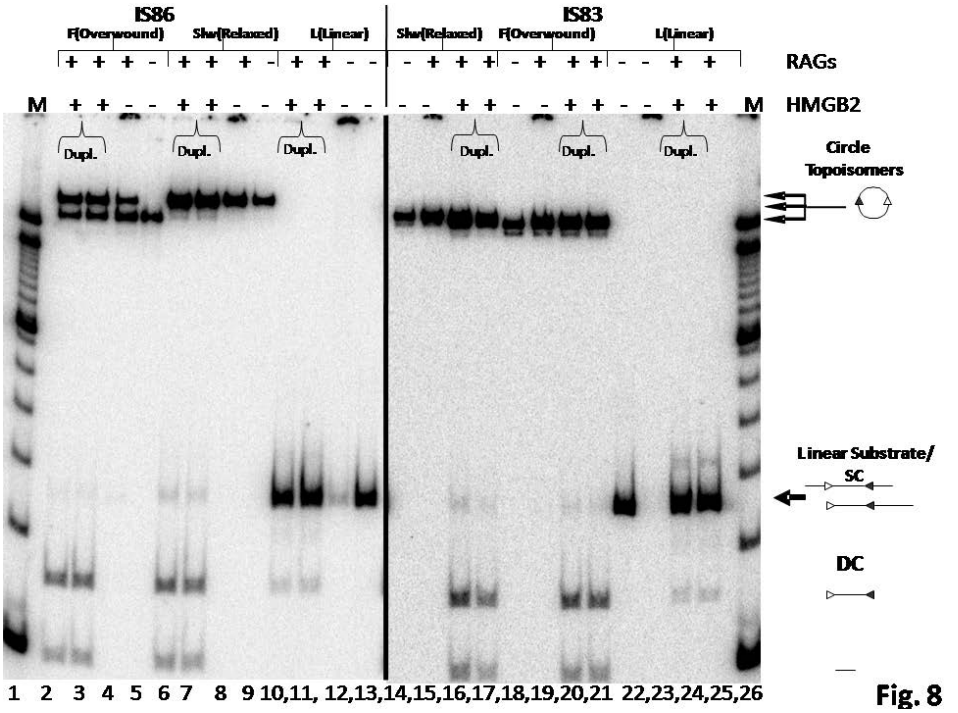


Fig. 8

Using preparative ring closure assays of each of these constructs in the presence of HMGB2 we gel isolated and purified each individual topoisomer and subjected them to RAG cleavage reaction with a complete RAG mix with wt MBPcRAG1(384-1008). The radiogram of a gel resolved RAG cleavage products of the linear and topoisomeric circles of IS83 and IS86 constructs is displayed in Fig. 8 (SC and DC designate double cleaved and single cleaved products). Two major observations strike at first glance. First, both types of circular topoisomers considerably enhance RAG coupled cleavage activity by comparison with that occurred with the linear substrates(in Fig.8 compare DC in lanes 2&3(duplicates) or 6&7 with 10&11 for IS86, or in lanes 16&17 or 20&21 with those in 24&25). Second, RAG intrinsic nicking activity converts the fast migrating overwound topoisomer F into that of slow migrating relaxed circle Slw even in absence of synapsis (see topoisomer interconvertion in lanes 4 and 19 with reactions without HMGB2). Similar reactions with those from Fig.8 were performed in 3 individual similar experiments, for each type of DNA

substrate and same observations were consistently found for all ten constructs. Before displaying and discussing the cumulative results of RAG coupled cleavage for all ten constructs we shall revisit the three hypotheses(H1-3) from the original project proposal (pg 26) and state their predictions with respect to such catalysis testing experiment. H1 the "DNA spring model " assumed that the entire supercoiled energy is converted into mechanical energy needed for DNA opening at the second hairpin formation step. If this were the case because nicking releases the torsional supercoiling energy in the step prior to hairpin formation H1 predicts that there will be no difference between the cleavage efficiencies of Slw vs. F topoisomers (as shown above nicking interconverts them with cleavage not yet occurring) of the same construct type. H2 "The protein-DNA elastic model" proposes a conversion of the whole supercoiled energy in a protein reconfiguration at the SC to synapsis transition, right before the hairpin step. Therefore, H2 proposes consistently higher cleavage efficiencies for all reactions with F than those with Slw topoisomers of the same DNA substrate type for all ten constructs. H3 "The fitted supercoiled RSS model" proposes that RAG requires a certain configuration of the RSS DNA which is easier for the RAG-HMGB complex to obtain from the supercoiled RSS than if the sites were in relaxed DNA form. The preferred helical phase orientations of the 12/23-RSS pairs displayed in the ring closure assays in our constructs should as well favor RAG coupled cleavage. In Fig. 9A are shown the cumulative coupled cleavage results obtained with the full complement of proteins for all constructs; with black linear DNAs, in red Slw- relaxed circles and in blue F-overwound topoisomers. In the same figure with dotted red line is depicted the trend of the ligation efficiencies/(synapsis formation) for all these constructs reported before (Ciubotaru et. al. 2007)(1). Most important we note that linear DNA substrates are cleaved almost with similar low 5-8% efficiency in all constructs. Unexpected for us in circular substrates IS77 to IS92 are generally better cleaved than those with longer inter RSS distance(IS 95 to IS104). However, for all constructs the circular topoisomers are cleaved remarkably better than their linear

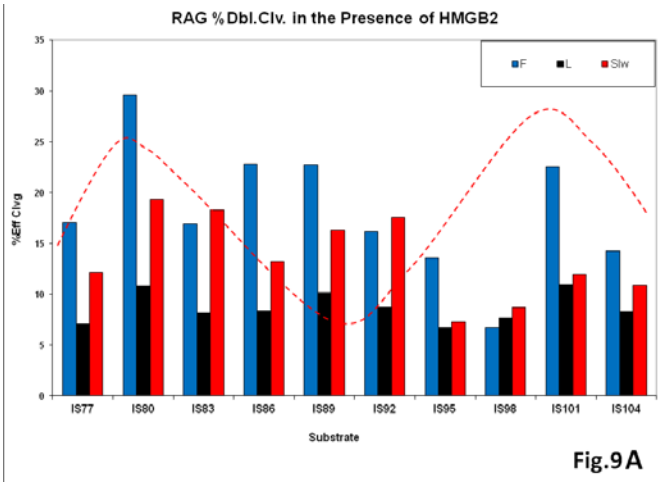


Fig.9A

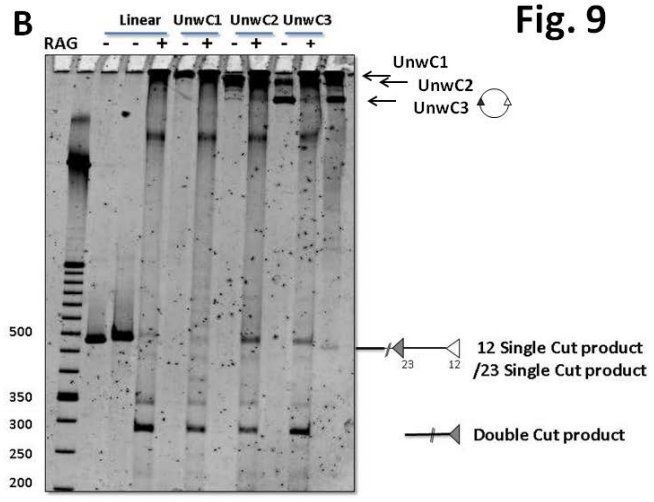
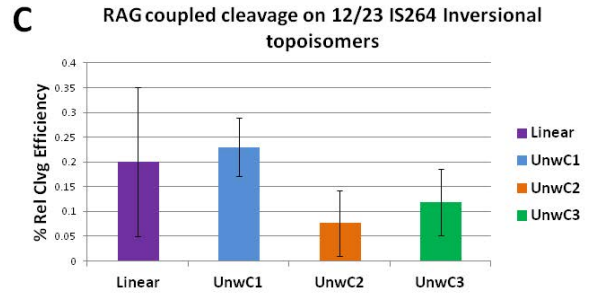
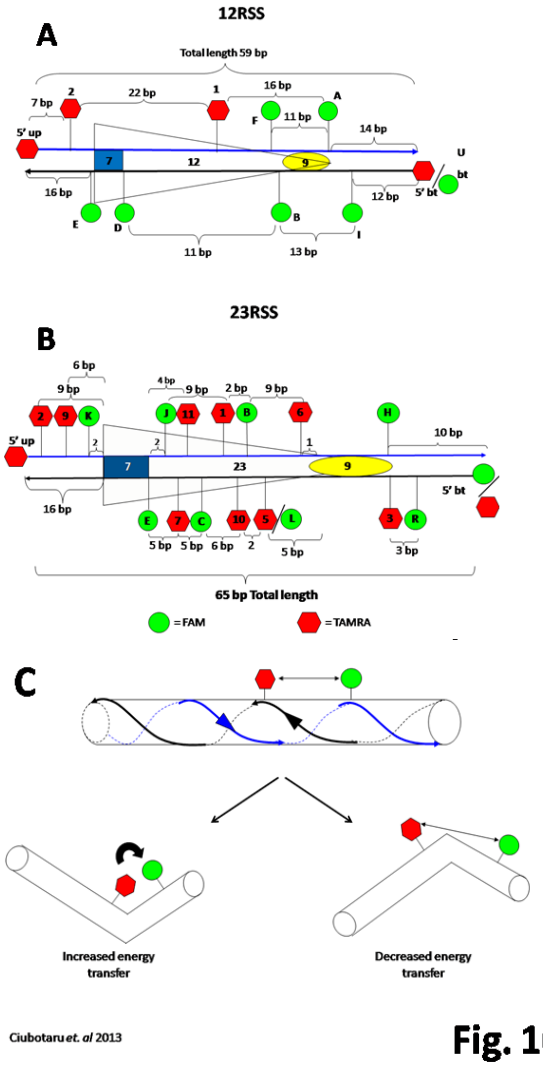


Fig. 9



counterparts. There are three constructs where RAG have little preference for F or Slw topoisomers (IS83, IS92, IS98), but for the other seven of them the overwound F topoisomers are better substrates for RAG cleavage. This observation invalidates both H1 and H2 hypothesis which as stated before require either consistent clevage equivalence of topoisomers (H1) or high catalytic preference only for overwound F circles (H2). Another strong argument against both H1 and H2 hypotheses is the high degree of variability in cleavage efficiency among the constructs. Although cleavage efficiency does not follow the exact strict "sinusoidal" periodicity displayed by the ligation assays (red dotted line), a similar trend is clearly present for the F overwound circles. Because the length of the intersignal IS distance affects the efficiency of RAG cleavage we have to interpret the IS80 and IS101 phase variants F results as local maxima. Taken together these findings offer most support for H3, showing that especially under supercoiling constraints RAG catalytic activity is guided by a preferred site topology, one that cannot be simply reproduced just by a writhe node sign of DNA crossing. Following these observations we performed in vitro RAG coupled cleavage experiments on a DNA substrate similar with IS264 (431bp in length presented in section D2A of the project proposal, Fig. 4A pg24) but having its pair of 12/23-RSSs in inversional orientation. The results of a representative gel are displayed in Fig.9B whereas those of the average of all 3 experiments quantified in Fig. 9C. These results simply show that negative supercoiling of UnwC1, UnwC2 and of UnwC3 circles of 12/23RSS inversional orientation largely either inhibit (UnwC2 and UnwC3) or do not change too much the efficiency of RAG cleavage from the results obtained on a linear substrate with this orientation of sites(UnwC1). This set of results are diametrically opposed from those obtained on the deletional orientation 12/23 sites IS 264 (D2A, Fig4A pg25) where UnwC3 substrate supports RAG cleavage at levels 4-5 fold higher than does its linear counterpart. These observations confirm our original hypothesis according to which deletional 12/23RSSs orientation would facilitate *in cis* PC assembly by different DNA sign supercoiling than that which would enhance it in DNA with inversional site configuration, whereas same type supercoiling would be stimulatory for one and inhibitory for the other. We decided to investigate in greater detail the intricacies and structural aspects of the two RSSs in the synaptic complex described at sections D3B1 of the proposal.



Ciubotaru *et. al* 2013

Fig. 10

<u>D3A4</u> The Configuration of 23 and 12-RSS in the synaptic complex. (objective D2A is addressed in the experimental section of the proposal at subsection D3A, and D3A4the last

addressing experimentally also the first part of the objective D2C)

D3A4. The Configuration of 23 and 12-RSS in the synaptic complex.

To be able to that address in *cis* versus in *trans* synapsis as well as to design metallohelicates binding to supercoiled RSSs one needs to fully understand the topological details of the two RSSs DNA paths inside the paired complex PC(or synaptic complex). Given the fact that a simple writhing node cannot reproduce the configuration of the two RSSs inside the synaptic complex we decided to use FRET methodology to investigate in greater detail these aspects both in 12 and 23-RSS. The FRET experiments were performed in collaboration with Dr. David G. Schatz's and Dr. Elizabeth Rhoades's groups from Yale University whereas the 3D molecular structural models were built in collaboration with Dr. Andrei J. Petrescu's group from the Institute of Biochemistry of The Romanian Academy from Bucharest. In a previous analysis, we examined the organization of the RSSs in the PC using fluorescence resonance energy transfer (FRET) and 12RSS and 23RSS oligonucleotide substrates in which the donor and acceptor fluorophores were placed in *trans* on the two different RSSs (1). The data suggested that the RSSs in the PC cross one another and are strongly bent, although we could not rule out a mixture of other, possibly less bent, configurations (Fig. 2 at pg. 19 of the original project proposal). Here, we set out to test the idea that the RSSs in the PC contain a large bend and to characterize in some detail the architecture of RSS DNA in this complex. Using FRET and numerous 23RSS and 12RSS oligonucleotide substrates labeled with both donor and acceptor fluorophores (in *cis*), we have obtained data that support large bends of both RSSs in the PC and which allow us to generate a working model for the trajectories of the two RSSs inside the PC.

We generated a series of consensus 12 or 23RSS oligonucleotide substrates containing a donor (6-carboxyfluoresceine (FAM)) and an acceptor (carboxytetramethylrhodamine (TAMRA)) fluorophore located in various positions relative to one another and to the 12 respectively 23RSS (depicted schematically in Fig. 10A and 10B). Fluorophores were coupled to a DNA base via a C_6 methylene linker. If RAG binding to the substrate bends or distorts the DNA in the region between the fluorophores, this can alter the interfluorophore distance and the efficiency of energy transfer. Depending on the location of the fluorophores and the magnitude and direction of the bend, the efficiency of energy transfer can increase or decrease compared to the free substrate (Fig. 10C). For each substrate, we measured the efficiency of energy transfer (E-FRET) in the absence or presence of proteins and partner RSS (to allow PC formation) and used this information to calculate the change in interfluorophore distance that occurred as a result of complex formation. The proteins used in most experiments were individually expressed MBP-RAG1c (aa 384-1008), GST-RAG2c (aa 1-387) and full length HMGB1.

While evidence for DNA bending was obtained with many of the substrates examined (see below), the $23RSS_dR2_a$ substrate (Fig.11-Fluorescence Emission Spectra) was particularly informative and was examined in greatest detail. In this substrate, the donor lies 3 bp 3' of the nonamer(yellow circle) and the acceptor 9 bp 5' of the heptamer (blue box) and are separated by a total of 51 bp, or a distance of 171-180Å in B form DNA (calculated using data from (32)). Since energy transfer is not detectable with these fluorophores at distances greater than about 90Å, no energy transfer was expected in the free substrate DNA. Indeed, as shown in Fig. 11A, the emission spectra for $23RSS_dR2_a$ (red line; indicated as "(d+a)") was similar to that of a

control reaction in which the donor and acceptor were in *trans* on different 23RSS substrates (an equimolar mixture of $23RSS_dR$ and $23RSS2_a$ substrates; solid black line; "(d) + (a)"). These spectra show a peak of emission at ~ 520 nm, as expected for the FAM donor, with no evidence of acceptor emission or quenching of donor emission when the fluorophores are in *cis* as compared to in *trans*. Note that all spectra have been corrected for residual acceptor fluorescence.

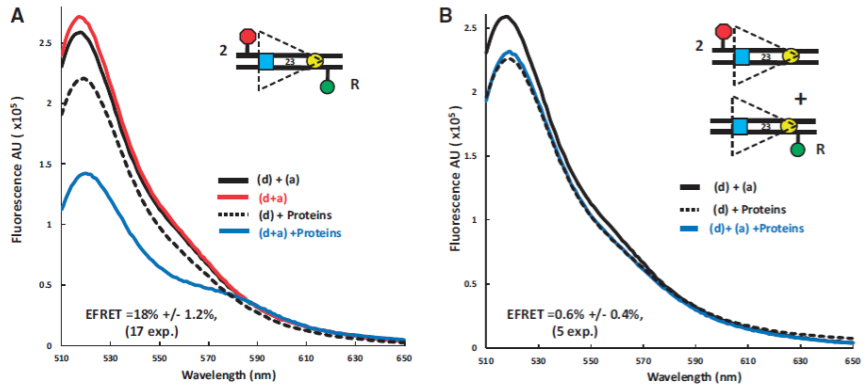
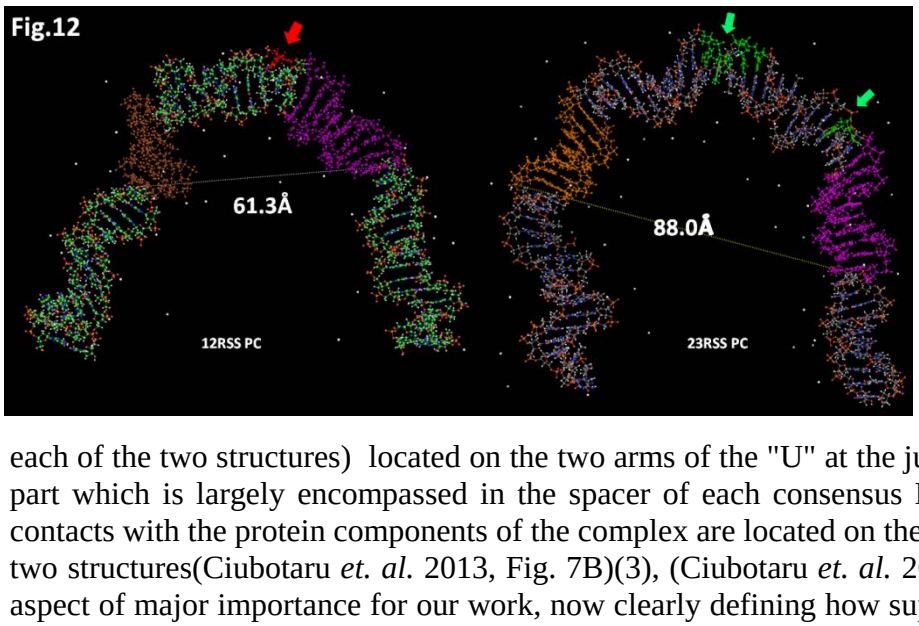
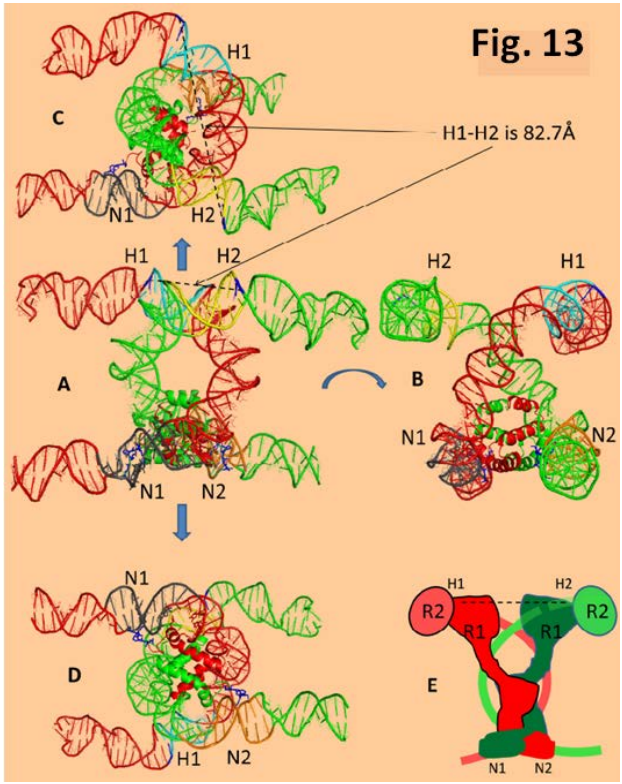


Fig. 11

Addition of the full complement of proteins and unlabeled 12RSS to the $23RSS_dR$ substrate that contains only the donor (dotted black line; "(d) + proteins") resulted in substantial quenching of donor fluorescence (Fig. 11A), which is due to interactions between the proteins and the FAM donor. When the doubly labeled $23RSS_dR2_a$ substrate was used (blue line; "(d+a) + Proteins"), we observed a further decrease in donor emission as well as acceptor sensitization (emission peak between 570-595 nm) (Fig. 11A, compare blue and dotted black lines). These changes in the spectra, which are directly attributable to the sensitization of the acceptor, indicate that energy transfer is occurring between the two fluorophores. This was consistently observed in 17 independent experiments, quantitation of which yielded an average FRET efficiency (E-FRET) of 18.0% ± 1.2%. This indicates that in the protein-DNA complex, the donor and acceptor are now separated by a distance less than 90Å. Using some simplifying assumptions (discussed below), we calculated the interfluorophore distance of $23RSS_dR2_a$ in the PC to be 71Å ± 10Å. This could only occur if the DNA undergoes substantial bending/distortion in the region between the two fluorophores. It was important to confirm that the observed energy transfer was occurring between the donor and acceptor on the same $23RSS_dR2_a$ substrate molecule (in *cis*), rather than by synapsis/aggregation of two (or more) $23RSS_dR2_a$ molecules, thereby allowing energy transfer in *trans*. 23/23 RSS synapsis was not anticipated to occur at substantial levels in our reactions given the preference of RAG/HMGB1 for 12/23 RSS synapsis and the 3-fold molar excess of the 12RSS. Nonetheless, numerous control reactions were performed to detect potential *trans* FRET using an equimolar mixture (7.5 nM each) of $23RSS_dR$ and $23RSS2_a$ under identical protein and partner 12RSS conditions (Fig. 11B). In no case was energy transfer detected in *trans*, results arguing that the energy transfer detected with the $23RSS_dR2_a$ substrate occurs in *cis*. Similar experiments were performed for other 14 *cis* pairs of fluorophores for 23-RSS and 16 *cis* pairs for 12-RSS mapping, shown in Fig. 10A and B. The detected interfluorophore distances in the PC(synaptic complex) and single complex SC(only for 12-RSS) were used to build structural molecular models of 23 RSS inside the PC and 12-RSS both in the PC and SC. The complete work describing all FRET measurements for the mapping of the 23-RSS inside only the PC and for 12-RSS in the PC and 12SC were published Ciubotaru *et. al.* 2013 and Ciubotaru *et. al.* 2015 in Nucleic Acids Research, Oxford University Press (3,8).

Fig. 12 shows by contrast the modeled structures of 23-RSS and 12-RSS in the PC as and with arrows are indicated the sites of DNAase I hypersensibility detected by Swanson et al 2002(4), where the core of the RAG-HMGB1 imparts the maximal bending curvature onto the sugar-phosphate backbones of the RSSs. Two major conclusions can be drawn from these modeled structures. For both 12 and 23-RSSs the DNAs are considerably bent in an upside-down"U" letter shape with the consensus heptamer and nonamer elements (colored in each of the two structures) located on the two arms of the "U" at the junctions with their arching part which is largely encompassed in the spacer of each consensus DNA. On both DNAs the contacts with the protein components of the complex are located on the concavities of each of the two structures(Ciubotaru *et. al.* 2013, Fig. 7B)(3), (Ciubotaru *et. al.* 2015, Fig. 7A and B)(8) an aspect of major importance for our work, now clearly defining how supercoiled looped RSSs are brought into close proximity in the PC by the RAG-HMGB complex. These aspects that we discovered first (our 2015 paper (8) was published in January) were later proposed to be incorporated in the model of how the resolved high resolution crystal RAG1-RAG2 recombinase alone structure would bind bent RSSs(10) and in Nov. of 2015 confirmed when the first high resolution structure of EM of the PC was published (11).





Although these structures add tremendously to our current understanding of the organisation of the RAG-RSS complexes with short free oligonucleotidic DNA, yet much remains to be uncovered with what happens in vivo where the recombinase binds to native negatively supercoiled chromosomal DNA from the V(D)J loci. To help and bring a more insightful perspective into the relationships between the two consensus RSSs and RAG in the PC complexes in 2014 we "in silico" anchored the nonamers of two PC complexed 23RSSs to the tight array of Aminoacid-DNA Base interactions described in the GGRPR DNA binding motifs of the of the 75 aminoacids peptide Nonamer Binding Domain NBD of RAG1 (389-456)(F. F. Yin et. al. 2009, the only high resolution structure of RAG with DNA available in 2014)dimer (2).

The structures were further in silico subjected to molecular dynamics and solvation/cationic dispersion minimization constraints. The resulted model published by us as a review in 2014(Ciubotaru *et. al.* 2014)(7), is displayed in Fig. 13 with N letters locating each of the two RSSs nonamers and H the heptamers. Because the modeled structure shown in Fig. 13 comes from just a small domain of RAG1, was built in 2014 with just 23RSS bent the only RSS structure available at that time, certainly assumes an oversimplified view of the PC complex. It is an alternative model to that of Ru *et. al.* 2015(11) which does account for the plectonemic possible PC assembly which could use supercoiled RSSs, thus for us does provide insightful clues helping us understanding the key results of IS264 ligation kinetics and RAG catalysis on constrained topoisomeric DNAs (presented in the previous part at pg.4-9 of this report). First, because the core of the RAG-HMGB protein complex seems to be at least partially wrapped outside the plectonemes of the two RSS, it considers a direct contact between their DNA spacers just as if RAG would simply bind a preformed writhed supercoiled positive or negative node.We ephasize that this situation is an alternative compatible with both RAG high resolution structures resolved in 2015(10,11) but considering the RSSs tightly wrapped instead of being relaxed as short free oligonucleotides. It explains though why RAG may preferentially bind with better affinity individually to each supercoiled RSS forming 12 or 23SC, than it would do to relaxed DNA-RSS. The torsion of the supercoil forces the DNA to adopt a rather bent configuration (closely resembling the one in its final PC configuration). However, the PC formation is not facilitated by the mere juxtaposition of the two RSSs in a single supercoiled node (as we wrongly hypothesized in our original project proposal, see pg.22, 25), but is rather mediated by protein-protein interactions that tether two complementary supercoiled SCs RSSs in the synaptosome. It is exactly this optimal orientation and in register adjacency of these protein-protein interactions (after SC complex formation) that facilitates better synapsis and catalysis in constructs such as IS80 or IS101. Because the two RSSs display considerable structural difference (as shown in Fig.12) our model raises an additional question of whether the synaptosome forms isotropically or anisotropically (requires a unique orientation of its elements 12 on top 23 at the bottom or vice versa). We decided to test this question using transcription as a supercoil sign regulator (see below and section D2C-D3C).

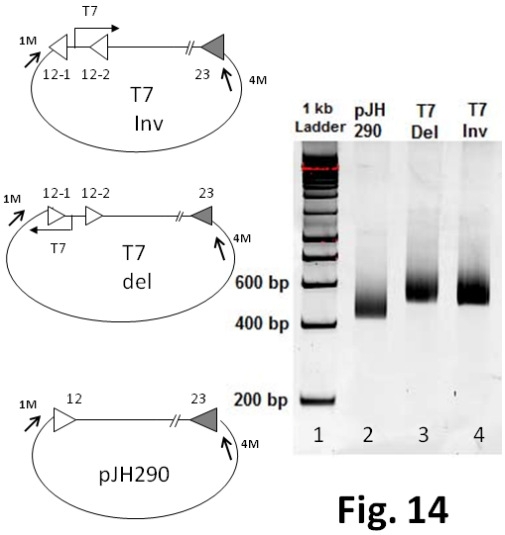


Fig. 14

<u>Obj3 -D3C. Prove transcription driven DNA supercoiling to be the selective factor in favoring PC formation *in cis*.</u>

(objective D2C is addressed in the experimental section of the proposal at subsection D3C of page 29).

Our hypothesis in the project proposal stated that transcription from promoters in the proximity of V, D or J segments embodies their RSSs into active supercoils which can activate them for recombination. To test this idea we proposed a dual intracellular/in vitro in test tube set of parallel experiments based upon a substrate DNA vector with two identical 12RSSs to be located adjacent to (reporters) a heterologous T7 phage promoter as an activator, in conjunction with a distal complementary constitutive 23RSS. We slightly modified the design of the construct

presented in our original project proposal (pg. 29, Fig.7), to better cope with our topoisomer and FRET findings. Given the fact that 23-RSS has lower affinity for RAG binding and requires the presence of HMGB proteins in our new construct both the reporter RSSs were chosen to be two identical 12-RSSs and the constitutive to be a 23-RSSs to pair in PC as partner to either one of the reporter 12s. We used as backbone the pJH290 vector were near its upstream 12-RSS we inserted another one identical with its original one and a T7 phage RNA polymerase promoter in between them. The construct was built with the 2x 12-RSSs in deletional and inversional orientation with respect to the distal 23-RSS, as shown in Fig. 14 constructs denoted 2x12RSST7Del or Inv. To test the effect of synaptosome orientation both 12-RSSs were positioned in close proximity to the T7 promoter such that transcription drives the one 12-RSS downstream its course in positive supercoiling(12-2 for T7Inv construct or 12-1 in T7 del) whereas the 12-RSS upstream to it will be negatively supercoiled (12-1 for T7Inv construct and 12-2 in T7del). To stop transcription we used a oop terminator 182bp downstream its promoter for the inversional orientation(in the middle of intersignal region) whereas for the deletional one a standard T7 terminator was inserted 108bp in the wake of 12RSS-1(distal from the 23 constitutive). In the right panel of Fig14 are shown amplified PCR products using a pair of 2 common designated primers designated1M and 4M, flanking approximately 50bp upstream of each of the distal 12-RSS1 reporter and downstream 23RSS on both constructs. The intracellular recombination assays were performed by cotransfecting 293T HEK human fibroblasts either with the 2x12RSS T7Inv or with 2x12RSS T7Del substrate, pEBB vectors expressing full length RAG1, RAG2 and a vector pAR3132T7RNP_NLS(6) expressing the T7 phage RNA polymerase. We will start the description of the experiments first with those obtained from the

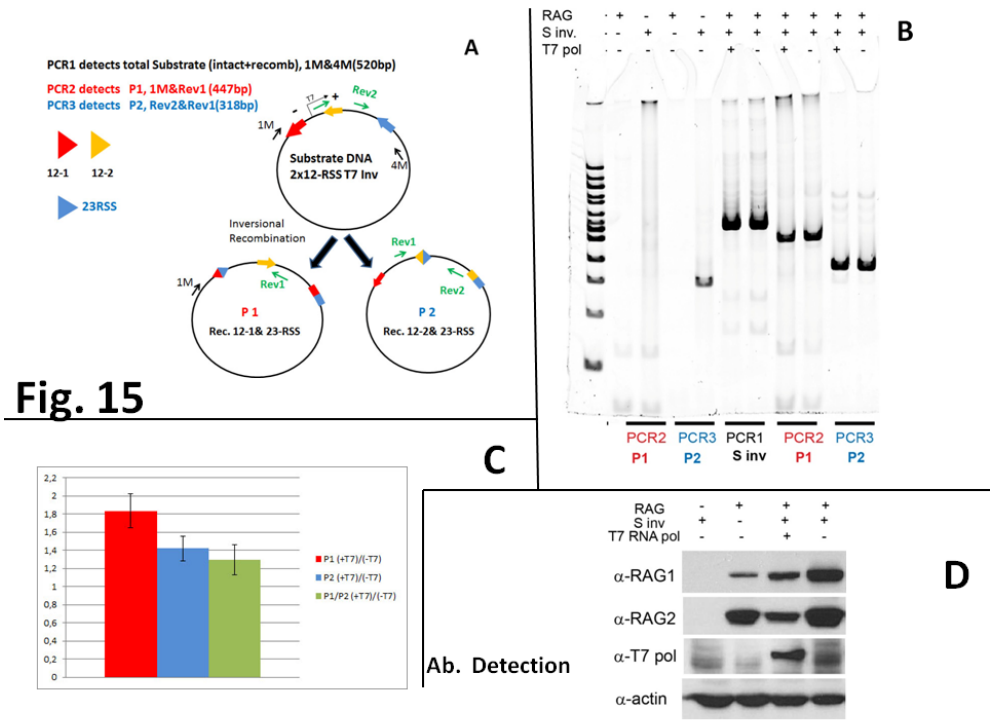
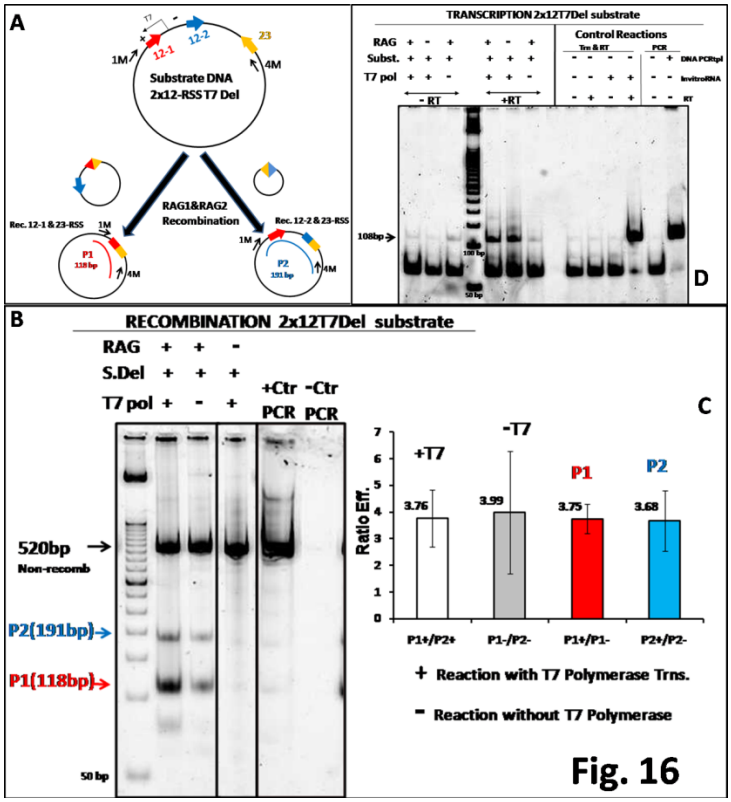


Fig. 15

inversional substrate (Fig. 15) and then we will contrast these results with those obtained from the deletional one (Fig. 16). Fig.15A depicts the process of recombination mediated by RAG undergone on the inversional and on Fig. 16A for the deletional substrate. These recombinations occur by pairing the constitutive 23RSS either with 12RSS-1(distal reporter with respect to the 23RSS partner, P1) or with 12-2 RSS(proximal to 23,P2), and are evidenced using in individualized PCR assays(primers positioned in Fig.15A and 16A diagrams) to detect each amplification products; P1(12-1&23) and P2(12-2&23). For both types of substrate three individual transfections and related control experiments were undertaken each analyzed from 5 distinct sets of PCR amplifications whose reaction products upon separation on 5% PAGE and SybrGreen staining were quantified using a Typhoon Fluorescent scanner(Fig.15Band 16B). In Fig.15Band 16B are illustrated representative gels on which the PCR amplification products are visualized (including the control reactions) reactions templated by plasmids extracted from the cells which were maintained in culture 48h. post-transfections. For both substrates the results of these quantifications are cumulatively shown as average values(and their SDs in error bars) in histograms from Fig.15C inversional and 16C for the deletional substrate. The results are first expressed as a relative ratio between detected products P1(red) or P2(blue) in the presence (+T7) or absence (-T7) of T7 RNA polymerase (P1+/P1-, P2+/P2-). Second, in the case of the inversional substrate the green bar reflects a ratio between two other ratios those between the two products (P1/P2) in the presence of T7 poly. divided by the ratio between products obtained in the reactions in the absence of T7, (P1/P2 "+T7")/(P1/2 "-T7")(Fig.15C). In the case of the deletional substrate each of these two ratios is individually represented P1+/P2+ in the presence(+T7)(white), or P1-/P2- in the absence of transcription (+T7)(grey)(Fig.16C). In the case of the inversional substrate PCR1 detects reference unrecombined (1M and 4M) signal from the vector, PCR2 evidences recombination product P1(between the distal 12-1 and 23RSS) whereas PCR3 the product P2(between the proximal 12-2 and 23RSS). In the case of the substrate with deletional orietation of the RSSs because all PCR products are evidenced with the same two primers(1M and 4M), a prerequisite normalization of the total amount of substrate vector template extracted from each cellular transfected lot was realized using a set of calibration PCR reactions with two primers which amplify in the region of the substrate not subjected to recombination(data not shown).Fig. 15D displays the immunoblot analysis showing the presence or absence of appropriate protein expression levels after 48hrs. post transfection. Fig. 15E and 16D display the RT-PCR reactions made from cell RNA extracts dilutions treated with DNAase

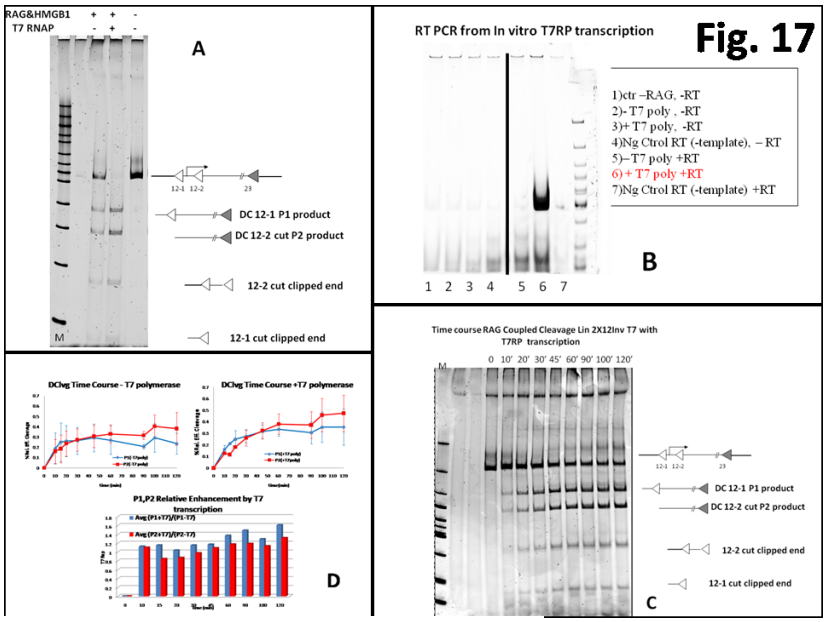


I, followed by +/-Reverse Transcription and PCR amplification with a set of primers flanking the T7 transcript from the promoter level up to the proximity of each terminator. One can clearly evidence specific T7 transcription initiated from the substrate promoter, in cells transfected with T7 RNA polymerase(+T7) versus those grown in absence of this vector transfection(-T7). In Fig. 15F are illustrated the effects of vector template dilutions used in the PCR reactions for the assessment of the quantifications displayed in Fig. 15C and 16C.

Two main conclusions can be drawn from the analysis of the experiments shown in Fig.15 and 16. **1)**First, very consistently and reproducibly for both orientations recombination occurs more efficiently at 12-RSS1(distal with respect to 23RSS, evidenced by P1.) than that detected at 12-RSS2(proximal 12 corresponding to P2) independent from the presence or absence of transcription(compare the wells labeled with PCR2-P1(red) with those identifying PCR3-P2(blue) in Fig.15D and F for the inversional substrate, and wells 2 and 3 from Fig. 16B(products indicated by arrows) for 2x12RSST7Del. For the inversional substrate the average value from all quantifications for the P1/P2 ratio is 1.54 +/- 0.2 in the absence of T7 transcription and 1.92+/- 0.19 in its presence, whereas the same ratio for the deletional substrate is 3.99+/-1.2 at the transfections without T7 polymerase and 3.76+/-0.6 for those with T7 transcription. This observation clearly shows that RAG preference for 12RSS1 versus 12RSS2 is more accentuated for the deletional substrate(where the values of the mentioned ratios are almost double from those values obtained on inversional substrate experiments). However, only in the case of the inversional substrate experiments a slight induced preference for recombination at 12RSS1 versus 12RSS2 is facilitated by transcription. The extreme degree of DNA bending required for RAG pairing both at the level of each RSS as well as for the intervening inter-site DNA combined with the known plasticity for bending of short B double stranded DNA(reflected by its persistence length of 150bp) may explain why two more distantly located sites like 12-1/23RSS may be favored for pairing versus two sites located rather closely(12-2/23RSS) (In 2x12RSST7Inv, 12RSS1 is 358bp from the 23RSS(nonamer-nonamer, n-n) and 12RSS2 is at 285bp(n-n) from it; similarly in 2x12RSST7Del, 12RSS1 is located at 335bp from the 23RSS(n-n), whereas 12RSS2 is only 262bp far from 23RSS(n-n)). **2)**T7 transcription enhances equivalently P1 and P2 recombination in the deletional substrate to a level almost 3.7 fold higher than that detected the reactions in absence of it(in 2x12RSST7Del the ratio P1+/P1- is 3.75+/-0.28 and P2+/P2- is 3.68+/- 1.2). Although in the inversional substrate T7 transcription has also a stimulatory effect for both recombination products, here the effect is significantly higher for P1 (P1+/P1- is 1.82+/-0.21) than that detected for P2(P2+/P2- is 1.4+/- 0.17). Remarkably these results show that T7 polymerase transcribing not only that it does not sterically hinder in any way the access of RAG to its RSSs(located within 75bp proximity from the transcription initiation site) but enhances its catalysis to them. **This notable effect is a direct facilitation of RAG activity induced by transcription, and must not be confused with the described accessibility phenomenon by which the local chromatin modifications are fostered by the endogenous RNA polymerase II action.**

To decipher the reaction mechanism by which recombination is activated by transcription we also have investigated in vitro in the test tube the effect of T7 transcription on RAG induced coupled cleavage reactions first on both linear 520bp 2x12RSS T7Inv and Del substrates and then on circular DNAs of different topologies. The coupled cleavage reactions were set in presence or absence of T7RNA polymerase under 8.5mM $MgCl_2$, and 2mM rNTPs, MBP-

cRAG1(384-1040), MBP-cRAG2(1-387)coexpressed and HMGB1. Each cleavage reaction is performed in duplicate with one reaction being analyzed after deproteinization on native PAGE with Sybr gold staining for RAG cleavage products, while the other is subjected to DNAaseI digestion, RNA extraction followed by Reverse transcription with a specific transcript complementary primer and the RT products further amplified by PCR. Because the results obtained on linear templates of both orientations are similar we will show in this report only those resulted from the inversional one. Fig. 17A shows the slightly stimulative effect of T7RNAP transcription activating and equivalently the level of P1 and P2 RAG coupled cleavage after 2hrs at 37°C (compare second -T7 RNAP lane with the third +T7RNAP). Fig. 17B shows the results of RT- PCR of these reactions where we included controls for absence of substrate (lanes 1&7), -RT(lanes 2&3) or +RT(lanes 4&5) with clear evidence in lane 6 of the expected transcript. Because the transcribed substrate is in linear configuration one does not expect considerable topological effects to accumulate near T7 promoter more than transiently, thus P1 and P2 must have under such circumstance similar cleavage efficiencies and kinetics, effects proved nicely Fig. 17C and D. We also investigated the time course effect of T7RNAP transcription on RAG cleavage to show transcription parallels RAG activity up to 90 minutes and then drastically decreasing for the last half hour of the experiment due to a scarcity in DNA template which was significantly processed by RAG cleavage(Fig. 18).



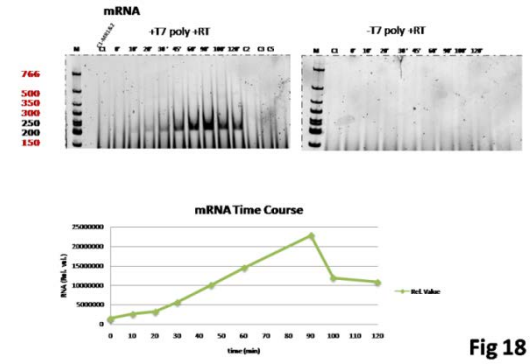


Fig 18

The next step was to similarly *in vitro* verify the T7 transcription local effect near the reporter 12RSSSs on RAG, using this time minicircles in which the 12/23 are present in both functional orientations (deletional and inversional Obj.3 described at D3C pg29-31 of the project proposal). To obtain such minicircles with various degrees of supercoiling we applied the same ligation strategy using T4DNA Ligase in the presence or absence of HMGB1 onto linear constructs with compatible XbaI cohesive ends of 520bp (2x12T7 Inv/Del), followed by PAGE separation and gel purification of the various resulted topoisomers. Fig. 19A depicts the steady state RAG coupled cleavage results +/-T7 RNA poly. transcription performed on four such negatively supercoiled topoisomers (UnwC1-UnwC4) as well as on the original linear substrate 2x12T7Del. Unlike the presented results from the original project proposal (Fig.4B, pg.25) where the

separation of the reaction cleavage products was done on acrylamide gels casted in squencing format (33x35cm) this time the PAGE separation was done on conventional gels casted in smaller format ProteanII XL Biorad (18.5x20cm), the only one available on Romanian market.

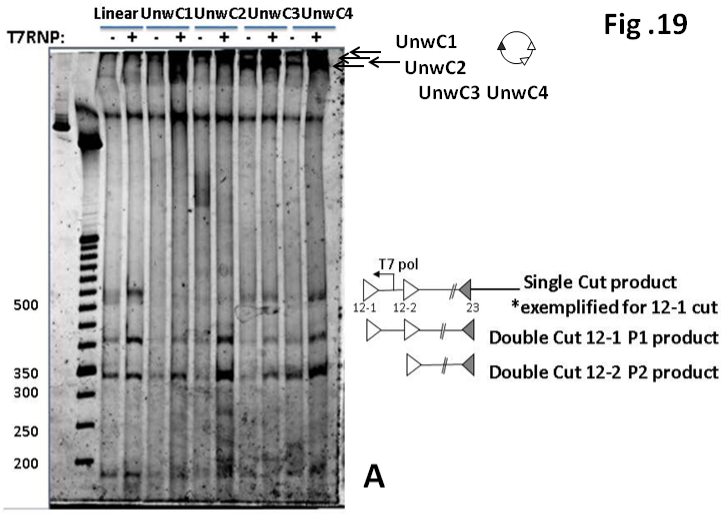
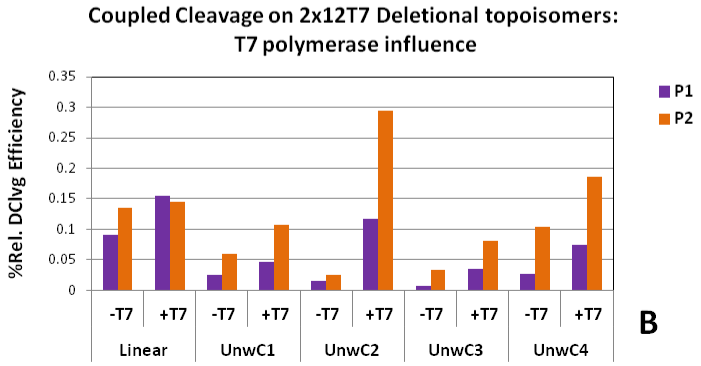


Fig .19



For this reason the gel cannot show distinctly on the same electrophoretic migration the difference of intact circle mobilities as well as the linear small products(P1, P2, SC) resulted from their RAG cleavage(see the arrows depicting the original positions UnwC1-UnwC4 on the top of the wells of the gel). Hence, the amount of topoisomers added in each reaction were normalized according to fluorescence readings (Qubit- fluorescence Life Technologies). Fig. 19B depicts the histograms resulted from the quantifications of the products obtained from these reactions. For the linear DNA the presence of T7RNAP induces a slight stimulation of cleavage at both reporter 12RSSs, consistent with the similar results shown for the inversional substrate in Fig.17. Remarkably topoisomers UnwC2 and to a lesser extent UnwC4 display a considerable enhancement of RAG coupled cleavage caused by T7 transcription with the maximal effect visible for 12RSS2 reporter(proximal to 23RSS generating P2; mounting to almost ten-fold P2 and 3.5xP1 fold increase upon transcription on UnwC2). The effect was consistently reproduced to the same topoisomers in three independent experiments. Single cleavage to one RSS(SC -single cut-which requires in trans PC on two DNAs) can be detected only for UnwC4 and linear DNA. **In conclusion such high enhancement of P2 by T7 transcription at UnwC2 can be explained by our original hypothesis that the negative supercoiling created behind the marching polymerase in the deletional orientation helps bending and reconfigures the proximal 12RSS2 making it more amenable for synapsis with its 23 partner.** Note though how subtle is this effect detection in terms of finding the appropriate energetic topoisomeric unwinding value to match such enhancement requirements(UnwC1 or UnwC3 do not support fully such effects).

We tested the effect T7 transcription has on RAG coupled cleavage with 2x12T7Inv circular substrates obtained as described above and the results are displayed in Fig. 20. On this substrate topoisomers one observes that coupled cleavage(supporting in cis site synapsis) is either strongly(UnwC1, UnwC2) or moderately (UnwC3) inhibited whereas in the case of UnwC2 completely blocked. Remarkably for UnwC3(Fig.20A) the RAG reaction is almost

completely funneled towards in trans PC formation which almost exclusively generates SC(see also histogram of quantified products Fig.20C). Interestingly though the observed effects seem unmodified by the simultaneous presence or absence of transcription effects which are probably attributable to the plectonemic topoisomer configuration, nonproductive for synapsis in this substrate even if T7 driven transcription is well supported in these reactions (data not shown). **Despite the absence of T7 transcription modulation of RAG activity on 2x12T7Inv minicircles, the experiments shown in Fig.20 are fully supportive to our original hypothesis that certain DNA topologies could be inhibitory/blocking to PC formation whereas allowing only recombination in trans with respect to DNA.**

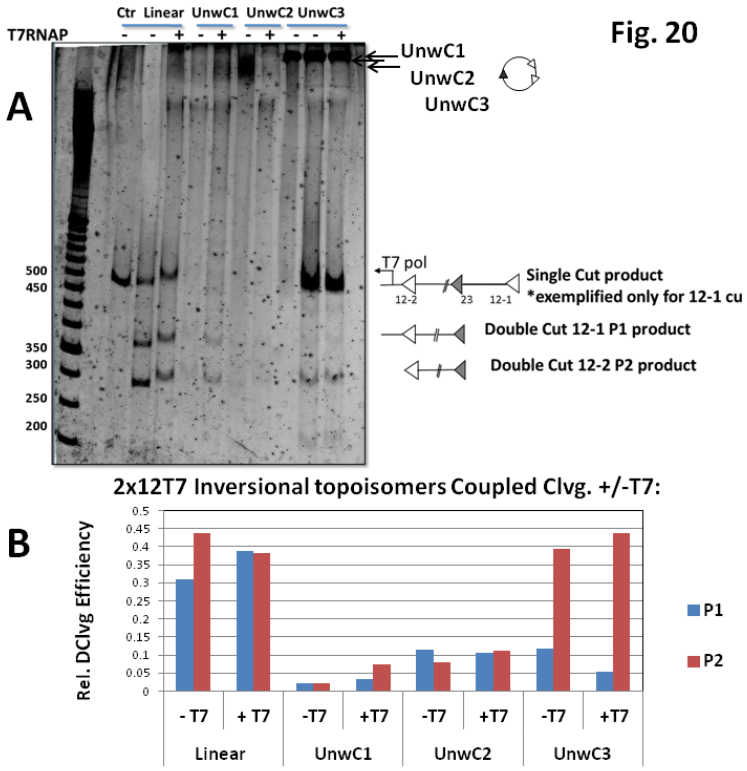
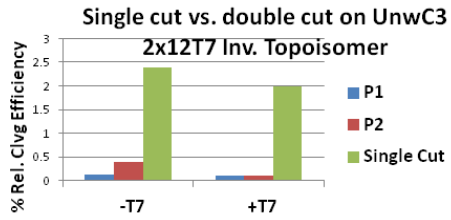


Fig. 20



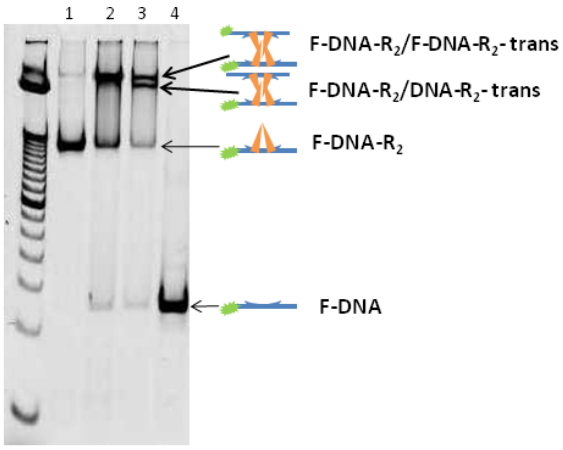


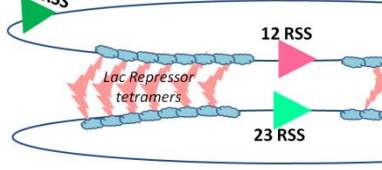
Fig. 21

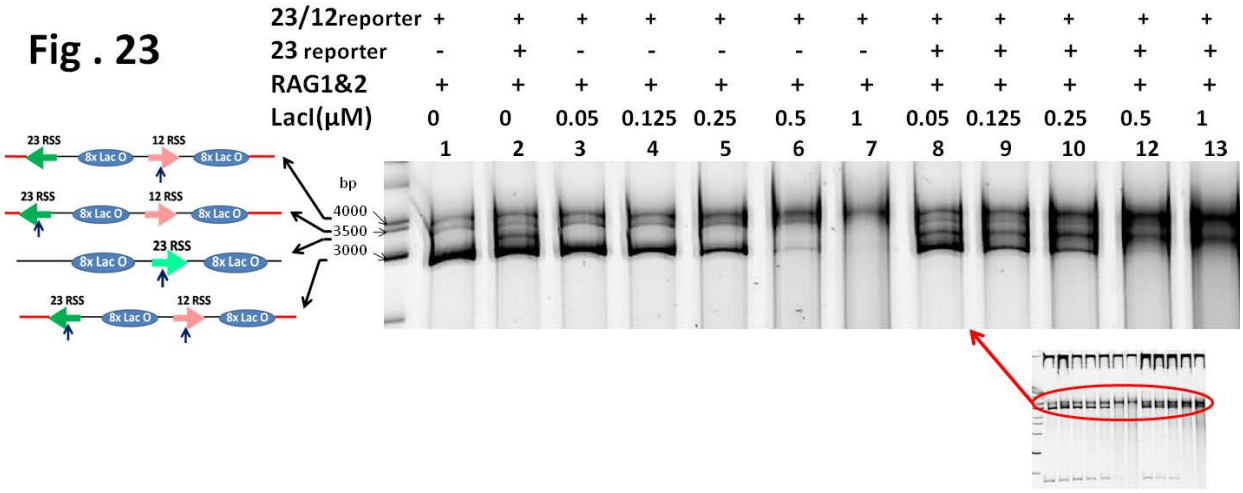
Ob.2 D3B-1. LacI repressor DNA operator site interactions modulate DNA-DNA tethering in trans.

As we described at pages 31-32 of our proposal to fulfill the IIndOb.it was needed to built an artificial system to anchor two distinct DNA macromolecules via an accessory protein capable to bind cooperatively in trans to both of them, hence bridging them in close proximity(simultaneous binding of two protein protomers onto two distinct DNa molecules). Our attempts to use the exact same strategy as that described in our original proposal using the 434 lambdoid phage repressor mutant D173G was not fully productive, since this protein although does engage in binding onto two separate operators located *in trans* it does so with low cooperativity and hence with low tethering affinity. For this reason to fulfill our Ob.4 we decided to change instead of 434R and we have replaced the accessory protein with the repressor of E. coli lactose operon, LacI and its DNA binding

operator Op1 consensus sequence. First to test this interaction we have purified LacI repressor from an E. coli transformed with a lacZ expressing vector.To build the small Nr polyoperator sequence elements which must be juxtaposed in tandem repeats, the same symmetric Op1 element(see pg31-32 D3D project proposal)was assebled by ligation of compatible cohesive ends oligonucleotides, one of which was labeled at its 5' end with Alexa 484 fluorophore. Fig. 21 shows an EMSA(electrophoretic mobility shift assay) experiment which evidences first the dimer LacI binding by preincubation to one operator site(DNA-R2-dimer 1st well, at a molar ratio 1.25:1 LacI:OpDNA). When the LacI binding to the fluorophore labeled OpDNA is obtained by adding an excess 1:1.6 LacI:OpDNA of Operator DNA(OpDNA fully labeled with fluorophores on both DNA strands -2nd well) or half labeled OpDNA(3rd well, one fluor singly labeled OpDNA mixed with unlabeled OpDNA) the major protein-DNA complex obtained, bridges *in trans* two DNAs latched by a LacI tetramer. These rather encouraging experiments allowed us to switch to LacI repressor strategy and build our polyoperator tethering system sequences described in the next session.

D3B-2. The LacI-OpDNA interaction guides V(D)J recombination *in trans.* Given the successful outcome of the tetramer LacI DNA tethering experiments presented in Fig. 21 next we have built plasmids containing 8 adjacently in tandem located operator sequences each such octamer 8xOp flanking either a 12 or a 23-RSS reporter(Fig. 22, or Ob4 D3D Fig.8 pg 31 from the proposal) either on each side or just on one side. Because we intended to show that LacI can inhibit RAG coupled cleavage *in cis* whenever is engaged in cooperative extended tetrameric binding onto 2x DNAs each with a 8x Op, we also added to the reporter 12RSSr flanked by two 8xOp a constitutive 23RSSc as a cis partner(See Fig.22 upper construct). The design of the experiments meant to test the *in trans* anchoring effect caused by tetramer repressors was described in detail at Ob4 D3D pg. 31-32 and illustrated on Fig.8 of the project proposal, thus we will just briefly outline here their logistic. First set of experiments tests if the LacI tetramers bound onto 4 *in trans* 8xOp tether appropriately in register 12 and 23RSSr such as to allow RAG to perform an *in trans* catalysis(versus the natural one occurring *in cis* between 12RSSr and 23RSSc, without LacI intervention). For this one has





to perform first in vitro RAG coupled cleavage on the two plasmid constructs shown in Fig. 22 in

the presence or absence of an appropriate concentration of purified LacI repressor. Fig. 23 shows a blown up image of the gel used for DNA RAG cleavage product separation(the zoomed area of interest is depicted from the whole gel in the right corner), products generated in 13 distinct reactions according to the legend shown above the gel. The left hand diagram points by arrows the positions of RAG cleavage on each of the four identifiable products: single cut at 12-RSSr (the lowest electrophoretic mobility product), 23-RSSc (immediately under the 12RSSrSC in mobility), 23-RSSr SC(with intermediary mobility), and the one of coupled *in cis* cleavage 12-RSSr & 23-RSSc (the smallest band with the highest mobility). Notably, *in trans* RAG cleavage requires besides the presence of both types of plasmid constructs (one with 12RSSr and the other with 23RSSr) also that of the docking accessory LacI repressor(compare wells 1,3,4,5,7 with those from reaction 2 and the set of reactions 8 to 13). We also have measured how the LacI concentration affects the degree of 12RSSr and 23RSSr mediated proximity(compare the reactions set 3 to 7 with those shown at 8 to 13. Since the 12RSSr cleavage levels are due to a summative effect from the *in trans* synapsis with 23RSSr and *in cis* pairing with 23RSSc its interpretation must be correlated with the other two clevage products (coupled cleavage and SC at 23RSSr). The RAG single cleavage product at 23-RSSr selectively indicates only the efficiency of in trans pairing and represents the product whose level reproduces with highest accuracy the LacI docking of the two types of constructs and consequently RAG in trans cleavage. To better follow the correlative effects of all the reaction products we have quantified the gel band fluorescence of 3 reactions similar with those depicted in Fig.23 and represented their relative increase in intensity(with respect to the control reaction in absence of LacI) as a function of LacI concentration present in each reaction Fig.24. In blue are shown the curves of the control reactions in which the partner RSS required for synapsis is omitted, whereas the red traces show the reactions containing both partner sites 12RSSr&23RSSr in trans. **Fig. 24C shows beyond any doubt the remarkable effect LacI has in mediating the *in trans* RAG cleavage at the reporter 23RSS via docking of the two poly8x operators. Within a range of concentrations between 0.1-1μM of LacI the single cleavage at 23RSS r increases almost five folds, increase which occurs only when the appropriate in trans 12RSSr partner is present in the reaction(compare the red with the blue trace in Fig.24C).** As pointed above one can correlatively follow to some extent the same effect reflected as well at the partner 12RSS reporter(Fig. 24A) but in this case *in cis* coupled cleavage also contributes to the reported product relative increase. In Fig. 24B the relative decrease in coupled cleavage between 12RSSr &23RSSc is represented as a function of LacI repressor concentration, an effect which is

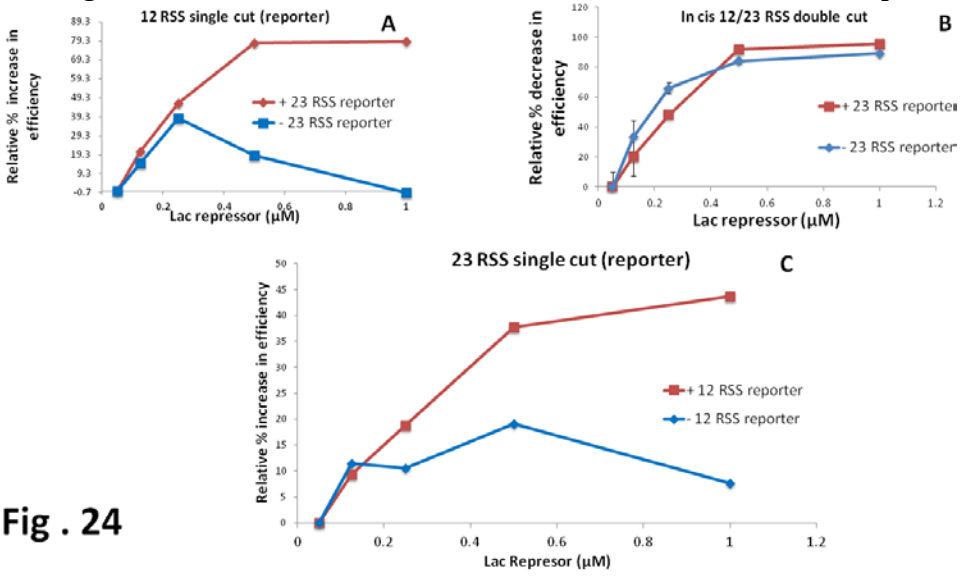
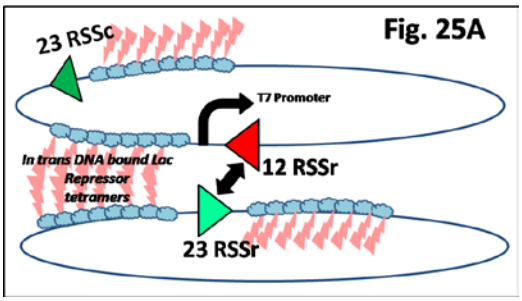
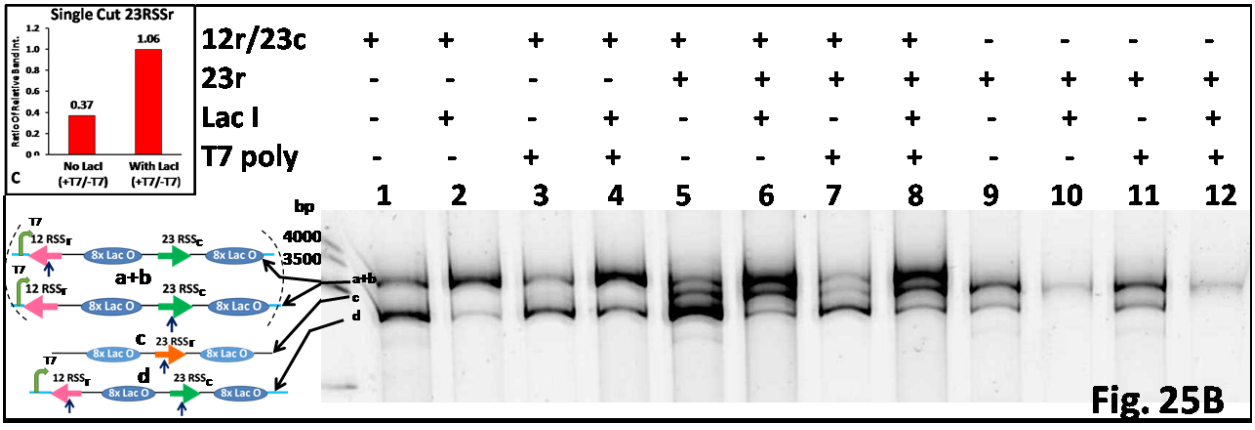


Fig . 24

expected since repressor docking also inhibits in cis synapsis regardless of whether the partner reporter 23RSS is present or not in the reaction. We attribute the LacI inhibitory effect in absence of 23RSS perhaps to a repressor induced dramatic bending in the inter-RSS intervening DNA where the poly 8xOp is located effect, which may decrease sterically the effective concentration of the two RSSs on the same DNA reducing their pairing chance. The major goal of these experiments has been reached, namely to reproduce in vitro conditions that drive RAG synapsis *in trans* thus mimicking those that occur pathologically in B or T cell lymphomas when chromosomal translocations initiate the malignancies. These results are pioneering representing the first experimental success in driving artificially in vitro the RAG pairing on two DNAs located *in trans*. The next step in our experiments is to see what influence transcription has on RSS configuration prior to synapsis in trans(see below).

Objective 2-D3B3 Test how DNA supercoiling affects RAG deleterious synapsis *in trans*.The positive results obtained at the experiments presented at sections D3D-1 and 2 allow us to investigate how local transcription from a heterologous T7 promoter influences the configuration of an adjacent reporter 12RSSr and its ability to synapse *in trans* with a reporter 23RSSr (Fig. 25A). For this experiment we have built a new construct shown in Fig. 25A(upper circle) which carries a 12RSSr reporter, under the influence of an upstream located T7 promoter, and a constitutive 23RSS flanked from both sides by two 8xOp. The other plasmid used in the experiment is the previously used 23RSS reporter flanked by two operator octamers 8xOp. The results of this experiment are shown in Fig. 25B with the





incubation scheme similarly displayed as in Fig.23 above the gel, whereas on the left hand side the cartoon diagram represents the legend of the band separation pattern. Notice here the lowest electrophoretic mobility band is a mixture of 12RSSr and 23RSSc unresolved singly cutted products (a+b), band c is the 23RSSr single cutted product whereas band d of highest mobility corresponds to the RAG *in cis* coupled cleavage product (12RSSr&23RSSc). All reactions containing T7 RNA polymerase and rNTP were also tested by RT-PCR to have supported the appropriate 180 bp transcripts with a specific set of primers which were used before in 2x12RSST7Del transcription reactions (the 12RSSr/23RSSc lends its entire T7 transcription block together with its 12RSSr from this original construct)(data not shown). First a direct

assessment of 12RSSr(near the T7 promoter) single cleavage cannot be made due to the band a+b resolving issue. Hence, we shall account for how 12RSSr reconfigured by T7 transcription influences afar its RAG pairing ability with 23RSSr(reflected in 23RSSr single cut product levels). Because RAG single cut at the reporter 23RSSr has significant background in the absence of partner 12RSSr (wells 9 to 12), to objectively quantify the band of this product under various reaction conditions these background bands were subtracted from their positive corresponding reactions(in presence of 12r/23c counterparts). The quantified band intensities of the 23RSSr single cut(SC) products were expressed as their relative ratios in Fig. 25 C inset with respect to +/- T7 polymerase transcription presence, first in the absence of LacI(left bar size 0.37) and then in presence of LacI(ratio1.06). In absence of LacI, 23RSSr and 12RSSr(where T7 transcription is running) are as expected uncoupled an effect reflected in the low 0.37 ratio background 23RSSrSC value. When LacI is present, the two reporter sites are being paired but the levels of 23RSSr SC production are unchanged by the T7 transcription occurrence at its partner pairing site 12RSSr(hence, the unitary ratio value between band c intensities under both conditions). The message is also visually compelling if we compare band c(23RSSr SC) intensity between well 6 (-T7) and that in well 8(+T7) in Fig.25B. **In conclusion this result suggests that in the context of LacI mediating *in trans* docking of a pair of reporter 12 and 23RSSs located on distinct DNAs, transcription of one site does not influence their RAG synapsing levels beyond the normal levels observed when the reaction is not assisted by transcription. Thus, it seems that T7 transcription requires a local effect to enhance RAG synapsis on transcribed RSSs located *in cis* as shown in the experiments described in Fig.16 and Fig.19.**

OBJECTIVE 4

D3D. Use DNA supercoiling to induce *in vivo* RAG translocations on a synthetic chromosomal mililocus. The most ambitious objective with which we proposed to conclude our project starts building in the nucleus of a murine fibroblast cell 3T3NIH an endogenous chromosoal synthetic minilocus containing a sequence of 21 polyoperator LacI repressor sites arranged in tandem(21xOp). The 21xOp in the minilocus must be flanked at either site by a 12 and a 23RSS sequence, in a flipped deletional orientation, each under the control of a T7 promoter which is oriented to run transcription in the direction 5'nonamer-heptamer 3'(Fig.26B). To fulfill this objective we have built a plasmid px21LacOpFRT construct containing the 21xOp with its 12/23RSSs with their two attached T7 promoters in an MLS site of the backbone commercial pcDNA5/FRT(Invitrogen-FlpIn System) vector flanked by its FRT(Flp recombinase transfer) sites(see Fig.26A). Using the Invitrogen system we started from a murine cell line Flp-In-3T3 which has in its nucleus an endogenous locus with a FRT casette adjacent to a promoter driving its Zeocin resistance gene. Such cells grow in our case after appropriate titration on media with 250μg/ml Zeocin. Any Flp recombination in this FRT casette switches off the Zeocin resistance gene and turns on a Hygromycin resistance gene from the inserted sequence(by using the appropriate SV40 FRT adjacent promoter and lending its ATG site, which is missing in pcDNA5/FRT).The rationale of the 21xOp minilocus insertion is to use Flp recombinase (provided from the pOG44 vector) in cotransfections with our px21LacOpFRT on Zeocin grown Flp-In-3T3 cells and 48 hours post transfections switch the 3T3 transfected cells on media supplemented with Hygromycin to select for our 21xOp positive clones integrants (see Fig.26A and B). After appropriate pOG44 & px21LacOpFRT cotransfections we selected our sets of Flp-In-3T3 transfected cells for two weeks in Dulbecco media supplemented with a wide range of

Hygromycin concentrations(10-400µg/ml Hygromycin B), with permanent media renewal at every three days.

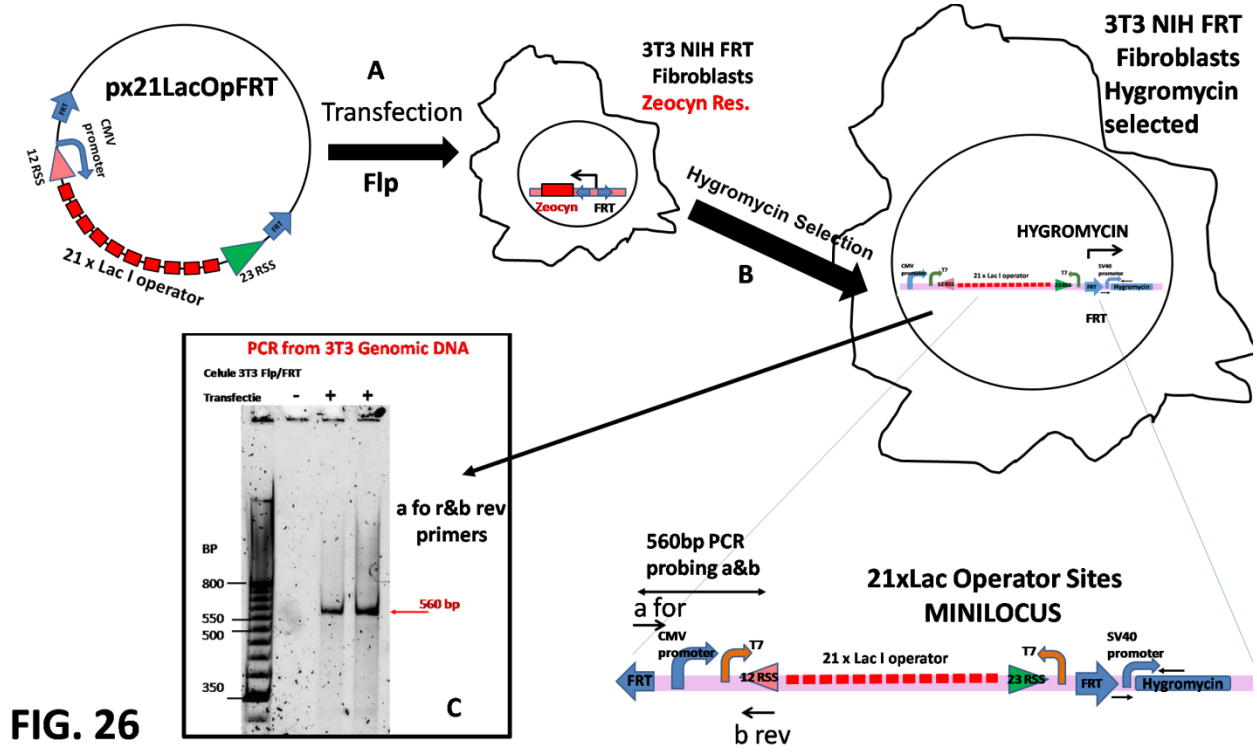


FIG. 26

After two weeks of growing in Hygromycin selecting media we succeeded selecting three individual cellular clones with resistance to this antibiotic. To test whether our selected clones contained the appropriate insertion 21xOp minilocus we extracted from each individually grown amplified Hygromycin resistant selected cell clones ($6x10^5$ cells each) genomic DNA and with it templated PCR reactions using a pair of specific primers (a for & b rev) which anneal specifically at the 5' end of the 21xOp minilocus (see Fig.26B and C). Figure 26C shows the result of two positive 3T3 21xOp minilocus clones which yield a positive 560bp amplification product versus the probed genomic DNA extracted from Flp-In-3T3 untransfected cells(first control well on the gel). Because the 21xpolyOp sequence has 21 repetitions of 20 bp(each operator site) the PCR method is not appropriate to amplify such repetitive DNA. To verify that the PCR tested clones have the minilocus in its complete configuration we next tested our

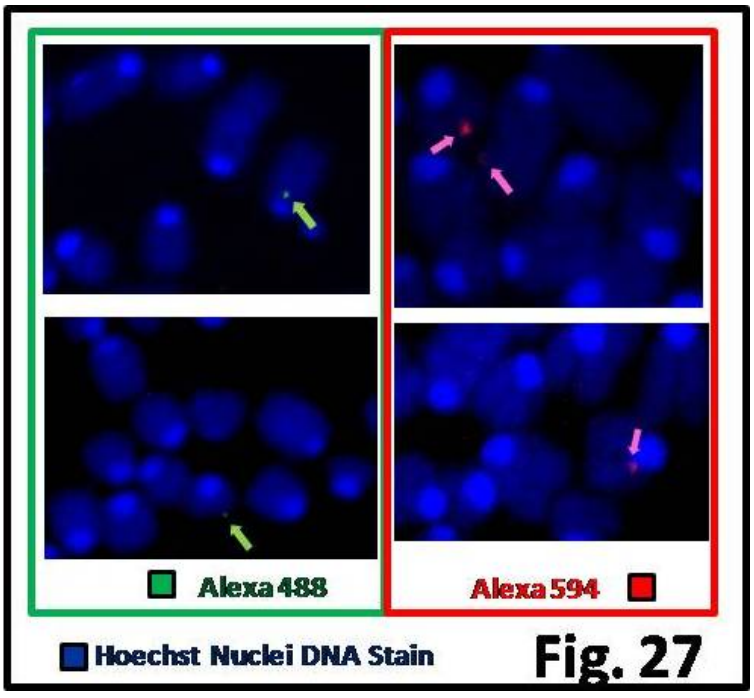


Fig. 27

selected 3T3_21xOp cells using fluorescence in situ hybridization (FISH) technique probing the cells with a 5.6kb very small probe which contains the whole sequence of px21OpLacFRT labeled by nick translation either with Alexa 488 or with Alexa 594 fluorophores. The FISH results shown in Fig. 27 on Hoechst DNA nuclear blue background show that we have positively identified with both types of labels specific signals (Fig.27 arrows indicate specific probe fluorescence) to our probes despite its short size(almost at the limit of detection). As described in our project proposal at pg 31 we have built

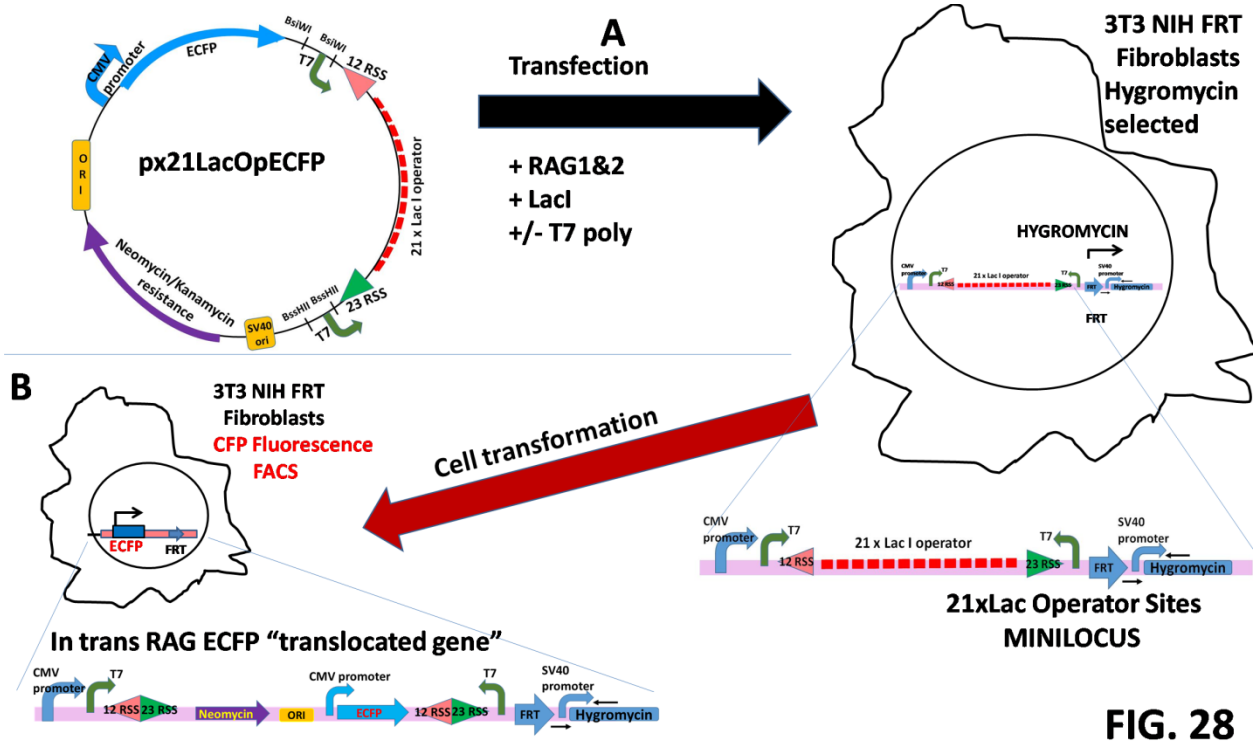


FIG. 28

another plasmid for mammalian cell transfection schematically depicted in the left corner of Fig.28A, namely px21LacOpECFP from a Novagen backbone vector pECFP in which the same 21xpolyOp flanked by 12/23RSS pair was inserted juxtaposed to a Cyan Fluorescent Protein gene transcribed from an independent CMV promoter. Another construct that we have built was a plasmid for mammalian cell transfection expressing LacI with a nuclear localization sequence NLS at its N terminus in a pEBB backbone vector. Fig. 28 schematically depicts the experiments that we performed (and we are currently still performing them) by cotransfecting the 3T3x21Op cells with px21LacOpECFP, pEBBRAG1&RAG2 and +/- pAR3132T7RNP_NLS. As shown in Fig. 28B bottom, the cells in which RAG recombine *in trans* the 12/23 RSSs of px21LacOpECFP with those of the endogenous 21xpolyOp locus this yields cells which gain intrinsic fluorescence by FACS. We are currently undergoing intense Neomycin selections for this recombination assay which should conclude our project with the final selection of CFP endogenously Fluorescent 3T3, cells which are the recipient of In trans RAG ECFP "translocated gene"

FINAL CONCLUSIONS

At the end of our project we succeeded to accomplish all our proposed goals fulfilling at each of the 4 objectives the following valuable fundamental findings and remarkable deliverables:

Objective 1- D3A. **Identify the mechanism by which DNA supercoiling activates RAG catalysis.**

a) Routinely express and purify RAG1, RAG2 and HMGB1/2 of various sources, active proteins that constitute valuable reagents for all our in vitro studies.

b) We have identified based upon IS264 RAG facilitated ligation assays, that RAG induces rapidly (<5min) substantial accumulation of supercoiled circles. Their production was initially postulated then proved by FRET experiments to be due to a mechanism by which RAG synapsis wraps plectonemically bent 12&23 RSS actively inducing DNA supercoiling. This effect facilitates dramatically in cis synapsis, which is evidenced by monomeric circles accumulating rapidly and efficiently, versus multimers which barely detectable in the assay.

c) We have investigated the mechanism by which RAG cleavage is activated by the degree of supercoiling of its substrate. We performed RAG coupled cleavage on 10 distinctly supercoiled topoisomers of various helical phase internal positionings of the 12/23RSSs. The cleavage is substantially increased in supercoiled topoisomers and respects the helical orientations of RSSs which synapse optimally. **This suggests that RSS supercoiling simply facilitates a lower activation energy for the formation of the synaptic complex, fostering its configuration.**

d) Using FRET methodology as an alternative to circumvent an unproductive design in our original proposal(IHF mediated DNA bending pg27) we have determined the 12&23RSS configuration in the RAG synaptic complex, and we further exploited these findings building a model of the two RSSs in the presence of the RAG complex. **The configuration of 23-RSS in the synaptic complex was published in 2013**

Ciubotaru, M. Trexler A. J., Spiridon L., Surleac M. D., Rhoades E, ., Petrescu A. J., Schatz D.G. "RAG and HMGB1 create a large bend in the 23RSS in the V(D)J recombination synaptic complexes. *Nucleic Acids Research,* 2013,vol **41**, 2437-2454

We have identified the configuration of 12-SC and that of 12RSS in the synaptic complex published in 2015

Ciubotaru M. , Surleac M. D., Metskas L. A., Koo P., Rhoades E., Petrescu A. J., Schatz D.G." The architecture of the 12RSS in V(D)J recombination signal and synaptic complexes ", Nucleic Acids Research, 2015, vol43, 917-931

Objective 2-D3B3 Test how DNA supercoiling affects RAG deleterious synapsis *in trans.*

e) We assembled 3 very important constructs in which an octamer of LacI poly operator sites is located in the vicinity of reporter 12 or 23RSS and with it we showed beyond any doubt the

remarkable effect LacI has in mediating the *in trans* RAG cleavage at the reporter 23RSS via docking of the two poly8x operators. Within a range of concentrations between 0.1-1μM of LacI the single cleavage at 23RSS r increases almost five folds, increase which occurs only when the appropriate in trans 12RSSr partner is present in the reaction(Fig.24 of the report).

f) Our results(Fig.25) suggest that in the context of LacI mediating *in trans* docking of a pair of reporter 12 and 23RSSs located on distinct DNAs, transcription of one site does not influence their RAG synapsing levels beyond the normal levels observed when the reaction is not assisted by transcription. Thus, it seems that T7 transcription requires a local effect to enhance RAG synapsis on transcribed RSSs located *in cis* as shown in the experiments described in Fig.16 and Fig.19. at Obj. 3

These results and those from Obj 3 were presented as an oral presentation: ***Ciubotaru M***,."V(D)J recombination targeted in cis by transcription induced DNA supercoiling", Invited talk presented at 14th edition of the Site-Specific Recombination, Transposition and DNA dynamics workshop, DNA transactions, Sept. 7-12, 2014 Isle d'Oleron, France.

Objective3 -D3C. Prove transcription driven DNA supercoiling to be the selective factor in favoring PC formation *in cis*.

g) We have shown both by in vivo and in vitro experiments that transcription driven supercoiling can activate dramatically RAG activity in the promoter adjacent RSSs. We have shown from our intracellular assays(Fig.15 and 16) that T7 polymerase transcribing not only that it does not sterically hinder in any way the access of RAG to its RSSs(located within 75bp proximity from the transcription initiation site) but enhances its catalysis to them. This notable effect is a direct facilitation of RAG activity induced by transcription, and must not be confused with the described accessibility phenomenon by which the local chromatin modifications are fostered by the endogenous RNA polymerase II action.

h) In our in vitro assays we show high RAG enhancement near a 12RSS reporter in deletional under the T7 transcription influence in underwound topoisomer UnwC2(Fig.19).This finding can be explained by our original hypothesis that the negative supercoiling created behind a RNA marching polymerase in the deletional orientation helps bending and reconfigures the 12RSS located behind the promoter, making it more amenable for synapsis with its 23 partner(Fig.19).

The findings described at d), e) and h) were presented first as a review paper on theoretical grounds in 2014

Ciubotaru M., Surleac M.D., Musat G. M., Rusu A. M., Ionita E., Albu C. C. Paul " DNA bending in the synaptic complex in V(D)J recombination: turning an ancestral transpososome upside down", Discoveries, 2(1):e13:1-15, 2359–7232(2014).

and later in an International conference :***Ciubotaru M.*** "V(D)J recombination targeted in cis by transcription induced DNA supercoiling", Invited talk presented at Albany 2015 19th Conversation, June 9th-13, 2015, University of Albany, NY, USA.

j)The data presented in Obj. 2 and Obj. 3 are currently incorporated in a manuscript :"**Transcription direct effects that influence V(D)J recombination "** Andreea Maria Rusu, Elena Ioniță, Cristina Jiglaru, Ioana Popa, Mihaela G. Musat, David G. Schatz , Ștefan E. Szedlacsek and Mihai Ciubotaru, which should be submitted for publication by the end of 2016.

Objective 4-D3D.Use DNA supercoiling to induce *in vivo* RAG translocations on a synthetic chromosomal mililocus.

i)We succeeded to assemble and make a 3T3 stable cell line 3T3 21xLacIOp carrying endogenously in a chromosome a 21x LacI Operator minilocus and we proved its presence and integrity by PCR and FISH technique(Fig.26 and 27).

k) Using the stable cell line 3T3 21xLacIOp we have developed an unprecedented intracellular in trans RAG assay with which we can test various endogenous DNA insertions in the chromosome of the host cell. This is a valuable reagent with which we intend to study the translocations involved in lymphoagenesis.

l) The data presented in Obj. 4 are currently incorporated in a manuscript :**"Intracellular Induced *in trans* V(D)J recombination "** Mihaela G. Musat, and Mihai Ciubotaru, which should be submitted for publication by the end of 2016.

References

1)Ciubotaru, M., et al. Fluorescence resonance energy transfer analysis of recombination signal sequence configuration in the RAG1/2 synaptic complex. Mol Cell Biol **27**, 4745-4758, (2007).

2) Yin, F. F. et al. Structure of the RAG1 nonamer binding domain with DNA reveals a dimer that mediates DNA synapsis. Nat Struct Mol Biol **16**, 499-508, (2009).

3) Ciubotaru, M. Trexler A. J., Spiridon L., Surleac M. D., Rhoades E, ., Petrescu A. J., Schatz D.G. "RAG and HMGB1 create a large bend in the 23RSS in the V(D)J recombination synaptic complexes. *Nucleic Acids Research,* 2013,vol **41**, 2437-2454

4) Swanson, P. C Fine structure and activity of discrete RAG-HMG complexes on V(D)J recombination signalsMol Cell Biol **22**, 1340-1351, (2002).

5) Zwilling, S. et al. High mobility group protein 2 functionally interacts with the POU domains of octamer transcription factors, EMBO J. 14, 1198-1208, (1995).

6) Dunn, J. J. et al. Targeting bacteriophage T7 RNA polymerase to the mammalian cell nucleus,Gene, vol68,259-266,(1988).

*7)*Ciubotaru M., Surleac M.D., Musat G. M., Rusu A. M., Ionita E., Albu C. C. Paul " DNA bending in the synaptic complex in V(D)J recombination: turning an ancestral transpososome upside down", Discoveries, 2(1):e13:1-15, 2359–7232(2014).

8) Ciubotaru M. **,** Surleac M. D., Metskas L. A., Koo P., Rhoades E., Petrescu A. J., Schatz D.G." The architecture of the 12RSS in V(D)J recombination signal and synaptic complexes ", *Nucleic Acids Research,* 2015, vol43, 917-931

9) Ciubotaru M."V(D)J recombination targeted in cis by transcription induced DNA supercoiling", *Journal of Biomolecular Structure and Dynamics,* 2015, vol 33, Sup1, p71. DOI: 10.1080/07391102.2015.1032746.

10) Kim M. S., Lapkouski M., Yang W., Gellert M. "Crystal structure of the V(D)J recombinse RAG1-RAG2", *Nature,* 2015, vol 518, p507.

11) Ru H., Chambers M. G., Fu T.M., Tong A.B., Liao M., Wu H. "Molecular mechanism of V(D)J recombination from synaptic RAG1-RAG2 complex structures", *Cell,* 2015, vol 163, p1138-1152.