

# Homologous recombination in mycobacteria

K. Muniyappa\*, N. Ganesh, N. Guhan, Pawan Singh, G. P. Manjunath, S. Datta<sup>†</sup>, Nagasuma R. Chandra<sup>‡</sup> and M. Vijayan<sup>†</sup>

Department of Biochemistry, <sup>†</sup>Molecular Biophysics Unit, and <sup>‡</sup>Bioinformatics Center, Indian Institute of Science, Bangalore 560 012, India

**In recent years, considerable effort and resources have been expended to develop targeted gene delivery methods, and generation of auxotrophic mutants of mycobacteria. The results of these studies suggest that mycobacteria exhibit a wide range of recombination rates, which vary from loci to loci. Here we review the methods developed for allele exchange and targeted gene disruption as well as the mechanistic aspects of homologous recombination in mycobacteria. The results of whole genome, functional and structural analyses of *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* RecA and SSB proteins provide insights into variations of the prototypic *Escherichia coli* paradigm. This variation of a common theme might allow mycobacteria to function in their natural but complex physiological environments.**

STUDIES of *Mycobacterium tuberculosis* are hindered by its long generation time (12 h), severe clumping of the bacilli and the safety risk involved with handling live cultures. Our understanding of the mechanisms of pathogenesis caused by the tubercle bacillus is inadequate, and the factors responsible for virulence are poorly understood. Although much research has focused on immunology, biochemistry, and microbiology of this pathogen, investigations into molecular interactions between specific gene products have not been possible because of the lack of defined mutants with specific phenotypes. It is believed that transfer of DNA into tubercle bacilli, either by allele replacement or transposon mutagenesis would provide insights into understanding of the role(s) of virulence determinants as well as mechanisms of pathogenesis. Therefore, understanding of the mechanistic aspects of homologous recombination may help molecular genetic manipulation of mycobacteria as well as knowledge needed to develop strategies to control tuberculosis.

## Introduction of foreign DNA into mycobacteria

Introduction of foreign DNA by transduction or conjugation has greatly facilitated the generation of mutant strains and the functional analysis of the genomes of *Escherichia coli* and *Salmonella typhimurium*<sup>1,2</sup>. Similarly, introduction of foreign DNA into mycobacterial strains via a genetic route has relied on the processes of transformation or transduction. Various plasmids, derived from

mycobacteriophages, such as TM4, L1, and D29, have proven useful for the development of transformation systems for mycobacteria<sup>3</sup>. The transfer of DNA by transduction by a virus was first demonstrated for *M. smegmatis*<sup>4,5</sup>. More recently, a single-step and relatively efficient allele exchange method was developed using a shuttle plasmid integrated into a specialized transducing mycobacteriophage TM4. This method was used to construct seven isogenic auxotrophic mutant strains of *M. smegmatis*, three substrains of *M. bovis* BCG and three strains of *M. tuberculosis*<sup>6</sup>. A number of investigators have ascertained the potential utility of this method for targeted gene disruptions at several loci in *M. tuberculosis*<sup>7-10</sup>. Mycobacteriophages have been used as vectors to generate luciferase reporter phages for the rapid detection of pathogenic species of mycobacteria and the assessment of their drug susceptibilities.

Bacterial conjugation is a process by which DNA is transferred from a donor to recipient cell through cell-to-cell contact mediated by energy-driven transport. The process is conceptualized as two sub-processes: DNA preparation, and mating bridge formation. Studies of conjugation in *E. coli* have played a crucial role in the development of bacterial genetics, and led to the isolation of the first recombination-deficient (*rec*) mutants. In *E. coli* the functions required for conjugation are mainly encoded by the F factor, which act at a unique *cis*-acting site to initiate and complete DNA transfer. By contrast, in the naturally occurring conjugation system of *M. smegmatis*, DNA transfer is chromosomally encoded<sup>11</sup>. In addition, unlike conventional plasmid transfer, recipient recombination functions are required to allow this plasmid, and derivatives of it, to re-circularize through a process similar to gap repair. Extended DNA homology with the recipient chromosome and the F factor is required to facilitate repair, resulting in acquisition of recipient chromosomal DNA by the plasmid. Together, these results show that DNA transfer in *M. smegmatis* occurs by a mechanism different from that of prototypical plasmid transfer systems<sup>11</sup>. Plasmid-mediated conjugative gene transfer has not been demonstrated in strains belonging to *M. tuberculosis* complex.

## Gene transfer in mycobacteria

In recent years, considerable effort and resources have been expended to the development of methods for targeted

\*For correspondence. (e-mail: kmbc@biochem.iisc.ernet.in)

gene delivery and mutation in mycobacteria. The methods are mechanistically similar to those developed for *E. coli* or yeast. Such approaches indicate that generation of genetically defined isogenic strains containing single or multiple mutations has been hampered by the lack of suitable tools. *M. tuberculosis* and *M. smegmatis* genomes exhibit a wide range of recombination rates as reflected in the efficiency of allele exchange, which is known to vary from loci to loci. It has been technically difficult to generate defined auxotrophic mutants of *M. tuberculosis* at high frequency in a routine manner. Mutagenesis of mycobacteria has been performed by random or targeted strategies. In organisms in which gene targeting has been observed at high efficiency, DNA molecules with broken ends have been shown to be more recombinogenic than covalently closed circular DNA. However, the stimulatory role of double-strand breaks in mycobacteria is poorly understood. Historically, the first recombinant DNA vectors developed for mycobacteria include shuttle plasmid vectors and chimeric DNA molecules that replicate in *E. coli* as plasmids and in mycobacteria as phages<sup>12</sup>. These are integrated into the bacterial genome, by recombination, so that encoded resistance genes may be maintained over time if the plasmid cannot replicate independently within that cell. The early studies of successful isolation of auxotrophic mutants for the *M. tuberculosis* complex strains used insertional mutagenesis systems, which resulted in illegitimate recombination<sup>13</sup>, transposon mutagenesis<sup>14</sup> or allele exchange<sup>15</sup>. Over the years, a variety of alternative gene transfer strategies have been developed to achieve high frequency of allele exchange in *M. smegmatis*<sup>16-20</sup>. On the other hand, similar studies in *M. tuberculosis* involving random shuttle mutagenesis using transposons displayed low frequency of mutations at allelic sites<sup>13,21,22</sup>. The difficulties encountered in these studies led to the conclusion that slow-growing species of mycobacteria promote a high frequency of illegitimate recombination<sup>13,23,24</sup>. Why is this the case? The probable answer stems from the fact that the methods used for detection of very rare allelic exchange events are hindered by low transformation efficiencies and high frequencies of illegitimate recombination, especially in the slow-growing species of pathogenic mycobacteria.

Advances in the construction of gene targeting vectors together with the improvement in the delivery systems have led to increased efficiency of generation of 'knockout' mutants of *M. tuberculosis* and *M. bovis* BCG. These investigations involved short<sup>25,26</sup> or long linear DNA fragments<sup>15</sup> as substrates. Several groups have demonstrated the use of 'suicide' plasmid vectors (using a non-temperature-sensitive plasmid) to achieve insertional mutagenesis in both fast- and slow-growing species of mycobacteria<sup>18-20,21,27-33</sup>. A two-step selection method using selectable and counterselectable markers, positioned on either replicating or non-replicating plasmids, has also been successfully used in *M. smegmatis*<sup>20,32</sup>, *M. bovis*

BCG and *M. tuberculosis*<sup>30,34-36</sup>. Interestingly, in the case of 'suicide' plasmid vectors, pretreatment of DNA with UV light or alkali enhanced homologous recombination (HR), and abolished illegitimate recombination in the recipient cells. The suicide vector approach is dependent upon the delivery of the gene targeting vectors by electroporation. Therefore, the HR frequencies are very close to the efficiency at which plasmids can be electroporated into slow-growing mycobacteria, whereas the suicide vector approach is limited to those cases where high transformation efficiencies can be obtained. Consequently, it is surmised that this electroporation limitation, rather inefficient HR, may be the reason for difficulties encountered in allele exchange experiments in slow-growing species of mycobacteria<sup>19</sup>.

An alternative strategy for gene targeting involves the use of replicating vectors. The method offers the advantage that high density of recombinant vectors increase the frequency of allele exchange, compared with that obtained with the suicide vectors. These vectors have greatly improved reproducibility of allele exchange in the slow-growing species of mycobacteria<sup>37</sup>. The possible reasons are (i) the availability of increased time for HR and (ii) DNA replication and recombination occur concurrently in the cell. A further increase in the efficiency of isolation of allelic replacements has been achieved by combining a counter-selection method with vectors bearing temperature-sensitive origin of replication<sup>38</sup>. Recently, a promising method has been developed for making targeted gene knockouts in *M. smegmatis* and *M. bovis* BCGs based on two-plasmid incompatibility system. This method uses a pair of replicating plasmids carrying a mutated allele of a targeted gene or a transposon, and has the advantage by providing prolonged time for HR<sup>37</sup>. When used for the generation of *M. smegmatis* *pyrF* mutant alleles, high frequency of recombinants was obtained by this method.

### **Analysis of *M. tuberculosis* and *M. leprae* genomes for *rec* genes**

The *M. tuberculosis* genome is 4.4 Mb long, which is exceedingly rich in genes for lipid biosynthesis and degradation<sup>38</sup>. In parallel, the 3.3 Mb genome sequence of *M. leprae* has been determined<sup>39</sup>. *M. tuberculosis* genome can potentially encode 3924 genes, while the *M. leprae* encodes 1604 proteins and contains 1116 pseudogenes, compared to six in *M. tuberculosis*<sup>38,39</sup>. Comparison of the genome sequence of *M. leprae* with that of *M. tuberculosis* indicates that the former has undergone massive gene decay, losing large number of genes since its divergence from a common mycobacterial ancestor<sup>40,41</sup>. It is possible that its genes were rendered inactive once their functions were no longer essential for survival, and this was followed by genome shrinkage through rear-

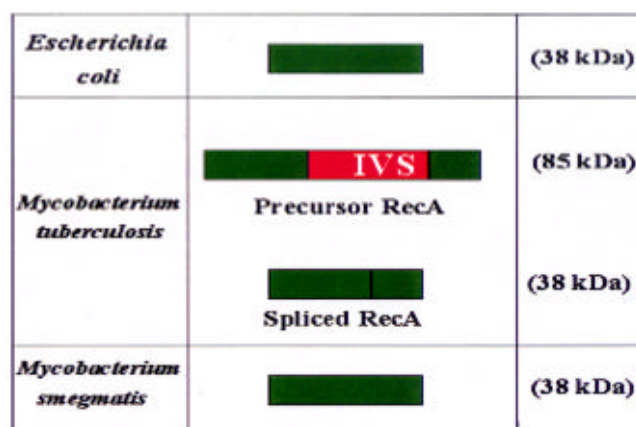
rangements and/or deletions. It has been proposed that downsizing of *M. leprae* genome, and mutations in several metabolic genes, may account for its exceptionally slow growth as well as its failure to grow *in vitro*. It seems to have completely dispensed with or substantively reduced certain metabolic pathways, including oxidative and anaerobic respiratory chains. The enzymes for breaking down host-derived lipids, a means by which many mycobacterial pathogens derive their energy, are also drastically reduced in *M. leprae*. By contrast, most anabolic pathways seem to be intact, indicating that *M. leprae* depends on these pathways to survive in the nutrient-poor micro-environment of phagosomes<sup>38-41</sup>. The availability of the mycobacterial genome sequences and the ability to generate transposon mutants, targeted gene disruptions, and complementation analyses provide an unprecedented opportunity for the elucidation of the functions of mycobacterial genes.

Comparative analysis of the genomes of *M. tuberculosis* and *M. leprae* has revealed a considerable decay or deletion of genes involved in recombination, especially of those encoding for alternate pathways of HR<sup>40-42</sup>. In *E. coli*, at least four alternate pathways exist for HR, each featuring the action of a distinct exonuclease and/or helicase<sup>43-44</sup>. These are required to generate 3' invasive ends for polymerization of RecA to initiate recombination. These include RecBCD, RecE/RecT or RecJ/RecQ. Most notably, the *M. tuberculosis* genome is devoid of homologues of *E. coli* *recE*, *recT*, *recQ*, *recJ*, *recO* and *rusA*<sup>42</sup>. RecQ helicase has been shown to disrupt illegitimate recombination in *E. coli*, and its absence could be one of the reasons for higher frequency of illegitimate recombination in *M. tuberculosis*. Intriguingly, *M. tuberculosis* *recB*, *recC* and *recD* genes resemble those of Gram-negative species rather than analogues of *addA* *addB* that exist in Gram-positive bacteria<sup>45</sup>. *M. leprae* genome possess neither of these systems, however, it contains an archaeal-type exonuclease and helicase similar to the *recB* family of exonuclease/helicase. Mutations are also found in *M. leprae* genes involved in DNA repair (the *mutT*, *dnaQ*, *alkA*, *dinX*, and *dinP* genes)<sup>45</sup>.

In *E. coli*, early steps of HR involve RecA, RecBCD enzyme, and the recombination hotspot called Chi (*c*) site. *E. coli* *c* sites are G-rich (5'-GCTGGTGG-3') asymmetric *cis*-acting regulatory sequences that modify the activities of the RecBCD enzyme, thereby leading to the generation of single-stranded DNA. This results in preferential loading of RecA onto the *c*-containing DNA strand. The RecA nucleoprotein filament then invades homologous double-stranded DNA to produce a D-loop structure. Although *recB*, *recC*, *recD* genes and putative *c*-like sites have been identified in the *M. tuberculosis* genome<sup>42,46</sup>, and are likely to exist in other mycobacteria, it has not been shown whether they constitute recombination hotspots in any of the mycobacterial species.

## Organization and expression of mycobacterial *recA*

The biochemical activities of many of the factors involved in HR in mycobacteria are poorly understood. However, two components of the pathway of HR in mycobacteria, RecA and SSB, have been studied in considerable detail. One of the primary functions of eubacterial *recA* is its role in recombinational DNA repair<sup>43,44</sup>. Recombination between similar DNA sequences contributes significantly to genome plasticity; post-replicative mismatch repair, and restricts recombination between homologous sequences. RecA is both ubiquitous and well conserved among a range of organisms. Unlike *M. smegmatis*, pathogenic species of mycobacteria display relatively low levels of HR<sup>47</sup>. In contrast to *M. smegmatis* *recA*, the *M. tuberculosis* and *M. leprae* *recA* contain an in-frame open reading frame encoding an intein<sup>48-49</sup>. RecA intein is removed from the precursor RecA by an autocatalytic protein splicing reaction, and active RecA is generated by ligation of amino- and carboxyl-terminal fragments mediated by intein (Figure 1). This post-translational processing is required for RecA activity: a mutant gene that no longer undergoes protein splicing fails to complement *E. coli* *recA* mutants, whereas the wild-type gene can<sup>49</sup>. Therefore, it is possible that this unusual arrangement for the production of mature, active RecA protein might affect its function in *M. tuberculosis*, either by regulating the splicing reaction or by subsequent interaction of the intein with RecA. The biological significance of the presence of intein in the *recA* gene in pathogenic mycobacteria is the subject of the on-going debate<sup>47,50-52</sup>. Given the fact that inteins are mostly found in *recA* of pathogenic mycobacteria, the advantage is unclear. Although the significance of *recA* intervening sequence is unknown, it has been shown that *M. leprae* or *M. tuberculosis* *recA* complements *M. smegmatis*  $\Delta$ *recA* strains for recombination and UV repair<sup>50-52</sup>.



**Figure 1.** Structural organization of *E. coli*, *M. tuberculosis* and *M. smegmatis* RecA proteins. IVS, intervening sequence.

The characterization of *M. tuberculosis* RecA intein revealed that it is a unique member of the LAGLIDADG family of homing endonucleases<sup>53</sup>. *M. tuberculosis* RecA intein displayed very novel characteristics: In the presence of Mn<sup>2+</sup> and ATP, it was able to cleave cognate site in the inteinless *recA* allele at 24 and 33/43 bases upstream of the intein insertion site, in the upper and lower strands respectively<sup>54</sup>. This property is consistent with the class of homing endonucleases that tolerate some sequence degeneracy within their recognition sequences<sup>53</sup>. Recent studies also provided a great deal of insight into the catalysis of DNA cleavage as well as ATPase activity of RecA intein<sup>55</sup>. Intriguingly, RecA intein displayed robust site-specific endonuclease activity with non-cognate DNA in the presence of Mg<sup>2+</sup> generating DNA fragments with blunt end or 1–2 base overhangs<sup>56</sup>. The latter activity has been implicated in the movement of RecA intein DNA sequence from one chromosome location to another in natural populations<sup>56</sup>. It is unknown whether *M. leprae* RecA intein possess similar activities.

In *E. coli*, *recA* is a part of the SOS response system<sup>43</sup>. SOS response is activated by agents or processes related to DNA metabolism that generate single-stranded DNA. The SOS box is the target for LexA binding, which exists upstream of all genes expressed in the SOS regulon. Under normal growth conditions, low levels of RecA and LexA exist in the cells<sup>43</sup>. LexA functions as a repressor to inhibit *recA*, *lexA* and many other repair operons involved in the SOS response. When RecA is activated by DNA damage, it promotes autocatalytic cleavage of LexA. In *M. tuberculosis*, in addition to *recA*, a number of DNA-damage induced genes are regulated by LexA-dependent mechanism<sup>57,58</sup>. However, few genes induced in response to DNA damage are not regulated by LexA binding, but by an alternate mechanism of gene regulation<sup>57</sup>. Recently, it has been shown that *M. tuberculosis recA* is expressed from two promoters: one is LexA-regulated, and the second remains DNA damage inducible in the absence of RecA or when LexA binding is prevented<sup>59</sup>. These findings indicate that the mycobacterial DNA repair system is different in many aspects compared to the prototypic model species, e.g. *E. coli* or *B. subtilis*. *recA* deletion mutant of *M. smegmatis* strain (HS42) exhibited enhanced sensitivity to UV irradiation and failed to display HR<sup>34</sup>. The deficiencies in UV survival and recombination were complemented by introduction of the cloned *M. smegmatis recA* gene<sup>34</sup>.

### Recombination activities of mycobacterial RecA and SSB proteins

Recombination is central to the identification of genes and to the understanding of the biology of any organism. Using *E. coli* as a model, the process of HR has been separated into four kinetically distinguishable phases:

presynapsis, synapsis, strand exchange and resolution<sup>43,44</sup>. *Presynapsis* involves cooperative binding of RecA protein on single-stranded DNA forming a helical nucleoprotein filament; *synapsis*, the homologous alignment of nucleoprotein filament comprised of RecA–ssDNA with naked duplex DNA; and unidirectional *strand exchange*, which creates long stretches of heteroduplex DNA. Finally, the heteroduplex DNA is expanded by RuvAB motor proteins and resolved by the RuvC endonuclease. *E. coli* RecA is the central component in these processes, and, because its functions are conserved from bacteriophage to humans, its study has provided a paradigm for understanding the biologically important process of HR. This complex process requires the action of > 20 gene products. In *E. coli*, the proteins that carry out all of the steps of HR have been purified and characterized *in vitro*. These studies are quite advanced in the case of *E. coli*, and portions of the HR pathway are being reconstituted *in vitro*. RecA-like proteins constitute a group of DNA strand transfer proteins ubiquitous in eubacteria, eukarya and archaea. However, the functional relationship among RecA-like proteins is poorly understood.

To understand the basis for inefficient allele exchange in mycobacteria compared to *E. coli*, to obtain gene targeting in mycobacteria with reasonable efficiency, and to understand the differences in allele exchange between *M. tuberculosis* and *M. smegmatis*, what is needed is greater insight into the molecular mechanism of HR in mycobacteria, and detailed characterization of the biochemical activities of the components of HR in these species. It is possible that the endogenous DNA repair and recombination machinery in mycobacteria is different from that of *E. coli*. To this end, *M. tuberculosis* RecA (38 kDa), but not its precursor (85 kDa), displayed the hallmark features of *E. coli* RecA, including binding to single-stranded DNA, ssDNA-dependent ATP hydrolysis, formation of D-loops, homologous pairing between single-stranded DNA with duplex DNA, and strand exchange<sup>60,61</sup>. There were, however, striking qualitative and quantitative differences in the activities and pattern of strand exchange promoted by *E. coli* and *M. tuberculosis* RecA on one hand, and between *M. smegmatis* and *M. tuberculosis* RecA on the other<sup>60–62</sup>. These include rates of ssDNA-dependent ATP hydrolysis, conditions and cofactors required for the display of maximum homologous pairing and strand exchange. Mycobacterial RecA proteins promoted much slower rates of ATP hydrolysis than the rates of the reactions catalyzed by *E. coli* RecA in side-by-side comparisons. Results of molecular modelling and the crystal structure of *M. tuberculosis* RecA indicated that the reduced affinity for ATP and relative catalytic inefficiency of *M. tuberculosis* RecA is related to the expansion of the P-loop region, compared to its homologue in *E. coli*<sup>60</sup>.

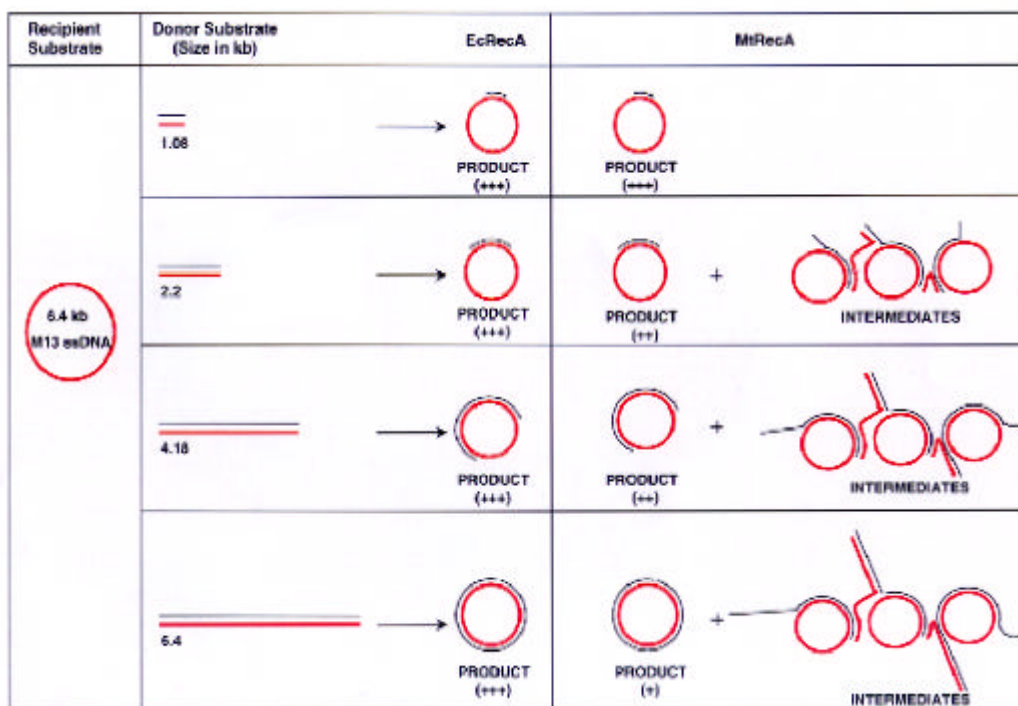
Another set of observations revealed significant differences in the pattern of homologous pairing and strand

exchange promoted by *M. tuberculosis* and *M. smegmatis* RecA proteins. *M. tuberculosis* RecA was able to effect maximum strand exchange in the alkaline pH range, whereas *M. smegmatis* RecA was around neutral pH<sup>61-63</sup>. Although the rates and the pH profiles of dATP hydrolysis catalysed by *M. tuberculosis* and *M. smegmatis* RecA were similar, only the latter was able to couple dATP hydrolysis to strand exchange. A number of studies have shown that single-stranded DNA binding proteins (SSB) serve as accessory factors in HR<sup>43,44,64</sup>. *M. smegmatis* SSB (165 aa) shares 84% identity and 89% similarity with the *M. tuberculosis* SSB (164 aa)<sup>65</sup>. While *E. coli* RecA promoted substantial strand exchange in the absence of SSB, mycobacterial RecA proteins were completely unable to do so in the absence of SSB<sup>61,63</sup>. This finding confirmed the absolute requirement of SSB for a HR in *M. smegmatis*. Significantly, unlike *E. coli* SSB, mycobacterial SSB proteins physically interacted with their cognate RecA proteins with high affinity. Further, DNA size played an important role on the ability of mycobacterial RecA proteins to synthesize extended lengths of heteroduplex DNA. For example, with duplex DNA length of < 2 kb, the efficiency of strand exchange was indistinguishable from that of the prototype *E. coli* RecA, whereas it decreased with increase in the size of duplex DNA (Figure 2)<sup>61,63</sup>. *E. coli* RecA was able to effect complete strand exchange between linear duplex DNA

(6.4 kb) and ssDNA (6407 nucleotides) to generate gapped or nicked circular duplex DNA. In contrast to this, *M. tuberculosis* RecA generated substantial amounts of intermediates and networks of DNA as the length of linear duplex DNA was increased from 1 kb to 6.4 kb. The direct correlation between the length of duplex DNA and accumulation of DNA intermediates and networks of DNA indicate that the ability of *M. tuberculosis* RecA to generate extended stretches of heteroduplex DNA is limited. In addition, strand exchange promoted by *M. tuberculosis* and *M. smegmatis* RecA displayed distinctly different pH profiles, suggesting functional diversity between RecA from pathogenic and non-pathogenic species of mycobacteria<sup>62,63</sup>.

### Structure of mycobacterial RecA proteins

In the absence of DNA, the crystal structure of *E. coli* RecA revealed a central core domain and two smaller domains at the amino and carboxyl termini. The core which is made up of twisted eight stranded **b**-sheet flanked by four **a**-helices contains domains for DNA-binding, designated as L1 and L2 loops, and P-loop containing the nucleotide triphosphate-fold<sup>66,67</sup>. In the *E. coli* RecA crystal structure, the monomers are packed so as to form a right-handed helical filament with 6 monomers



**Figure 2.** Model for the effect of the length of duplex DNA on strand exchange promoted by *E. coli* or *M. tuberculosis* RecA. Reactions were performed with linear duplex DNA (donor; 1–6.407 kb) and nucleoprotein filaments of RecA-circular single-stranded DNA (recipient; 6407 nucleotide residues) as described<sup>61</sup>. ‘Plus’ symbols correspond to the extent of strand exchange: ‘+++’ denotes maximum strand exchange; ‘++’ half-maximal; and ‘+’ one-third of the maximum value. EcRecA, *E. coli* RecA; MtRecA, *M. tuberculosis* RecA.

per turn<sup>66,67</sup>. The overall structures of *M. tuberculosis* and *M. smegmatis* RecA are nearly identical to the *E. coli* RecA structure with r.m.s. deviations of 0.6 to 1.2 Å<sup>68,69</sup>. In comparison, the molecular surface of the *M. tuberculosis* and *M. smegmatis* RecA filaments possess negative electrostatic potential, compared to *E. coli* RecA (Figure 3). The ligand-bound structures of *M. tuberculosis* RecA revealed subtle variations in nucleotide conformations. Furthermore, the neighbouring filaments in the bundle are involved in several hydrogen bonds in *E. coli* RecA, whereas are hardly any in the case of *M. tuberculosis* RecA. As a consequence, the association of filaments of mycobacterial RecA into bundles is much weaker. The DNA-binding loops, L1 and L2, were undefined in the *E. coli* RecA crystal structure. On the other hand, the conformation and orientation of L1 and L2 loops in the mycobacterial RecA structures were defined, and appear to be different<sup>70</sup>. More importantly, the nucleotide binding by the *M. smegmatis* RecA was accompanied by the movement of Gln196 in the L2 loop, which has been implicated in the propagation of the signal induced by the binding of nucleotide cofactor to the DNA-binding loops<sup>69</sup>.

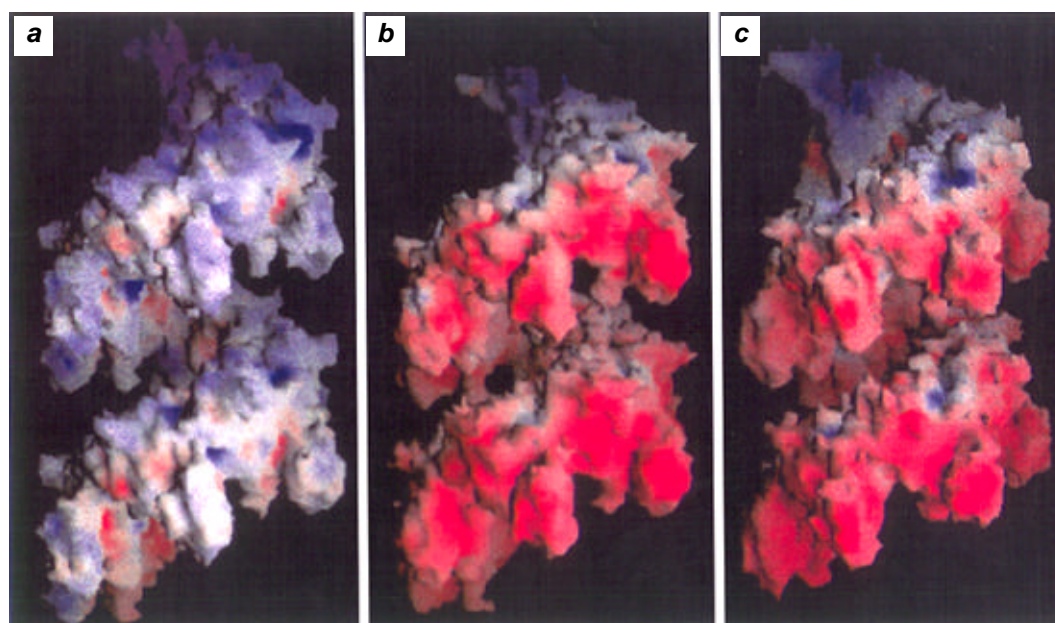
### Regulation of recombination

Recombination is vital for various cellular processes related to DNA metabolism, but it must be tightly controlled. The insight into regulation of HR has been derived from studies on *recX* in eubacteria. In *M. smegmatis*, *M. tuberculosis*, *Pseudomonas aeruginosa*, *Streptomyces lividans*, or *Thiobacillus ferrooxidans*, the ORFs of *recA* and *recX*

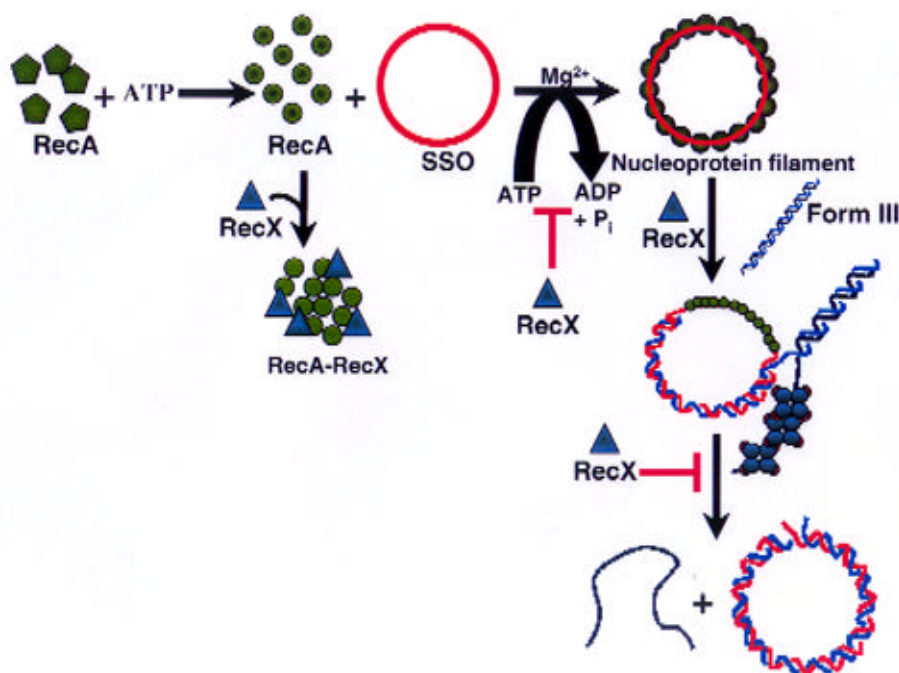
overlap and the two genes are co-transcribed<sup>71–76</sup>. It is known that overexpression of *recA* in *recX* mutants of *S. lividans*, *M. smegmatis*, or *P. aeruginosa*, but not mutant RecA, lead to induction of deleterious effects<sup>72–73</sup>. However, the molecular mechanisms by which *recX* attenuates the deleterious effects induced by *recA* overexpression have remained unknown. Using *M. tuberculosis* as a model, it has been shown that *M. tuberculosis* RecX binds directly to *M. tuberculosis* RecA as well as *M. smegmatis* and *E. coli* RecA proteins *in vivo* and *in vitro*, but not SSB<sup>77</sup>. The direct association of RecX with RecA failed to regulate the specificity or extent of binding of RecA either to DNA or ATP, ligands that are central to activation of its functions. Significantly, RecX severely impeded ATP hydrolysis and the generation of heteroduplex DNA promoted by homologous as well as heterologous RecA proteins<sup>77</sup>. These findings reveal a novel mode of negative regulation of RecA, and imply that RecX might act as an anti-recombinase to repress inappropriate recombinational repair events during normal DNA metabolism (Figure 4). Consistent with these observations, *E. coli* RecX was shown to inhibit strand exchange as well as ATPase activities of its cognate RecA<sup>78</sup>, indicating that negative regulation of HR might be a general phenomenon.

### Perspectives

Understanding of the biology of tubercle bacillus requires inputs from comparative analysis of non-pathogenic species of mycobacteria as well. From the earliest studies of HR, it was recognized that *E. coli* is the best model for



**Figure 3.** Molecular surface representation of the RecA filament: *a*, *E. coli*; *b*, *M. tuberculosis*; *c*, *M. smegmatis*. The surfaces are colour-coded according to the electrostatic potential: Red (negative values), blue (positive values), and white (neutral values).



**Figure 4.** Schematic representation of three-strand exchange promoted by the concerted action of RecA and SSB and inhibition by RecX. Both in *in vivo* and *in vitro* conditions, RecX (triangle) has been shown to interact with RecA (circle)<sup>77</sup>. The displaced single-stranded DNA is sequestered by tetramers of SSB. The site of action of RecX is indicated by the T symbol. Form III, linear duplex DNA.

elucidation of the mechanism of HR at the molecular level. The results of whole genome, functional and structural analyses of *M. tuberculosis* and *M. smegmatis* RecA and SSB proteins provide insights into variations of the prototypic *E. coli* paradigm. This variation of a common theme might allow mycobacteria to function in their natural but complex physiological environments. However, further functional and structural studies will be required in understanding the activities of the recombination machinery as well as the mechanistic aspects of HR in mycobacteria.

- Zinder, N. D. and Lederberg, J., *J. Bacteriol.*, 1952, **64**, 679–699.
- Lennox, E. S., *Virology*, 1955, **1**, 190–206.
- Snapper, S. B., Lugosi, L., Jekkel, A., Melton, R. E., Kieser, T., Bloom, B. R. and Jacobs, W. R., Jr., *Proc. Natl. Acad. Sci. USA*, 1988, **85**, 6987–6991.
- Raj, C. V. and Ramakrishnan, T., *Nature*, 1970, **228**, 280–281.
- Tokunaga, T., Mizuguchi, Y. and Suga, K., *J. Bacteriol.*, 1973, **113**, 1104–1111.
- Bardarov, S. *et al.*, *Microbiology*, 2002, **148**, 3007–3017.
- Glickman, M. S., Cox, J. S. and Jacobs, W. R. Jr., *Mol. Cell*, 2000, **5**, 717–727.
- Raman, S., Song, T., Puyang, X., Bardarov, S., Jacobs, W. R., Jr. and Husson, R. N., *J. Bacteriol.*, 2001, **183**, 6199–6125.
- Sirakova, T. D., Thirumala, A. K., Dubey, V. S., Sprecher, H., Kolattukudy, P. E., *J. Biol. Chem.*, 2001, **276**, 16833–16839.
- Steyn, A. J., Collins, D. M., Hondalus, M. K., Jacobs, W. R. Jr., Kawakami, R. P. and Bloom, B. R., *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 3147–3152.

- Parsons, L. M., Jankowski, C. and Derbyshire, K. M., *Mol. Microbiol.*, 1998, **28**, 571–582.
- Jacobs, Jr., W. R., Tuckman, M. and Bloom, B. R., *Nature*, 1987, **327**, 532–534.
- Kalpna, G. V., Bloom, B. R. and Jacobs, W. R. Jr., *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 5433–5437.
- McAdam, R. A. *et al.*, *Infect. Immun.*, 1995, **63**, 1004–1012.
- Balasubramanian, V. *et al.*, *J. Bacteriol.*, 1996, **178**, 273–279.
- Boshoff, H. I. and Mizrahi, V., *J. Bacteriol.*, 2000, **182**, 5479–5485.
- Braunstein, M., Brown, A. M., Kurtz, S. and Jacobs, W. R., Jr., *J. Bacteriol.*, 2001, **183**, 6979–6990.
- Pavelka, M. S. Jr. and Jacobs, W. R., *J. Bacteriol.*, 1999, **181**, 4780–4789.
- Frischkorn, K., Sander, P., Scholz, M., Teschner, K., Pramnanan, T. and Bottger, E. C., *Mol. Microbiol.*, 1998, **29**, 1203–1214.
- Knipfer, N., Seth, A. and Sharder, T. E., *Plasmid*, 1997, **37**, 129–140.
- Sander, P., Meier, A. and Bottger, E. C., *Mol. Microbiol.*, 1995, **16**, 991–1000.
- Gordhan, B. G., Andersen, S. J., De Meyer, A. R. and Mizrahi, V., *Gene*, 1996, **178**, 125–130.
- Aldovini, A., Husson, R. N. and Young, R. A., *J. Bacteriol.*, 1993, **175**, 7282–7289.
- McFadden, J., *Mol. Microbiol.*, 1996, **21**, 205–211.
- Azad, A. K., Sirakova, T. D., Rogers, L. M. and Kolattukudy, P. E., *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 4787–4792.
- Reyrat, J.-M., Berthet, F.-X. and Gicquel, B., *Proc. Natl. Acad. Sci. USA*, 1995, **92**, 8768–8772.
- Husson, R. N., James, B. E. and Young, R. A., *J. Bacteriol.*, 1990, **172**, 519–524.
- Berthet, F. X. *et al.*, *Science*, 1998, **282**, 759–762.
- Fitzmaurice, A. M. and Kolattukudy, P. E., *J. Biol. Chem.*, 1998, **273**, 8033–8039.

30. Parish, T. and Stoker, N. G., *Microbiology*, 2000, **146**, 1969–1975.
31. Pelicic, V., Jackson, M., Reyrat, J. M., Jacobs, W. R., Gicquel, B. and Guilhot, C., *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 10955–10960.
32. Pelicic, V., Reyrat, J. M. and Gicquel, B., *J. Bacteriol.*, 1996, **178**, 1197–1199.
33. Pelicic, V., Reyrat, J. M. and Gicquel, B., *Mol. Microbiol.*, 1996, **20**, 919–925.
34. Papavinasasundaram, K. G., Colston, M. J. and Davis, E. O., *Mol. Microbiol.*, 1998, **30**, 525–534.
35. Hinds, J., Mahenthalingam, E., Kempell, K. E., Duncan, K., Stokes, R. W., Parish, T. and Stoker, N. G., *Microbiology*, 1999, **145**, 519–527.
36. Parish, T., Gordhan, B. G., McAdam, R. A., Duncan, K., Mizrahi, V. and Stoker, N. G., *Microbiology*, 1999, **145**, 3497–3503.
37. Pashley, C. A., Parish, T., McAdam, R. A., Duncan, K. and Stoker, N. G., *Appl. Environ. Microbiol.*, 2003, **69**, 517–523.
38. Cole, S. T. *et al.*, *Nature*, 1998, **393**, 537–544.
39. Cole, S. T. *et al.*, *Nature*, 2001, **409**, 1007–1011.
40. Young, D. and Robertson, B., *Curr. Biol.*, 2001, **11**, R381–383.
41. Brosch, R., Gordon, S. V., Eiglmeier, K., Garnier, T. and Cole, S. T., *Res. Microbiol.*, 2000, **151**, 135–142.
42. Muniyappa, K., Vaze, M. B., Ganesh, N., Reddy, M. S., Guhan, N. and Venkatesh, R., *Microbiology*, 2000, **146**, 2093–2095.
43. Kowalczykowski, S. C., Dixon, D. A., Eggleston, A. K., Lauder, S. D. and Rehrauer, W. M., *Microbiol. Rev.*, 1994, **58**, 401–465.
44. Cox, M. M., *Annu. Rev. Genet.*, 2001, **35**, 53–82.
45. Dawes, S. S. and Mizrahi, V., *Lepr. Rev.*, 2001, **72**, 408–414.
46. Kyuma, D., Arakawa, K., Uno, R., Nakayama, Y. and Tomita, M., *Genome Informatics*, 2002, **13**, 533–534.
47. McFadden, J., *Mol. Microbiol.*, 1996, **21**, 205–211.
48. Davis, E. O., Sedgwick, S. G. and Colston, M. J., *J. Bacteriol.*, 1991, **173**, 5653–5662.
49. Davis, E. O., Jenner, P. J., Brooks, P. C., Colston, M. J. and Sedgwick, S. G., *Cell*, 1992, **71**, 201–210.
50. Davis, E. O., Thangaraj, H. S., Brooks, P. C. and Colston, M. J., *EMBO J.*, 1994, **13**, 699–703.
51. Frischkorn, K., Sander, P., Scholz, M., Teschner, K., Prammannan, T. and Böttger, E. C., *Mol. Microbiol.*, 1998, **29**, 1203–1214.
52. Frischkorn, K., Springer, B., Böttger, E. C., Davis, E. O., Colston, M. J. and Sander, P., *J. Bacteriol.*, 2000, **182**, 3590–3592.
53. Guhan, N. and Muniyappa, K., *Crit. Rev. Biochem. Mol. Biol.*, 2003, **38**, 199–248.
54. Guhan, N. and Muniyappa, K., *J. Biol. Chem.*, 2002, **277**, 16257–16264.
55. Guhan, N. and Muniyappa, K., *Nucleic Acids Res.*, 2003, **31**, 4184–4191.
56. Guhan, N. and Muniyappa, K., *J. Biol. Chem.*, 2002, **277**, 40352–40361.
57. Brooks, P. C., Movahedzadeh, F. and Davis, E. O., *J. Bacteriol.*, 2001, **183**, 4459–4467.
58. Durbach, S. I., Anderson, S. J. and Mizrahi, V., *Mol. Microbiol.*, 1997, **26**, 643–653.
59. Davis, E. O., Springer, B., Gopaul, K. K., Papavinasasundaram, K. G., Sander, P. and Böttger, E. C., *Mol. Microbiol.*, 2002, **46**, 791–800.
60. Kumar, R. A., Vaze, M. B., Chandra, N. R., Vijayan, M. and Muniyappa, K., *Biochemistry*, 1996, **35**, 1793–1802.
61. Vaze, M. B. and Muniyappa, K., *Biochemistry*, 1999, **38**, 3175–3186.
62. Ganesh, N. and Muniyappa, K., *Proteins Struct. Funct. Genet.*, 2003, **53**, 6–17.
63. Ganesh, N. and Muniyappa, K., *Biochemistry*, 2003, **42**, 7216–7225.
64. Reddy, M. S., Vaze, M. B., Madhusudan, K. and Muniyappa, K., *Biochemistry*, 2000, **39**, 14250–14262.
65. Reddy, M. S., Guhan, N. and Muniyappa, K., *J. Biol. Chem.*, 2001, **276**, 45959–45968.
66. Story, R. M., Weber, I. T. and Steitz, T. A., *Nature*, 1992, **355**, 318–325.
67. Story, R. M. and Steitz, T. A., *Nature*, 1992, **355**, 374–376.
68. Datta, S., Prabhu, M. M., Vaze, M. B., Ganesh, N., Chandra, N. R., Muniyappa, K. and Vijayan, M., *Nucleic Acids Res.*, 2000, **28**, 4964–4973.
69. Datta, S., Krishna, R., Ganesh, N., Chandra, N. R., Muniyappa, K. and Vijayan, M., *J. Bacteriol.*, 2003, **185**, 4280–4284.
70. Datta, S., Ganesh, N., Chandra, N. R., Muniyappa, K. and Vijayan, M., *Proteins: Struct. Funct. Genet.*, 2003, **50**, 474–485.
71. De Mot, R., Schoofs, G. and Vanderleyden, J., *Nucleic Acids Res.*, 1994, **22**, 1313–1314.
72. Vierling, S., Weber, T., Wohlleben, W. and Muth, G., *J. Bacteriol.*, 2000, **182**, 4005–4011.
73. Papavinasasundaram, K. G., Movahedzadeh, F., Keer, J. T., Stoker, N. G., Colston, M. J. and Davis, E. O., *Mol. Microbiol.*, 1997, **24**, 141–153.
74. Sano, Y., *J. Bacteriol.*, 1993, **175**, 2451–2454.
75. Horn, J. M. and Ohman, D. E., *J. Bacteriol.*, 1988, **170**, 1637–1650.
76. Guilani, N., Bengrine, A., Borne, F., Chippaus, M. and Bonnefoy, V., *Microbiology*, 1997, **143**, 2179–2187.
77. Venkatesh, R., Ganesh, N., Guhan, N., Reddy, M. S., Chandrasekhar, T. and Muniyappa, K., *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 12091–12096.
78. Stohl, E. A., Brockman, J. P., Burkel, K. L., Moromatsu, K., Kowalczykowski, S. C. and Seifert, S. H., *J. Biol. Chem.*, 2003, **278**, 2278–2285.

ACKNOWLEDGEMENTS. This work was supported by a grant from the Indian Council of Medical Research, New Delhi.