

The Need for Speed: Run-On Oligomer Filament Formation Provides Maximum Speed with Maximum Sequestration of Activity

Funding Source: Research reported in this publication was supported by the National Science Foundation under Grant No. MCB-1410355 (to NCH), as well as the Office of the Director, National Institutes of Health of the National Institutes of Health under award number S10OD013237 (to CKP), and by the National Institute of General Medical Sciences of the National Institutes of Health under award number T32GM008659 (to JLS). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS, NIH or NSF.

Claudia Barahona, L. Emilia Basantes, Kassidy J. Tompkins, Desirae M. Heitman, Barbara I. Chukwu,
Juan Sanchez, Jonathan L. Sanchez, Niloofar Ghadirian, Chad K. Park & N. C. Horton*

Department of Molecular and Cellular Biology, University of Arizona, Tucson, Arizona 85721

Running title: Advantages of the run-on oligomer filament mechanism

¹*To whom correspondence should be addressed:

Dr. Nancy C. Horton

Department of Molecular and Cellular Biology

Telephone: (520) 626-3828.

E-mail: nhorton@u.arizona.edu.

Abbreviations

bp, base pair

BMe, 2-mercaptoethanol

BPB, bromophenol blue

DTT, dithiothreitol

EDTA, ethylenediaminetetraacetic acid

HEPES, 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid

HEPES-NaOH, HEPES titrated to a desired pH with NaOH

kDa, kilodalton

mw, molecular weight

nt, nucleotide or nucleotides

OAc, acetate

PAGE, polyacrylamide gel electrophoresis

RE, restriction endonuclease

RM system, restriction-modification system encoding cognate endonuclease and methyltransferase

RT, room temperature

SDS, sodium dodecyl sulfate

Tris, Tris(hydroxymethyl)aminomethane

Tris-HCl, Tris titrated to a desired pH with HCl

XCFF, xylene cyanol FF

ABSTRACT:

Herein we investigate an unusual anti-viral mechanism developed in the bacterium *Streptomyces griseus*. SgrAI is a type II restriction endonuclease which forms run-on oligomer filaments when activated, and which possesses both accelerated DNA cleavage activity and expanded DNA sequence specificity. Mutations disrupting the run-on oligomer filament eliminate the robust anti-phage activity of wild type SgrAI, and the observation that even relatively modest disruptions completely abolish this anti-viral activity shows that the greater speed imparted by the run-on oligomer filament mechanism is critical to its biological function. Simulations of DNA cleavage by SgrAI uncover the origins of the kinetic advantage of this newly described mechanism of enzyme regulation over more conventional mechanisms, as well as the origin of the sequestering effect responsible for the protection of the host genome against the damaging DNA cleavage activity of activated SgrAI.

Key words: run-on oligomer, protein filament, restriction endonuclease, DNA binding protein, enzyme mechanism, phage, anti-phage mechanisms

IMPORTANCE

This work is motivated by the interest in understanding the characteristics and advantages of a relatively newly discovered enzyme mechanism involving filament formation. SgrAI is an enzyme responsible for protecting against viral infections in its host bacterium, and was one of the first such enzymes shown to utilize such a mechanism. In this work, filament formation by SgrAI is disrupted and the effects on the speed of the purified enzyme as well as its function in cells are measured. It was found that even small disruptions, which weaken but do not destroy filament formation, eliminate the ability of SgrAI to protect cells from viral infection, its normal biological function. Simulations of enzyme activity were also performed and show how filament formation can greatly speed up an enzyme's activation compared to other known mechanisms, as well as better localize its action to molecules of interest such as invading phage DNA.

INTRODUCTION

The co-evolution of phage and anti-phage activities in what's been called the “phage-host arms race” is thought to be among the oldest and largest in scale co-evolutionary system on Earth^{12, 3}. From this system, many useful biomacromolecules have been discovered. For example, the type II restriction endonucleases (REs) have shown great utility in recombinant DNA technology due to their very high sequence specificity and rapid double stranded DNA cleavage abilities⁴. Newer technologies include the CRISPR enzymes, which allow relatively more convenient programming of DNA cleavage specificity⁵. Still the great diversity of REs contrasts with the relatively limited number that have been fully studied, suggesting that much is left to be discovered in this interesting class of enzymes⁶. Our studies with the type II RE SgrAI from *Streptomyces griseus* led us to propose a new mechanism of enzyme regulation involving filament formation^{7, 8}.

Filament and run-on oligomer formation by non-cytoskeletal enzymes is a relatively newly discovered phenomenon, being first described in 2009-2010 for such diverse enzymes as Ire1 (the unfolded protein response nuclease-kinase)⁹, CTP synthase^{10, 11}, ACC (acetyl-coA carboxylase)¹², and SgrAI⁷. At approximately the same time, large-scale screens for protein localization using fluorescence microscopy showed unexpectedly that many enzymes formed filaments in response to particular metabolic conditions or other stimuli in cells^{11, 13-15}. The term “run-on oligomer” (ROO) filament is used here to describe an assembly of an enzyme into a filament by the successive addition of enzymes at either end, and which in principle could extend indefinitely^{8, 16}. ROO filament formation by SgrAI was first proposed in 2010 based on behavior in analytical ultracentrifugation and native gels⁷, and subsequently using ion-mobility mass spectrometry¹⁷. The enzymatic activity of SgrAI was found to be activated in the ROO and to possess an altered (expanded) DNA sequence specificity^{7, 8}. The three-dimensional CryoEM structure of the ROO filament formed by the assembly of SgrAI/DNA complexes shows a left handed helical arrangement with approximately four DNA bound dimers of SgrAI per turn (**Fig. 1A, Fig. 1B-D** show different views of an individual SgrAI/DNA complex)⁸. In the ROO filament helix, both the DNA and SgrAI form stabilizing interactions with neighboring SgrAI/DNA complexes (**Fig. 1A, E-F**)⁸.

SgrAI is a type II RE cleaving primary sites (CR|CCGGYG, R indicating A or G, Y indicating C or T, |

indicating cleavage site) and secondary sites (CR|CCGGGG and CR|CCGGY(A or C or T)) in duplex DNA in a magnesium ion dependent mechanism^{18, 19}. Its cleavage of secondary sites occurs only under particular conditions, namely when present on the same DNA molecule as a primary site, or alternatively, when in the presence of high concentrations of both SgrAI and primary site DNA sequences on separate DNA molecules^{7, 20-23}. The same conditions leading to cleavage of secondary sites by SgrAI also accelerate the cleavage of primary sites by SgrAI over 200-fold (the acceleration of secondary site cleavage is approximately 1000-fold)^{7, 16, 20, 22}. Further, under these conditions, SgrAI forms the ROO filament described above, thought to stabilize the activated state of SgrAI⁸. The role of this unexpected structure was not known and has been the subject of recent investigations.

Being a relatively newly described enzyme mechanism, several fundamental questions are of interest including: 1) how the ROO filament accelerates the formation of the product of the reaction (*i.e.* cleaved DNA) without trapping it in the filament, 2) whether or not the assembly and/or the disassembly of the ROO limits the overall rate of reaction, 3) what the growth and dissolution mechanism of the ROO filament is (*e.g.* from the ends only or occurring anywhere in the filament?), and 4) are there special advantages of the ROO filament mechanism (over more conventional mechanisms) that evolved due to the particular biological niche of SgrAI. The first three questions were addressed in prior work^{24, 25}, which showed that ROO filament assembly is rate limiting in *in vitro* reactions at low concentrations of SgrAI and DNA. It was also found that DNA cleavage is rapid in the ROO filament, faster than dissociation of the ROO filament, making the reaction pathway efficient since DNA cleavage is much more likely with each addition to the ROO filament prior to its dissociation. But the relatively rapid dissociation kinetics limits any trapping of cleaved DNA within the filament. As for the growth mechanism, the previous work concluded that disassembly of ROO filaments can occur at any junction between adjacent SgrAI/DNA complexes, and assembly must also be possible from two ROO filaments of any size.

In the current work, we address the fourth question of whether there are particular advantages to the ROO filament mechanism, perhaps relating to requirements and challenges of specialized biological niches. First, we show that mutations that disrupt the ROO filament also eliminate the ability of SgrAI to protect a host

bacterium from invading phage. We find that even relatively moderate disruptions of ROO filament formation appear to render the protection to nearly that of the parent strain, as if the enzyme were not even present in the cell. This indicates that the speed of SgrAI, in terms of rapid activation and DNA cleavage, is critical to the biological function of SgrAI.

Next, using kinetic modeling and rate constants derived and determined from prior work^{24, 25}, we simulate the *in vivo* kinetics of SgrAI activity. We also build an alternative, non-ROO model to use side-by-side in simulations in order to discover any advantages or limitations inherent to the ROO filament mechanism. In order to simulate the reaction *in vivo*, we estimate concentrations of SgrAI and DNA in the cell, and also estimate the “local” concentration of SgrAI when bound to sites on the same contiguous DNA molecule. Using these concentrations, and the kinetic model and rate constants derived in prior work^{24, 25}, we discover that while the relatively slow association rate constant of SgrAI/DNA complexes into the ROO filament does in fact limit the rate of reaction, it importantly is also the source of the proposed “sequestration effect” that protects the host genome from the potentially damaging activity of SgrAI. This is because it limits ROO filament formation within the cell to only those SgrAI bound to the same DNA molecule, and meaning *in vivo* SgrAI would cleave only invading phage DNA and not the host genome. SgrAI, a type II RE, acts as a bacterial defense system to protect its host bacteria from such invading and pathogenic phage. However, the activity of SgrAI must be controlled such that it does not perform damaging DNA cleavages on the host DNA.

Comparing simulations of the ROO filament reaction to a non-ROO (Binary association) reaction mechanism, we find that the ROO filament mechanism has a significant kinetic advantage over the non-ROO mechanism in the speed of DNA cleavage. Further, the advantage appears to derive from the two distinct ways SgrAI/DNA complexes and ROO filaments can associate (such as at either end of the ROO filament). In addition, investigation of the cleavage of secondary sites by SgrAI, which differ from primary by 1 base pair, shows that even an accelerated Binary reaction (with an increased assembly rate constant) is inferior to the ROO filament mechanism, due to greater host damage resulting from the less effective sequestration found in the non-ROO mechanism.

The “phage-host arms race” is complex with diverse mechanisms to evade infection on the part of the host, and evade restriction on the part of the phage³. For example, negative selective pressure should result in the reduction of the number of restriction sites on phage, and genome analyses suggest this is the case^{3, 26-29}. However most bacteria contain multiple RM (for restriction-modification) systems, and a positive correlation exists between genome size and the number of such systems³. The genome size of *S. griseus* predicts the presence of 4-5 RM systems, each possibly containing a unique recognition sequence³. The evasion of multiple RM systems simultaneously through mutations to eliminate recognition sequences, without affecting genome functions in coding, replication, and expression, would be challenging. The evolutionary pressure exerted on the SgrAI restriction-modification system, which has produced the unusual filament forming enzyme mechanism, may originate with the relatively large genome of its host, *Streptomyces griseus*. The larger genome results in many more potential recognition sites which must be methylated by the SgrAI methyltransferase for protection, lest be cleaved by the SgrAI endonuclease. The surprisingly slow DNA cleavage rate of SgrAI (0.1 min⁻¹), as well as the longer and therefore more rarely occurring 8 bp recognition sequence may both have evolved to reduce this pressure. However, these same properties, the slow cleavage rate and rare cleavage site, would also limit the effectiveness of SgrAI against phage infection. The 200-1000 fold activation of DNA cleavage activity, and expansion from 3 DNA recognition sites to 17 upon ROO filament formation results in many more possible cleavages in the phage genome, which may limit phage infection to a greater extent. For example, phage genomes with more restriction endonuclease cleavage sites are restricted to a greater degree than those with fewer³⁰⁻³². The ROO filament stabilizes the activated conformation of SgrAI, yet such stabilization could also in principle be performed by a simple binary mechanism involving the assembly of only two copies of SgrAI. The specific advantage of forming a filament, compared to a finite oligomer such as in a binary system, is indicated by this work to be in both the combined (and competing) properties of speed (faster activation) and sequestration (limiting secondary site cleavage to the invading phage DNA and not the host), resulting from the particular biological niche found in *Streptomyces griseus*.

RESULTS

DNA Cleavage

Single turnover DNA cleavage assays were performed to measure the basal rate of DNA cleavage by SgrAI (*i.e.* that in the absence of activating conditions) as well as under activating conditions (*i.e.* where ROO filaments normally form, which have accelerated DNA cleavage properties). The basal, unactivated rate of DNA cleavage was measured with an 18 bp DNA containing a single primary recognition site (18-1, see Methods for sequence), which does not activate SgrAI but can be cleaved, albeit at a slow rate ($0.094 \pm 0.015 \text{ min}^{-1}$ in the case of wild type SgrAI). This DNA cleavage rate constant should not be perturbed by the mutations, if the mutations affect only formation of the ROO filament. **Table 1** shows this basal DNA cleavage rate constant for wild type and each mutant of SgrAI. The values range from $0.014 \pm 0.003 \text{ min}^{-1}$ (R24E) to $0.14 \pm 0.03 \text{ min}^{-1}$ (I59E). Most are within one standard deviation of the wild type value, and all show measurable activity. Hence the mutations did not disrupt DNA binding by SgrAI, or the unactivated DNA cleavage activity, as expected.

The single turnover DNA cleavage rate constant under activating conditions was also measured for each mutant SgrAI. These assays include $1 \mu\text{M}$ PC DNA, which activates SgrAI into forming ROO filaments with accelerated DNA cleavage activity^{7, 8, 16}. PC DNA is a “pre-cleaved” 40 bp DNA containing a single primary site sequence (see Methods for sequence). Two copies of PC DNA self-anneal to form a 40 bp DNA with nicks (missing the phosphate due the absence of 5' phosphates on synthetic DNA) at the SgrAI cleavage sites (CR|CCGGYG, | indicates cleavage site). This 40 bp DNA binds to SgrAI the same as an uncleaved version would, and favors ROO filament formation by forming stabilizing interactions with other SgrAI/DNA complexes (see **Fig. 1**), including those containing the 18-1 DNA. In this way, accelerated cleavage of 18-1 is induced. Mutations which disrupt the ROO filament by removing favorable, or introducing unfavorable, interactions between SgrAI/DNA complexes will result in less activation, and therefore a slower DNA cleavage rate under normal activating conditions (*i.e.* with $1 \mu\text{M}$ PC DNA). **Table 1** shows the results; most mutations severely affect this rate constant (wild type is $22 \pm 7 \text{ min}^{-1}$, most mutations reduce this rate constant to $\sim 1 \text{ min}^{-1}$ or less, **Table 1**). Those with intermediate effects on this rate constant include S56Q, M62E, R84E and R127A (rate constants of 3-8 min^{-1}). Only the mutation, A57Q, gives a wild type rate constant

(within 1 standard deviation of that of wild type SgrAI, $15 \pm 8 \text{ min}^{-1}$), hence appearing to have minimal effects on disrupting activation (and ROO filament formation) of SgrAI.

Residues S56, A57, and R127 are all at a protein-DNA interface occurring between neighboring SgrAI/DNA complexes (**Fig. 1**). As described previously for S56 and A57 mutations¹⁶, the introduction of a negative charge at the protein interface (*i.e.* mutations S56E and A57E) creates electrostatic repulsion with the DNA, but mutation to a neutral side chain (*i.e.* mutations S56Q and A57Q) does not, hence the smaller effects on activated DNA cleavage by those mutations. In the case of R127A, removal of a positive charge at this interface also weakens the ROO filament, but perhaps not as effectively as those nearby R131A and R134A (both with rate constants measured under accelerating conditions of less than 1 min^{-1} , **Table 1**). As can be seen in the CryoEM model of the ROO filament (**Fig. 1E**), R127 is further from the protein-DNA interface than R131 and R134, providing an explanation for the smaller effect when mutated. M62 and R84 are at a different interface, one formed between the protein chains of two adjacent SgrAI/DNA complexes in the run-on oligomer filament (**Fig. 1E-F**). R84 and M62 are both more distant from this interface than the other residues mutated in this study. Residues T4, S6, I7, R24, N25, P27, Q34, I51, I59 are also at the protein-protein interface between adjacent SgrAI in the ROO filament (**Fig. 1E-F**), and mutation of these residues results in severe effects on the activation of SgrAI (**Table 1**).

Phage Challenge

The phage challenge assay uses cells (*i.e.* Tuner (DE3) *E. coli*, Novagen, Inc.) expressing wild type or mutant SgrAI, MspI.M (the MspI methyltransferase with CCGG specificity), and a mutant of the *E. coli* lambda phage, λ JL801, which is incapable of lysogeny³³. The SgrAI protein is expressed from an inducible T7 promoter on a high copy number plasmid, while the MspI.M methyltransferase is expressed from its natural promoter showing near consensus -35 and -10 sequences^{34, 35} in a plasmid derived from the low copy number plasmid pACYC184. The phage challenge assay involves mixing λ JL801 phage at different dilutions with cells expressing both MspI.M and wild type or mutant SgrAI, and counting the number of plaques after

overnight growth. Each plaque represents a successful infection by λ JL801 phage, and is quantified as the number of plaque forming units (*i.e.* PFU) per μ l of our stock of purified phage. The parent bacterial strain (having only MspI.M plasmid) gave $5.7 \pm 1.1 \times 10^4$ PFU/ μ l. When cells express wild type SgrAI, no plaques were found, even using the highest concentration of phage available. In contrast, all mutant SgrAI but one led to the formation of plaques, most near the count found with the parent strain (**Table 1**). The mutation A57Q, which also had the least effect (if any) on the accelerated (activated) DNA cleavage rate (**Table 1**), also appeared to completely protect SgrAI from phage infection, showing no plaques even at the highest concentration of phage available (**Table 1**). **Figure 2** plots the Plaque Titer (PFU/ μ l) vs. the activated DNA cleavage rate constant (min^{-1}) measured for that mutant (or wild type) SgrAI. As can be seen, the greatest protection (*i.e.* NP, no plaques observed, at least 10^4 -fold protection relative to the parent strain) is found with the fastest enzymes (wild type and A57Q). Some protection may occur with enzymes showing activation levels at 10-35% of that of wild type (R84E, S56Q, R127A and M62E), although only slight, if any, protection is found (**Fig. 2**).

Verification of protein expression

Cells used in the phage challenge assay were analyzed by Western blot to confirm expression of SgrAI proteins (**Fig. 3**). Levels of expression were assessed by determining the relative concentration of protein in each lane and normalizing to that found for wild type SgrAI (see Methods). The values were corrected for dilution and the estimate of cells used in each lane (**Table S1**). Values varied from 29%-108% (relative to expression of wild type SgrAI) showing that all mutant SgrAI proteins were expressed. Though the expression level varied, it is uncorrelated with the protection against phage measured by the phage challenge assay (**Fig. 4**).

Simulation of in vivo reactions: Local concentration and sequestration

One of the important distinctions between the *in vitro* and *in vivo* reactions is the concentration of DNA. *In vitro*, the concentration of DNA is experimentally controlled, and the average size of the ROO filament

and the activation of DNA cleavage by SgrAI are thereby controlled as well. *In vivo*, DNA concentrations will be limited to one copy of the bacterial genome and one or perhaps more copies of the invading phage DNA. In terms of concentration, this is approximately 3 nM, for one copy of DNA per *S. griseus* cell (see **Table S2**). The estimation of the concentration of SgrAI in the cell is less certain, since the number of copies per cell is unknown, but with 100 copies the concentration would be approximately 300 nM (**Table S2**). At this concentration sufficient binding should occur between SgrAI and its recognition site in DNA ($K_D=0.057\pm0.009$ nM)⁷, however, at 3 nM of SgrAI/DNA complex virtually no ROO assembly (when sites are on separate DNA molecules) is expected and therefore no activation of SgrAI mediated DNA cleavage would occur⁷. The reason for this, as well as the dependence of the observed DNA cleavage rate constant on the concentration of SgrAI bound to DNA⁷, is the very slow association rate constant for of SgrAI/DNA complexes into the ROO filament (*i.e.* 1.3×10^5 M⁻¹s⁻¹)^{24, 25}, giving a calculated rate of association of only 1.3×10^5 M⁻¹s⁻¹)(3×10^{-9} M)(3×10^{-9} M) = 1×10^{-12} M/s or 1×10^{-3} nM/s.

However, this slow association rate constant, though preventing reactions between SgrAI bound to sites on separate DNA molecules, is sufficiently fast to allow association when on the same DNA molecule, owing to the “local concentration” effect. When two species, such as SgrAI, are bound to the same molecule, like contiguous DNA, they are constrained in space relative to each other and can act as if their concentrations were much higher. This higher concentration is termed the local concentration and is calculated considering the average distance between the two species, in this case, the two DNA bound SgrAI. **Table S2** provides the calculations, which include estimation of the average number of SgrAI recognition sites within a typical phage genome, as well as their distance apart in base pairs. A radius of gyration was used to estimate the average distance in Å between them, then used as the diameter of a sphere for a volume to be calculated³⁶. Given the number of DNA bound SgrAI (two in this example) and the volume of the sphere, a concentration can be calculated (80 nM). This means that although their actual concentration is 3 nM each (each bound to one molecule of DNA in the cell, *i.e.* one phage genome), they occupy the same space two separate complexes would occupy if at 80 nM. Note, this concentration would be even greater when considering secondary site sequences, which occur much more frequently (with 14 different sequences) and hence closer

together. At 80 nM, much faster association is expected, $(1.3 \times 10^5 \text{ M}^{-1}\text{s}^{-1})(80 \times 10^{-9} \text{ M})(80 \times 10^{-9} \text{ M}) = 8 \times 10^{-10} \text{ M/s}$ or 0.8 nM/s, 800 times faster than association of complexes on separate DNA molecules (*i.e.* phage and host genomes). Both the relatively slow association rate constant ($1.3 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$), and the requirement for an association step between SgrAI/DNA complexes before DNA cleavage, are responsible for the control of SgrAI activation by the local concentration effect, which results in sequestration. Sequestration of activated SgrAI on the DNA that contains the primary sites is important, since although primary sites on the host DNA are protected from SgrAI by methylation, secondary sites could in principle be cleaved by activated SgrAI causing damage to the host genome. Hence, this elegant mechanism may have evolved to ensure sequestration of activated SgrAI (which cleaves secondary as well as primary recognition sites) on only the invading phage DNA (see below).

Simulation of in vivo reactions of SgrA with phage DNA

Because cleavage of sites on the same DNA molecule is different in some respects from the experimental system used in our prior kinetic investigations²⁵, a new kinetic model was created to simulate the cleavage of SgrAI recognition sites on phage (and/or host) DNA (**Table S3**). This model differs in using one type of DNA site rather than two, which can both be cleaved and can also result in activation of other SgrAI/DNA complexes when bound by SgrAI. Previous kinetic models^{24, 25} use two types of DNA, the reporter DNA (*i.e.* 18-1) which can be cleaved, and an activator DNA (*i.e.* PC DNA), as also used in the single turnover DNA cleavage reactions described above. In the *in vivo* case, and considering only primary sites, only one type of site will be found, and that site can both be cleaved by SgrAI as well as activate it by stimulating ROO filament formation. Hence the model for activity of SgrAI on a phage DNA contains only one type of DNA recognition site, which is capable of both activating SgrAI (by inducing filamentation), and can do so whether or not the bound DNA is cleaved¹⁶.

For comparison to a non-ROO filament mechanism, equations for what we refer to as a “Binary” mechanism are included to allow side-by-side comparisons between the two mechanisms. The same rate constants (derived from prior work²⁵) were used for analogous steps in both kinetic mechanisms (**Table S4**).

Finally, to simulate the *in vivo* case, the DNA concentrations used were 80 nM to mimic the local concentration of sites when present on the same contiguous DNA, and 3 nM when simulating reactions occurring on separate DNA molecules. Due to software limitations, the ROO filament model was limited to a size of 3 SgrAI/DNA complexes, however this is not unrealistic given the predicted number of primary sites on the typical phage DNA (see **Table S2**). Dissociation of cleaved DNA from SgrAI is considered reversible in this simulation and is discussed further below.

Figure 5A shows the results when the ROO filament mechanism (red and blue) is compared to the closed ended non-ROO Binary mechanism (green and pink). At 80 nM SgrAI/DNA (E/DNA for the Binary reaction, E being the hypothetical enzyme in the Binary mechanism), both ROO filament and Binary mechanisms show robust DNA cleavage. However, the ROO filament mechanism has a very clear advantage (approximately two-fold) in the rate of cleaving DNA (compare the red dotted lines to the green dotted lines of the Binary model, **Fig. 5A**). Some cleaved DNA is released (~25%, solid red and green lines, from the ROO filament and Binary models, respectively, **Fig. 5A**). Some of the cleaved DNA is still held in ROO filaments or Binary assemblies (blue and pink lines, respectively, **Fig. 5A**). The remainder of the cleaved DNA is bound to SgrAI, but is not in an assembly.

To investigate the origin of the advantage of the ROO filament mechanism, the Binary mechanism was altered to now have one feature unique to the ROO filament mechanism, namely the two ways each SgrAI/DNA (or E/DNA in the case of the Binary reaction) can come together (compare the red complexes of the ROO mechanism to the green ones of the Binary mechanism, **Fig. 6**). Most (but not all) of the advantage of the ROO model is lost (**Fig. 5B**). This implies that the run-on nature of the ROO filament (where SgrAI/DNA complexes may bind to either end of a filament, or even a single SgrAI/DNA complex) is the origin of the kinetic advantage over the Binary mechanism.

The simulations of **Fig. 5A-B** allowed for rebinding of the cleaved DNA, however, the simulation did not allow for the inclusion of the change in relative concentrations (local to actual). The initial concentration was set to 80 nM, to simulate the local concentration, however, upon cleavage and dissociation, local concentration effects no longer apply. The actual concentration of DNA in the cell was estimated at 3 nM

(for 1 copy per cell, see above). The K_D for SgrAI binding to PC DNA⁷ in the presence of 10 mM Mg^{2+} has been measured to be 14 nM. To simulate the change in concentration from 80 nM to 3 nM actual concentration, the rebinding rate constant was made ten-fold lower. **Figure 5C** shows the result. The total amount of cleaved sites with time is not affected by this change (dotted lines, **Fig. 5C**), however more of the cleaved DNA is shown free of SgrAI, as expected. This leads to lower final ROO filament and Binary complex concentrations (blue and purple lines, **Fig. 5C**). This simulation actually over-estimates the amount of ROO and Binary assemblies, since those would also change from local to actual concentrations upon separation of the DNA sites to which they bind, though this change in concentration was not included in the simulation. Further, cleaved DNA is expected to dissociate into the two cleaved products (an estimated K_D of the self-association of 375 nM²⁴) limiting reassociation to SgrAI.

To simulate DNA cleavage on separate DNA molecules (*i.e.* SgrAI bound to a site on a phage DNA and a site on the bacterial host genome), the simulation shown in **Fig. 5D** was performed. In this case, the initial DNA concentration was set to 3 nM. Only a small amount of DNA cleavage is seen in 200 s, the time it takes to cleave nearly 90% of the sites in **Fig. 5A**, indicating that minimal cleavage of sites on the host due to SgrAI activated by binding to sites on the phage DNA. Further, primary sites on the host DNA would be methylated and thereby protected, and unable to be cleaved via this pathway. Therefore, the sequestration effect is likely more important for protecting from cleavage of secondary sites on the host, since those would require ROO filament formation with SgrAI bound to unmethylated primary sites, only available on the invading DNA (see below for further discussion).

Finally, **Fig. 5D** also shows that the sequestration effect is not unique to the ROO filament mechanism, as the non-ROO Binary mechanism (green, **Fig. 5D**) also shows very low DNA cleavage at the low concentrations of DNA in the cell, but rapid DNA cleavage at the estimated local concentration when sites are present on the same contiguous DNA molecule (green, **Fig. 5A**).

Simulations with secondary site DNA sequences

To investigate cleavage of secondary sites, a modified kinetic model was created (**Table S5**). Secondary

sites differ from primary in one base pair, at either the 7th or 8th position of the recognition sequence (primary sites include CRCCGGYG, where R is A or G and Y is C or T, secondary sites include CRCCGGGG and CRCCGGYX, where X is A, C or T). Cleavage of secondary sites by SgrAI is nearly undetectable unless assembly into a ROO filament with SgrAI (bound to a primary site) occurs^{7, 16}. In that case, the observed rate constant for cleavage of secondary site DNA is generally ~two-fold slower than that of primary site and is dependent on the concentration of SgrAI bound to primary site DNA (hence the cleavage rate is also limited by ROO filament assembly). To model the cleavage of secondary sites by SgrAI, two changes were made to the kinetic model. First, assembly of SgrAI/secondary site DNA complexes into the ROO filament (or Binary complex) was set to two-fold slower than that of SgrAI/primary site complexes. Second, an additional equation is included to allow for the two types of Binary complexes: that with two SgrAI/primary site DNA complexes, and that with one primary and one secondary site bound SgrAI (the ROO mechanism already allows for both types of associations). **Tables S6-S7** give the rate constants and concentrations of species used in the simulations. In addition, using the same logic as shown in **Table S2**, the local concentration of secondary sites relative to primary was calculated to be 1 μ M (for the nearest pairs of sites).

Simulations were carried out at 1 μ M DNA concentration (solid lines, **Fig. 7A-B**) to simulate cleavage of secondary sites in the phage DNA, and 3 nM (dotted lines, **Fig. 7A-B**) to model cleavage of secondary sites on the host genome. In these simulations, it is assumed that the only source of primary sites (which induce the formation of assemblies, *i.e.* ROO filaments or Binary complexes) are from the phage DNA. Orange lines represent total DNA cleavage with the ROO filament mechanism, light blue represents that for the Binary. **Figure 7A** shows the results when the two mechanisms have the same rate constants. Cleavage at 1 μ M DNA (representing phage DNA) is faster (2-fold) in the ROO filament mechanism (orange solid line, **Fig. 7A**), and very little cleavage occurs with 3 nM DNA (representing host DNA) in both mechanisms (dotted lines, **Fig. 7A**), demonstrating the sequestration of activated DNA cleavage away from the host DNA. Although very little, **Fig. 7A** shows that some cleavage of the host DNA (orange and blue dotted lines) is predicted. Upon cleavage of primary sites in phage, the phage DNA should dissociate from SgrAI and separate into fragments and undergo degradation³⁷, with little reassociation with SgrAI and into ROO filaments (blue line, **Fig. 5C**),

ending the threat to host DNA. At 100 s, the time it takes for most primary sites in phage to be cleaved (red dotted line, **Fig. 5A**), ~4% of host secondary sites are predicted to be cleaved as well (orange dotted line, **Fig. 7A**). There are ~2500 predicted secondary sites in *S. griseus*, and 4% is ~100 secondary sites. Studies have shown that such “autoimmunity” (*i.e.* cleavage of host DNA by REs) does occur in RM systems, particularly in the case of more “efficient” endonucleases, and that double stranded breaks in the host genome are repaired via the SOS/RecA repair pathway³⁸. This “autoimmunity” is tolerated under rapid growth conditions and high nutrient availability, although is less tolerated when resources are limited³⁸⁻⁴⁰.

Figure 7B shows the results when the assembly rate constant of the Binary mechanism is increased until cleavage at 1 μ M DNA (solid lines, **Fig. 7B**) matches that of the ROO filament mechanism. Now the Binary mechanism shows much greater cleavage at 3 nM (light blue dotted line, **Fig. 7B**) compared to that for the ROO filament mechanism (orange dotted line, **Fig. 7B**). Hence these simulations show that the ROO filament mechanism is also superior to the Binary in sequestering activated DNA cleavage away from the host DNA. Though DNA cleavage by the ROO filament mechanism is twice as fast as that by the Binary mechanism when utilizing all the same rate constants (**Fig. 7A, Table S6**), a 4.6-fold increase in the rate constant for assembly of enzyme-DNA complexes into binary complexes is necessary to achieve the same rate of secondary site cleavage by the two mechanisms (**Fig. 7B**, see rate constants k_3 and k_4 , **Table S6**). This is likely due to the competition for association that occurs between the two types of enzyme complexes (bound to primary or to secondary sites) that necessarily occurs in the Binary mechanism, and the fact that enzymes bound to primary sites preferentially self-assemble rather than associating with enzymes bound to secondary (**Fig. 8, left**). Competition is not necessary in the ROO filament mechanism, since any complex can assemble at either end of the ROO filament (**Fig. 8, right**).

DISCUSSION

Compared to other bacterial immune proteins, SgrAI is unusual in several respects (its low unactivated DNA cleavage rate and allosteric activation via filament formation with sequence specificity expansion) though exhibits similar DNA cleavage rates when activated. For example, the unactivated SgrAI DNA

cleavage rate is very slow (0.1 min^{-1}), but similar to other type II REs under activating conditions (22 min^{-1} , compared to 20 min^{-1} and 36 min^{-1} for the type II REs EcoRI and EcoRV, respectively^{7, 41, 42}). Note that the single turnover DNA cleavage rate constant measures all steps from DNA binding to DNA cleavage, however global kinetic data fitting indicates that the chemical cleavage step (once within the filament) is faster yet in SgrAI, estimated at 48 min^{-1} ²⁵. Cas9, a bacterial immune defense enzyme from the *Streptococcus pyogenes* CRISPR system, and of interest in genome engineering applications^{43, 44}, has been found to have single turnover DNA cleavage rate constants of 60 min^{-1} and 30 min^{-1} for cleavage by the two endonuclease domains (HNH and RuvC), respectively⁴⁵. Cas9, however, remains tightly bound to the cleaved product DNA, while SgrAI (and other type II REs) rapidly dissociate the cleaved DNA product ($>24 \text{ min}^{-1}$), following dissociation from the ROO filament (1.8 min^{-1}), freeing it to perform multiple rounds of enzymatic turnover^{25, 45, 46}.

Herein we include both experiments and simulations to show that the run-on oligomer (ROO) filament mechanism of SgrAI possesses unique characteristics that appear to have evolved to perform the requirements of its biological niche. First, based on the CryoEM structural model of the ROO filament formed by the assembly of SgrAI/DNA complexes⁸, point mutations were designed with the intention of disrupting interfaces and thereby destabilizing this assembly. As a result, the DNA cleavage properties of SgrAI, which are stimulated 200-1000 fold^{7, 16} in the ROO filament, are predicted to be hampered by these mutations since they weaken filament formation. In fact, most mutations diminished the activated DNA cleavage rate (**Table 1**), as predicted. Further, they did not affect the basal, unactivated DNA cleavage rate, showing that the mutations do not disrupt DNA binding or normal, unactivated DNA cleavage by SgrAI when not in the ROO filament. Hence these mutations must disrupt ROO filament formation since activation seen under conditions where wild type SgrAI forms the ROO filament are not found.

The mutant SgrAI enzymes were then tested for their ability to protect host cells from invading phage DNA (*i.e.* the phage challenge assay). Western blots were performed to verify expression of mutant proteins (**Fig. 3**). Expression of wild type and mutant SgrAI was induced overnight, then cells mixed with a modified form of lambda phage incapable of forming lysogens (λ JL801)³³ so that all productive infections can be counted by counting plaques formed from cell lysis. The parent host strain, without SgrAI, showed levels of

phage infection of $5.7 \pm 1.1 \times 10^4$ PFU/ μ l with our stock of purified λ JL801, however the expression of wild type SgrAI in these cells resulted in complete protection from infection as far as we could measure as no plaques were found, even with the highest concentrations of phage tested (this represents a $>10^4$ -fold protection relative to the parent strain, within ranges seen with other RM systems⁴⁷). In the case of the mutant SgrAI enzymes, all but one allowed phage infection to proceed, indicating dramatically diminished protection. Many SgrAI mutants showed plaque numbers within error of the unprotected parent strain (**Table 1**), and these mutant SgrAI also showed the lowest levels of activated DNA cleavage (**Fig. 2**). Some mutants with intermediate levels of activation (*i.e.* activated DNA cleavage rate constants 3-8 fold slower than that of wild type but still accelerated relative to the basal DNA cleavage rate), such as R84E, M62E, R127A, S56Q, showed perhaps some protection from phage infection, though the protection is very weak. Only A57Q, which showed activated DNA cleavage within error of that of wild type, gave complete protection from phage infection (**Table 1**, **Fig. 2**). Hence, mutations that disrupt the ROO filament of SgrAI, disrupt accelerated DNA cleavage, and importantly, these mutations also diminish the enzymes' ability to protect its host cell from phage infection. We find that rather than a linear relationship between the activated DNA cleavage rate of SgrAI and protective ability, a step function better describes the plot in **Fig. 2**, with the most active of the affected mutations (R127A, at 35% of WT activity) still completely ineffective at protecting its host cell against phage infection. This suggests that SgrAI must be faster than the R127A mutant in order to be effective against phage infection.

Biological relevance of the phage challenge assay

The phage challenge assay was used to measure the ability of SgrAI (wild type and mutant) enzymes to protect host cells from phage infection. Because phage of the natural host of SgrAI, *Streptomyces griseus*, have not been well characterized, this assay was instead performed in *E. coli*. Further, the cognate SgrAI methyltransferase has not been identified, therefore the methyltransferase MspI.M was used instead to create a functional RM system in *E. coli*. MspI.M methylates the C5 position of the first C in CCGG sequences, hence protects both primary and secondary site sequences within the *E. coli* genome. It is expressed from a plasmid derived from the low copy

number plasmid pACYC184, and from its natural promoter which shows near consensus -35 and -10 sequences³⁴.³⁵. In contrast, SgrAI proteins are expressed from a high copy number vector with an inducible T7 promoter giving high levels of expression⁴⁸. Relative expression levels of R (endonuclease) and M (methyltransferase) enzymes are important, too much R leads to “autoimmunity” (cleavage of host DNA)³⁸. Too little risks poor anti-phage activity due to slower cleavage (from low levels of R) and higher possibilities of escape via methylation of the invading phage DNA by the methyltransferase. Interestingly, studies show that some “autoimmunity” occurs in natural systems with some REs, and is repaired efficiently via the SOS/RecA pathway under nutrient rich conditions, but not nutrient limited³⁸⁻⁴⁰. The RM system used here is more likely to be tilted towards the R enzyme due to the overexpression of SgrAI proteins, compared to the constitutive expression of MspI.M. Even so, most mutant SgrAI enzymes were unable to show any protection against phage infection.

The second key element of the assay is the phage, and lambda phage λ JL801 was chosen due to its well-studied biology and the inability of this mutant to form lysogens. Lambda phage contains 6 primary sites and 32 secondary sites in its 50 kb genome, similar in size to two known phage of *Streptomyces coelicolor*, R4 and PhiC31 (50 and 40 kb, respectively)^{49, 50}. Hence, the phage challenge assay developed here recapitulates the necessary elements to test the ability of SgrAI to protect its host organism from phage infection.

Simulation of in vivo activity of SgrAI

Previous work used kinetic modeling to investigate the ROO filament mechanism of SgrAI, and when applied to *in vitro* data, allowed for the extraction of microscopic rate constants for each step of the reaction mechanism^{24, 25}. Using those mechanistic models and experimentally determined rate constants, simulations were performed here to model reactions as they may occur *in vivo*. To better mimic the *in vivo* reaction, only a single type of recognition site is used in this simulation, rather than the two types of DNA used in *in vitro* reactions (*i.e.* the reporter and the activator DNA). Also, to better mimic the *in vivo* reaction, DNA concentrations used were those estimated for the concentration of a single copy of DNA in the cell (3 nM), and that estimated for the “local concentration” of recognition sites with respect to each other on the same molecule of DNA (80 nM for primary to primary sites, 1 μ M for primary to secondary sites). In addition, to

compare the ROO filament mechanism to non-ROO mechanisms, we have also constructed a simple “Binary” model. In the Binary model, activation occurs when two hypothetical enzymes (“E”), each bound to a DNA site, associate into a dimer. All rate constants are otherwise identical in the simulations with both mechanisms. Finally, simulations to predict cleavage rates of secondary site sequences were also performed.

Advantages of the ROO mechanism

Finally, the question of advantage of the ROO filament mechanism is addressed by comparing DNA cleavage kinetics via both the ROO filament and non-ROO (Binary) mechanisms. Clearly, the ROO filament mechanism has the advantage (red vs. green solid and dashed lines, **Fig. 5A**) in both the rate of cleavage of DNA, and cleavage and release of DNA, under these reaction conditions. Such an advantage may be necessary where speed is required, such as in the race against viral replication (and methylation by the SgrAI methyltransferase) of invading DNA. The rate limiting assembly of SgrAI/DNA complexes into the ROO filament, which is required to provide sequestration of activated SgrAI, may be a limiting parameter for speed but is compensated for by this kinetic advantage of the ROO mechanism, explaining why a simple Binary mechanism is inferior (though it also possesses the sequestration effect). The origin of the kinetic advantage is found to be due to the two ways SgrAI may come together (**Fig. 6**). Hence the formation of a large ROO filament is not necessary for this kinetic advantage. Indeed, allowing the Binary (non-ROO) mechanism two ways for enzyme association to occur reduced most of the advantage of the ROO filament mechanism (**Fig. 5B**). A small additional advantage derives from the third or more enzyme additions to the ROO filament. Our results with mutants that disrupt the ROO filament of SgrAI show that the very fast DNA cleavage performed by activated SgrAI is critical for protection against phage infection, since even the fastest effected mutant, R127A, with 35% of WT activity, is completely ineffective (**Fig. 2**).

Why not merely evolve a faster Binary reaction?

Secondary sites differ from primary in one base pair, occurring at either the 7th or 8th position of the 8 bp recognition sequence. These sites are not appreciably cleaved by SgrAI without assembly into ROO

filaments composed of SgrAI bound to primary site DNA^{7, 16}. Primary sites on the host DNA are methylated, thereby protected from cleavage by SgrAI, but secondary sites are likely not, and therefore potentially susceptible to cleavage by activated SgrAI. Hence the need for sequestration is most relevant to the prevention of cleavage of secondary site sequences on the host DNA. We hypothesized that the ROO filament mechanism may have evolved to increase its ability to sequester the DNA cleavage activity of SgrAI to the same copy of DNA containing the activating primary sites, which *in vivo* would be the invading phage DNA.

This sequestration would serve to protect the host DNA from damaging cleavage at secondary site sequences. The simulations show that the ROO filament mechanism is two-fold faster than the Binary due to the two ways assembly may occur (**Fig. 6**) each time a SgrAI/DNA complex adds to the ROO filament. In fact, the simulations show that when the Binary reaction is allowed a faster (4.6-fold) assembly rate constant such that its accelerated DNA cleavage rate matches that of the ROO filament mechanism (solid lines, **Fig. 7B**), greater cleavage of secondary sites on the host DNA compared to the ROO filament mechanism is predicted (light blue dotted line, **Fig. 7B**). Thus, these simulations predict that the ROO filament mechanism is superior in both rapid DNA cleavage and in sequestration. The origin of both advantages derives from the multiple ways SgrAI/DNA complexes can assemble in the ROO filament (**Fig. 6, Fig. 8**). Secondary site bound SgrAI (orange, **Fig. 8**) need not compete with primary site bound SgrAI (red, **Fig. 8**) for association into ROO filaments, as assembly interfaces are always available. In contrast, the Binary complex associates more strongly when both SgrAI/DNA complexes contain primary site sequences (green, **Fig N8**), and SgrAI bound to secondary site DNA (light blue, **Fig. 8**), must compete with this complex for association with SgrAI/primary site DNA complexes.

CONCLUSION

The ROO filament mechanism of SgrAI appears to have evolved out of the phage-host competition or “arms race”, one of the oldest evolution-coevolution systems in evolutionary history¹, to accommodate specific challenges. These include reduced activity on the host genome through both a slower DNA cleavage rate and a longer (hence rarer) recognition site, as well as the ability to become rapidly activated on DNA

containing multiple unmethylated primary sites, such as invading DNA, through enzyme assembly into ROO filaments. SgrAI must sequester its DNA cleavage activity on the invading DNA, and does so by a slow association rate constant that limits assembly into the ROO filament to only those SgrAI bound to sites on the same molecule of DNA. However, speed is critical to performing SgrAI's biological role (as evidenced by the mutant study described herein), hence ROO filament formation allows for multiple ways for the enzymes to assemble, thereby increasing the rate of this step significantly. The ability to form longer ROO filaments provides for additional speed, and relatively fast dissociation prevents trapping of what may be limiting amounts of SgrAI in the cell²⁵. Merely increasing the assembly rate constant in a Binary mechanism does not match the advantage of the ROO filament mechanism, and loss of sequestration results. Hence the ROO filament mechanism is superior to the Binary (non-filament) mechanism in both speed and sequestration, both of which are important to the biological function of SgrAI.

As discussed in prior work^{24, 25}, recognition of the filament forming enzyme mechanisms has until recently been largely limited to the cytoskeletal ATPases and GTPases such as actin and tubulin. However, large-scale screening using newer imaging technologies has allowed for the observation of filament formation by metabolic and other enzymes, previously unknown to form such structures^{11, 13, 14}. The details of the roles of the filaments in enzyme activity have been investigated in only a handful of such enzymes^{9, 12, 51-54}, and here we provide the most detailed kinetic and mechanistic investigation of an enzyme filament mechanism to date. The ROO filament of SgrAI no doubt has unique features compared to the other systems, for example, SgrAI is unlikely to form the large-scale filaments seen in many of the fluorescence microscopy studies which persist over minutes to hours, yet common features likely also result. The detailed structures of run-on oligomers and filaments of most filament forming enzymes are also not known, exceptions include SgrAI, CTP synthase, Acetyl-CoA carboxylase⁵⁵ and Ire1, the unfolded response nuclease-kinase^{8, 9, 51, 52}. Our kinetic investigations complements our structural work⁸ in elucidating the features of the SgrAI ROO filament that make it unique, optimized for its biological niche, and advantageous over other mechanisms.

MATERIALS AND METHODS

Protein Purification

Wild-type and mutant SgrAI proteins were expressed with a C-terminal his tag in Tuner (DE3) *E. coli* which also contained the pLysS plasmid (Novagen, Inc.) and the MspI.M expression plasmid (pBAK.MspI)³⁴.⁵⁶ MspI.M methylates at the C5 position of the first cytosine of CCGG sequences and is expressed from its natural promoter in plasmid pBAK.MspI, a derivative of pACYC184 with the natural coding sequence of MspI.M^{34, 35}. Mutagenesis was performed as described previously¹⁶, and all expression vectors were sequenced fully in the SgrAI gene to verify the point mutation. The proteins were purified using Talon metal affinity resin (Clontech, Inc.) followed by ion-exchange FPLC using heparin resin (GE Healthcare Life Sciences). First, the cell lysate was incubated with Talon resin in lysis buffer (50 mM sodium phosphate buffer (pH 8.0@RT), 800 mM NaCl, 10 mM imidazole, and 1 mM BMe) 30 min to overnight. The unbound cell lysate was washed away using wash buffer (50 mM sodium phosphate buffer (pH 8.0@RT), 300 mM NaCl, 20 mM imidazole, and 1 mM BMe) followed by high salt wash buffer (50 mM sodium phosphate buffer (pH 8.0@RT), 2 M NaCl, 20 mM imidazole, and 1 mM BMe). Finally, the protein was eluted using elution buffer (50 mM sodium phosphate buffer (pH 8.0@RT), 300 mM NaCl, 250 mM imidazole, and 1 mM BMe). For ion exchange FPLC purification, the protein was excessively dialyzed into Heparin A buffer (50 mM Tris-HCl (pH 8.0@RT), 50 mM NaCl, 0.1 mM EDTA, 10 mM BMe), then purified using Heparin FF chromatography (GE Healthcare Biosciences) and a gradient of Heparin B buffer (50 mM Tris-HCl (pH 8.0@RT), 1 M NaCl, 0.1 mM EDTA, 10 mM BMe). Purity of the protein was confirmed using SDS-PAGE. The purified protein was then aliquoted into single use aliquots, flash frozen in liquid nitrogen, and stored at -80°C.

DNA Preparation

The oligonucleotides were made synthetically and purified using C18 reverse phase HPLC or denaturing PAGE⁵⁷. The concentration was measured spectrophotometrically, with an extinction coefficient calculated from standard values for the nucleotides⁵⁸. The self-complementary DNA strands, or equimolar quantities of complementary DNA, were annealed by heating to 90°C for 10 minutes at a concentration of 0.1-1 mM,

followed by slow-cooling to 4°C over 4-5 hours in a thermocycler or heat block. Sequences of the DNA used are given in Table 2 (recognition sequences are shown in red, | marks the SgrAI specific cleavage site).

The 5' ³²P end labeling of DNA was performed with T4 polynucleotide kinase (New England Biolabs) and [γ -³²P]-ATP (Perkin-Elmer, Inc.), followed by removal of excess ATP using P-30 spin columns (Bio-Rad Laboratories, Inc.).

Western analysis to measure protein expression levels

Western blots were performed with lysates from cells used in the phage challenge assays, with the OD of the cells noted before pelleting 1 ml of overnight growths, reconstituting in protein loading buffer (25 mM Tris-HCl pH 6.8, 2% SDS, 6% glycerol, 0.1 M DTT, 0.004% bromophenol blue), heating to 90°C for 5 minutes, and centrifugation at 10,000 rpm for 10 minutes. SDS-PAGE was performed on the samples followed by transfer to PVDF membrane (ThermoFisher Scientific) after soaking in Towbin buffer (25 mM Tris, 192 mM glycine, 20% methanol, 0.1% SDS). Blots were blocked with 3% BSA in PBST (80 mM Na₂HPO₄, 20 mM NaH₂PO₄, 100 mM NaCl, 0.2% Tween), followed by probing with HRP labeled primary antibody (mouse anti-his monoclonal antibody, MA1-135, ThermoFisher, Inc.) at a 1:1000 dilution, then washed 3 times in PBST. For visualization of bands, the blot was soaked in chemiluminescence solution (made from mixing equal parts of 0.78 mg/ml luminol, 0.95 mg/ml p-iodophenol, 0.1 M Tris-HCl pH 9.35 with 0.03% H₂O₂ in 0.1 M Tris-HCl pH 9.35) for 1 minute followed by imaging using a Chemidoc scanner (BioRad Laboratories, Inc.). Bands corresponding to SgrAI proteins were integrated using Image Lab (BioRad Laboratories, Inc.), corrected for dilution, OD, and normalized to that quantitated for wild type SgrAI.

DNA Cleavage Assays

Single turnover kinetic measurements of DNA cleavage were performed using ³²P-labeled oligonucleotides substrates (1 nM), under conditions of excess enzyme (1 μ M SgrAI dimer), with and without the addition of unlabeled PC DNA. All DNA cleavage reactions were performed at 37°C in 20 mM Tris-

HOAc (pH 8.0@RT), 50 mM KOAc, 10 mM Mg(OAc)₂, and 1 mM DTT. 5 µl aliquots were withdrawn at specific time intervals after mixing the enzyme and labeled DNA (100 µl total reaction volume), quenched by addition to 5 µl of quench solution (80% formamide, 50 mM EDTA, 1 mg/ml XCFF dye, and 1 mg/ml BPB dye), and electrophoresed on denaturing polyacrylamide gels (20% acrylamide:bisacrylamide (19:1 ratio), 4 M urea, 89 mM Tris base, 89 mM boric acid, 2 mM EDTA). Autoradiography of gels was performed without drying using a phosphor image plate exposed at 4°C for 12-17 hours. Densitometry of phosphor image plates was performed with a phosphorimager (GE Healthcare Life Sciences, Pittsburgh, PA, USA or Bio-Rad, Inc., Hercules, CA, USA), and integration using Image Lab software (Bio-Rad, Inc., Hercules, CA, USA). The percent of product formed as a function of time was determined by integrating the density of both cleaved and uncleaved DNA bands and normalizing to the total amount cleaved. The percentage of cleaved DNA was then fit to a single exponential function to determine the single turnover rate constant of DNA cleavage using Kaleidagraph (Synergy Software, Reading, PA, USA):

$$\text{Percentage of Cleaved DNA} = C_1 + C_2 \times (1 - e^{-kt})$$

where C_1 is a constant fitting the baseline, C_2 is the total percent of DNA predicted to be cleaved by SgrAI, k is the cleavage rate constant, and t is the length of incubation in minutes.

Phage Challenge Assay

Tuner (DE3) *E. coli* (Novagen) were transformed with (pBAK.MspI)(coding for MspI methyltransferase expression, New England Biolabs, Inc.) and pET.21a_SgrAIR (coding for his-tagged SgrAI expression, wild type or mutants), grown overnight in 6 ml LB culture with 50 µg/ml ampicillin, 30 µg/ml kanamycin, 0.2% maltose, and 1 mM MgSO₄, then grown overnight with induction in LB (with 0.4 mM IPTG). Cells were chilled to 4°C, centrifuged at low speed (4000 rpm for 10 minutes), supernatant removed and the cells resuspended in 1/10 volume of TMG (10 mM Tris-HCl, pH 8, 10 mM MgSO₄, 0.01% gelatin, sterile). Top agar (10 g/L agar, 10 g/L NaCl, 10 g/L tryptone) was prepared and aliquoted (3 ml) into screw-cap culture tubes, autoclaved, and kept at 50°C prior to use. Phage (λJL801, with the first 4 codons of

the cI protein coding region deleted to result in a purely lytic form of lambda phage without lysogeny³³) was prepared from large scale infections of Tuner (DE3) (without MspI.M or SgrAI) and kept in TMG at 4°C. 10-fold dilutions of the phage solution were prepared in TMG. For plating cells with each dilution of phage, 0.1 ml of the cells and 1 µl of the phage solution were added to the side of a tube containing the top agar, vortexed gently to mix, and quickly poured onto the top of pre-warmed agar plates (20 g/L agar, 10 g/L NaCl, 10 g/L tryptone in 100 mm x 15 mm dishes). Plates were incubated overnight at 37°C and plaques counted to give PFU (plaque forming units) per µl of phage stock. Triplicate measurements were done for each SgrAI protein (wild type or mutant) at the best dilution of phage (giving plaques between 25-250 per plate).

The phage stock was prepared by infecting the parent strain (Tuner (DE3)) with λJL801 (provided by J. Little³³) and plating in top agar as described above. Plates were incubated 2 hours with 3 ml of TMG at 37°C and gentle shaking to elute the phage particles. 1 ml of eluate was then transferred to a 1.5 ml tube, and 50 µl of chloroform added followed by brief vortexing. The samples were then centrifuged for 10 minutes at 10,000 rpm, and the aqueous fraction transferred to a new 1.5 ml tube with 30 µl chloroform added. Concentration (in terms of PFU per µl) was measured as described above for the phage challenge assay, and the same stock was used in all tests.

In vivo DNA cleavage reaction simulations

Kintek Global Kinetic Explorer (version 6.2.170301) (Kintek Global Kinetic Explorer Corp.)^{59, 60} was used for the simulations. Equations for modeling are given in the Supporting Information. Rate constants are also given in the Supporting Information and are those derived from prior work²⁵. Modeling of cleavage of primary sites in phage and in host DNA (as a result of activation via primary sites on phage DNA) used the equilibria found in **Table S3**. Two mechanisms are present in this case, that for the ROO filament mechanism, and that for the Binary. The Binary mechanism is simpler, and involves DNA (“site”) binding by the hypothetical enzyme “E” to create the enzyme/DNA complex “R”. Two Rs may associate to give the Binary complex “RR”. DNA cleavage only occurs in this Binary complex, and is symbolized by the conversion of

“R” to “X”. This occurs independently for the two “R” in the Binary complex “RR”. The Binary complexes with cleaved DNA may also dissociate before or after cleavage occurs, but dissociation of cleaved DNA (“cleaved DNA”) from “X” only occurs when X is isolated from the Binary complex. The forward and reverse rate constants for each equilibria are numbered and the values given in **Table S4** (and are those derived from fitting experimental data of SgrAI reactions in prior work²⁵). The ROO filament mechanism is more complicated, although software limitations prevented the modeling of ROO filaments longer than 3 SgrAI/DNA complexes long, similar to the predicted number of primary sites in the typical *Streptomyces* phage (see **Table S2**). Identical nomenclature is used in **Table S3** for this mechanism, however complexes of “R” and “X” include those of size 3.

The kinetic model used for simulating the cleavage of secondary sites is given in **Table S5**, and corresponding rate constants and starting concentrations given in **Tables S6-S7**. Slightly different nomenclature is used for species to denote the difference between SgrAI/DNA (or E/DNA) complexes: P for that bound to primary site DNA, and S for that bound to secondary. In this model complexes with primary site DNA (*i.e.* “P”) may self-associate and may also associate with complexes containing secondary site DNA (*i.e.* “S”). Complexes with secondary site do not self-associate, consistent with experimental observations of SgrAI activity^{7, 16}. To reduce the number of equilibria in the modeling, only cleavage of secondary sites is considered, however this is realistic since even cleaved primary sites will bind SgrAI and induce ROO filament formation⁸. This allows modeling up to ROO of 4 (in the ROO filament mechanism). Again, cleavage of secondary site DNA in complexes is symbolized by the conversion of “S” to “X” (**Table S5**). To estimate rate constants for complexes with secondary site DNA, a preliminary fitting of single turnover DNA cleavage data was performed¹⁶. As with primary site DNA, the apparent DNA cleavage rate constants of secondary site cleavage by SgrAI are dependent on the concentration of SgrAI bound to primary site DNA, hence are also rate limited by the association step (k_4 in **Table S5**) in ROO filament formation, however is approximately two-fold slower. Hence the rate constant k_4 was set to 2-fold lower than that for complexes with only primary site DNA (*i.e.* k_3). All other rate constants were held the same for complexes

with secondary site as those with primary site DNA. This assumes that the only effect of secondary site DNA on SgrAI is on the rate of assembly of ROO, and not on its dissociation or DNA cleavage.

ACKNOWLEDGEMENTS

We would like to acknowledge Prof. John Little, University of Arizona, for his kind gift of λ JL801.

TABLES

Table 1. DNA cleavage rate, phage titer, and protein expression level analyses of SgrAI enzymes.

SgrAI Enzyme	Phage Titer ^a (PFU/μl)	Unactivated DNA Cleavage Rate Constant ^b (min ⁻¹)	Activated DNA Cleavage Rate Constant ^c (min ⁻¹)
WT	NP ^d	0.094±0.15 ^e	22±7 ^e
T4D	3.8±0.5x10 ⁴	0.058±0.002	0.097±0.009
S6D	4.3±1.0x10 ⁴	0.03±0.01	0.19±0.06
I7E	4.5±0.3x10 ⁴	0.091±0.008	0.47±0.09
R24E	3.7±0.6x10 ⁴	0.014±0.003	0.028±0.004
N25E	5.4±0.7x10 ⁴	0.05±0.01	1.2±0.2
P27W	6±1x10 ⁴	0.037±0.05 ^f	0.14±0.01 ^f
P27G	4.6±0.4x10 ⁴	0.06±0.002 ^f	0.12±0.003 ^f
Q34D	6.4±1.2x10 ⁴	0.04±0.02	1.180±0.2
I51E	4.9±0.1x10 ⁴	0.05±0.01	0.26±0.07
S56E	4.9±0.3x10 ⁴	0.08±0.01 ^g	0.17±0.03 ^g
S56Q	2.9±0.3x10 ⁴	0.08±0.02 ^g	5.5±1.8 ^g
A57E	4.6±1.4x10 ⁴	0.098±0.002 ^g	0.39±0.06 ^g
A57Q	NP ^d	0.09±0.01 ^g	15±8 ^g
I59E	6.3±0.2x10 ⁴	0.14±0.03	0.026±0.006
M62E	2.6±0.8x10 ⁴	0.021±0.001	3.2±0.4
R84E	3.6±0.5x10 ⁴	0.082±0.005	2.6±0.2
R127A	6±3x10 ⁴	0.04±0.04	7.6±0.6
R131A	5.6±0.5x10 ⁴	0.10±0.01 ^g	0.28±0.02 ^g
R134A	5.2±0.8x10 ⁴	0.10±0.013 ^g	0.8±0.2 ^g

^aParent strain (Tuner (DE3) (Novagen, Inc.) with MspI.M) is 5.7±1.1x10⁴ PFU/μl.

^bSingle turnover DNA cleavage rate constant of 1 nM ³²P labeled 18-1 with 1 μM SgrAI enzyme.

^cSingle turnover DNA cleavage rate constant of 1 nM ³²P labeled 18-1 with 1 μM SgrAI enzyme and 1 μM PC DNA (activating DNA).

^dNP, no plaques detected.

^eFrom Park, *et al.*, 2010⁷.

^fFrom Park, *et al.*, 2010⁶¹.

^gFrom Shah, *et al.*, 2015¹⁶.

Table 2. DNA sequences used in the single turnover DNA cleavage assays

Name	Sequence
PC-top	5'-GATGCGTGGGTCTTCACA-3'
PC-bot	3'-CTACGCACCCAGAAGTGTGGCC-5'
18-1-top	5'-AAGTCCA CCGGTGGACTT-3'
18-1-bot	3'-TTCAGTGTGGCC ACCTGAA-5'

FIGURES

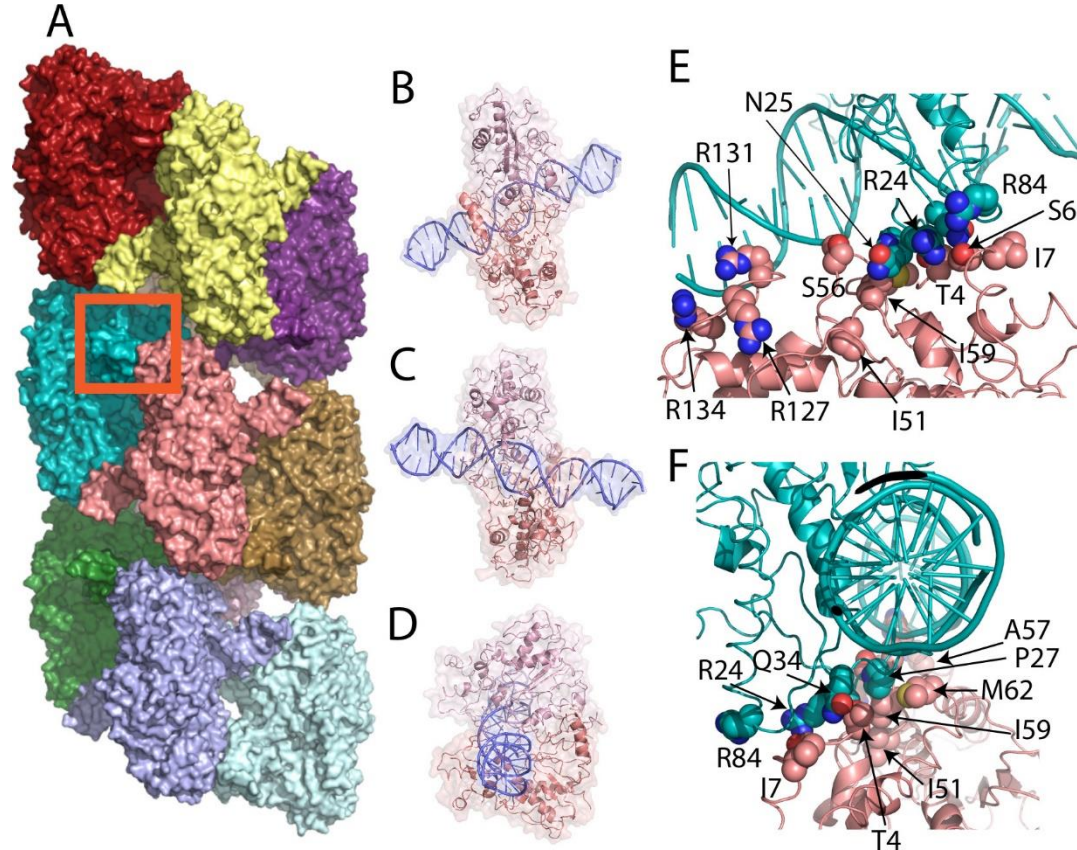


Figure 1. Run-on oligomer (ROO) filament of SgrAI/DNA complexes and sites of point mutations. A. The run-on oligomer filament structure with 11 SgrAI/DNA complexes, each colored separately. The red box indicates the area shown in E-F. **B.** Cartoon and surface rendering of one SgrAI/DNA complex, in the same orientation as the salmon colored complex in A. **C.** View rotated 180° about a vertical axis relative to that in B. **D.** View rotated 90° from that in C, showing the same orientation as the lightest blue colored complex in A. **E.** Zoom in on the boxed area of A, showing the interface between two adjacent SgrAI/DNA complexes within the run-on oligomer filament. Selected residues mutated in this study are indicated. **F.** Approximately 90° rotation from B, showing the positions of selected residues mutated in this study.

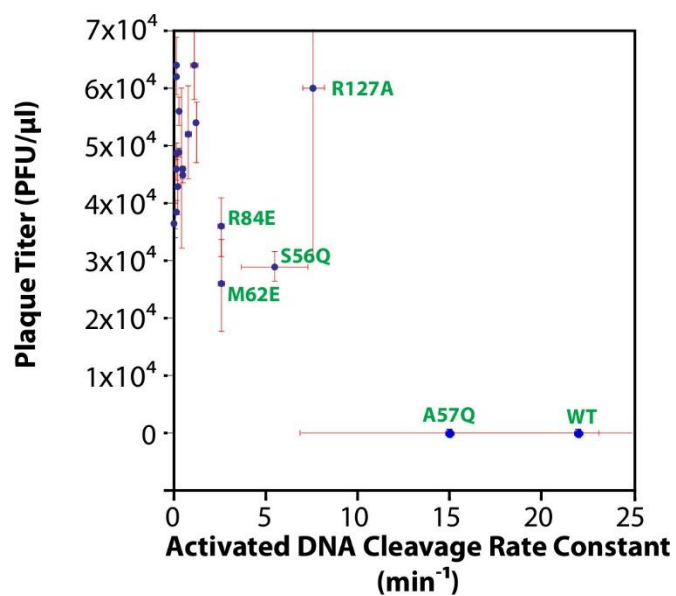


Figure 2. Plot of Phage Titer vs. Activated Single Turnover DNA Cleavage Rate Constant. Error bars show ± 1 standard deviation (red). Data points for selected mutant or wild type SgrAI enzyme are labeled (green).

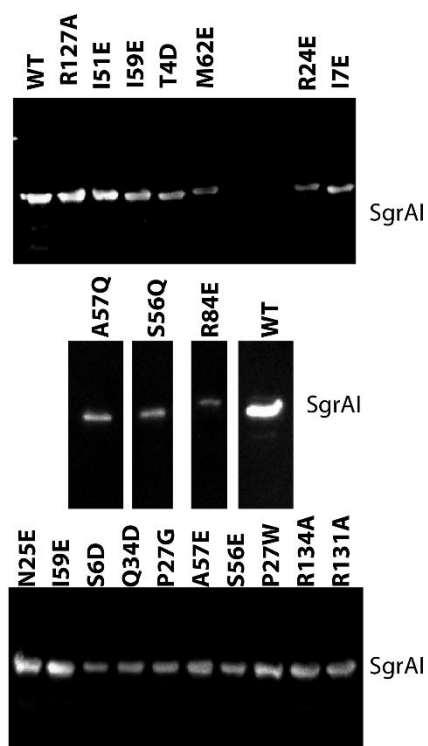


Figure 3. Western analysis of wild type and mutant SgrAI expression levels. Middle gel is cut to remove irrelevant lanes.

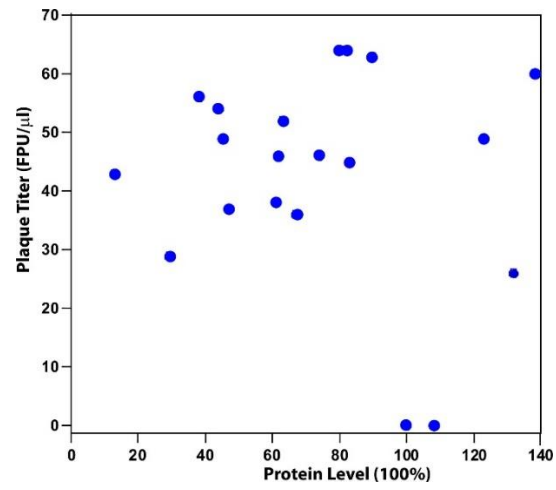


Figure 4. No correlation is found between plaque titer and protein expression level (levels normalized to that of wild type SgrAI).

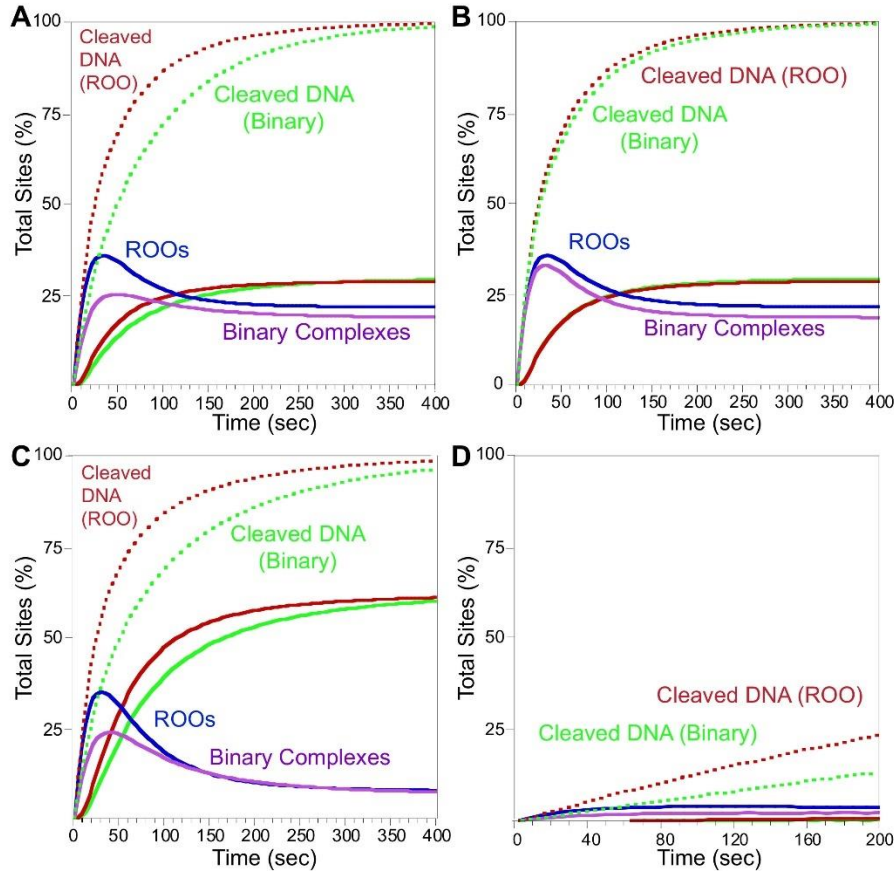


Figure 5. Model of Phage DNA cleavage and comparison of ROO and Binary mechanisms. Cleaved DNA bound to SgrAI or the Binary enzyme shown in dotted red (ROO mechanism) or green (Binary mechanism). Cleaved DNA released from SgrAI or the Binary enzyme shown in solid lines (red, ROO mechanism, green, Binary). Concentrations of ROO filaments (with cleaved and uncleaved DNA) shown blue, and Binary complexes in purple. See **Tables S3-S4** for model details and rate constants. **A.** Simulation of DNA sites in phage DNA. Starting concentrations of DNA 80 nM, the estimated local concentration of SgrAI bound to two primary sites present on the same DNA molecule. **B.** As in A, but the Binary non-ROO mechanism is now set to allow for two ways to form assemblies. **C.** As in A, however rebinding of cleaved DNA set to a lower off rate constant, to mimic the lower concentrations of DNA free in the cell (see text for details). **D.** As in A, but with 3 nM DNA to mimic reactions between separate molecules in the cell (*i.e.* host genome and phage DNA).

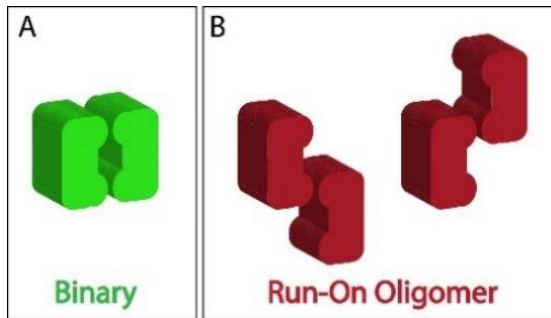


Figure 6. Possible association patterns in a simple Binary association and the run-on oligomer filament mechanism. **A.** Only one type of association is found with the closed-ended Binary model. **B.** Two ways are found for enzyme complexes to associate in the run-on oligomer (ROO) filament mechanism.

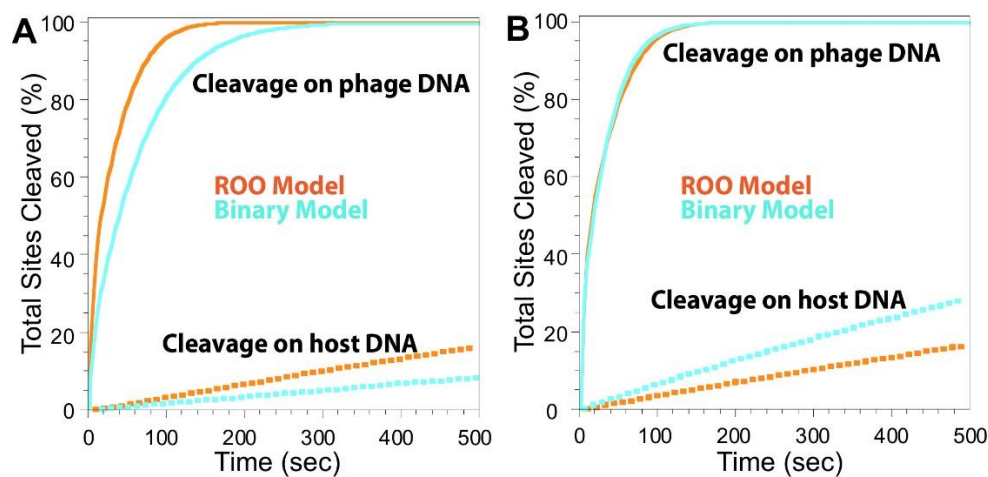


Figure 7. Total cleaved secondary site DNA from the ROO filament and Binary mechanisms with same and different assembly rate constants. A. Solid lines simulate cleavage on a phage DNA, and dotted lines simulate cleavage of sites on host DNA due to activation by phage DNA. Light blue, total cleavage sites from Binary mechanism, orange, total cleaved sites from ROO filament mechanism. Equations, rate constants and concentrations given in **Table S5-S7**. **B.** As in A, but with rate constants of the Binary mechanism increased 4.6-fold to give the same rate of cleavage on the phage DNA as the ROO filament mechanism. More cleavage of host DNA occurs with the Binary model (compare light blue dotted line to orange dotted line).

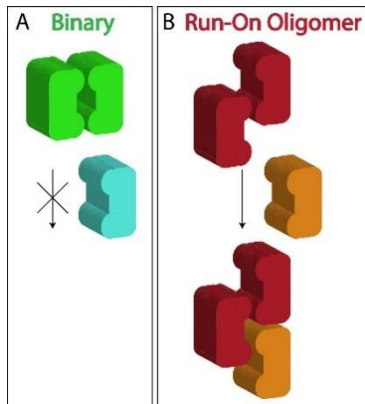


Figure 8. Competition occurs in the Binary mechanism for assemblies, but not in ROO filament mechanism. SgrAI bound to primary site (green, red) binding to a second copy of the same, or to SgrAI bound to secondary site (light blue or orange).

REFERENCES

- [1] Stern, A., and Sorek, R. (2010) The phage-host arms race: shaping the evolution of microbes, *Bioessays* 33, 43-51.
- [2] Nechaev, S., and Severinov, K. (2008) The elusive object of desire--interactions of bacteriophages and their hosts, *Current opinion in microbiology* 11, 186-193.
- [3] Vasu, K., and Nagaraja, V. (2013) Diverse functions of restriction-modification systems in addition to cellular defense, *Microbiol Mol Biol Rev* 77, 53-72.
- [4] Pingoud, A., Wilson, G. G., and Wende, W. (2014) Type II restriction endonucleases--a historical perspective and more, *Nucleic Acids Res* 42, 7489-7527.
- [5] Terns, M. P. (2018) CRISPR-Based Technologies: Impact of RNA-Targeting Systems, *Molecular cell* 72, 404-412.
- [6] Roberts, R. J., Vincze, T., Posfai, J., and Macelis, D. (2010) REBASE--a database for DNA restriction and modification: enzymes, genes and genomes, *Nucleic Acids Res* 38, D234-236.
- [7] Park, C. K., Stiteler, A. P., Shah, S., Ghare, M. I., Bitinaite, J., and Horton, N. C. (2010) Activation of DNA cleavage by oligomerization of DNA-bound SgrAI, *Biochemistry* 49, 8818-8830.
- [8] Lyumkis, D., Talley, H., Stewart, A., Shah, S., Park, C. K., Tama, F., Potter, C. S., Carragher, B., and Horton, N. C. (2013) Allosteric regulation of DNA cleavage and sequence-specificity through run-on oligomerization, *Structure* 21, 1848-1858.
- [9] Korennykh, A. V., Egea, P. F., Korostelev, A. A., Finer-Moore, J., Zhang, C., Shokat, K. M., Stroud, R. M., and Walter, P. (2009) The unfolded protein response signals through high-order assembly of Ire1, *Nature* 457, 687-693.
- [10] Ingerson-Mahar, M., Briegel, A., Werner, J. N., Jensen, G. J., and Gitai, Z. (2010) The metabolic enzyme CTP synthase forms cytoskeletal filaments, *Nat Cell Biol* 12, 739-746.
- [11] Noree, C., Sato, B. K., Broyer, R. M., and Wilhelm, J. E. (2010) Identification of novel filament-forming proteins in *Saccharomyces cerevisiae* and *Drosophila melanogaster*, *The Journal of cell biology* 190, 541-551.
- [12] Kim, C. W., Moon, Y. A., Park, S. W., Cheng, D., Kwon, H. J., and Horton, J. D. (2010) Induced polymerization of mammalian acetyl-CoA carboxylase by MIG12 provides a tertiary level of regulation of fatty acid synthesis, *Proc Natl Acad Sci U S A* 107, 9626-9631.
- [13] Narayanaswamy, R., Levy, M., Tsechansky, M., Stovall, G. M., O'Connell, J. D., Mirrielees, J., Ellington, A. D., and Marcotte, E. M. (2009) Widespread reorganization of metabolic enzymes into reversible assemblies upon nutrient starvation, *Proc Natl Acad Sci U S A* 106, 10147-10152.
- [14] Werner, J. N., Chen, E. Y., Guberman, J. M., Zippilli, A. R., Irgon, J. J., and Gitai, Z. (2009) Quantitative genome-scale analysis of protein localization in an asymmetric bacterium, *Proc Natl Acad Sci U S A* 106, 7858-7863.
- [15] Liu, J. L. (2010) Intracellular compartmentation of CTP synthase in *Drosophila*, *J Genet Genomics* 37, 281-296.
- [16] Shah, S., Sanchez, J., Stewart, A., Piperakis, M. M., Cosstick, R., Nichols, C., Park, C. K., Ma, X., Wysocki, V., Bitinaite, J., and Horton, N. C. (2015) Probing the Run-On Oligomer of Activated SgrAI Bound to DNA, *PLoS One* 10, e0124783.
- [17] Ma, X., Shah, S., Zhou, M., Park, C. K., Wysocki, V. H., and Horton, N. C. (2013) Structural Analysis of Activated SgrAI-DNA Oligomers Using Ion Mobility Mass Spectrometry, *Biochemistry* 52, 4373-4381.
- [18] Tautz, N., Kaluza, K., Frey, B., Jarsch, M., Schmitz, G. G., and Kessler, C. (1990) SgrAI, a novel class-II restriction endonuclease from *Streptomyces griseus* recognizing the octanucleotide sequence 5'-CR/CCGGYG-3' [corrected], *Nucleic Acids Res* 18, 3087.
- [19] Bilcock, D. T., Daniels, L. E., Bath, A. J., and Halford, S. E. (1999) Reactions of type II restriction endonucleases with 8-base pair recognition sites, *J Biol Chem* 274, 36379-36386.
- [20] Bitinaite, J., and Schildkraut, I. (2002) Self-generated DNA termini relax the specificity of SgrAI restriction endonuclease, *Proc Natl Acad Sci U S A* 99, 1164-1169.
- [21] Daniels, L. E., Wood, K. M., Scott, D. J., and Halford, S. E. (2003) Subunit assembly for DNA cleavage by restriction endonuclease SgrAI, *J Mol Biol* 327, 579-591.
- [22] Hingorani-Varma, K., and Bitinaite, J. (2003) Kinetic analysis of the coordinated interaction of SgrAI restriction endonuclease with different DNA targets, *J Biol Chem* 278, 40392-40399.
- [23] Wood, K. M., Daniels, L. E., and Halford, S. E. (2005) Long-range communications between DNA sites by the dimeric restriction endonuclease SgrAI, *J Mol Biol* 350, 240-253.
- [24] Park, C. K., Sanchez, J. L., Barahona, C., Basantes, L. E., Sanchez, J., Hernandez, C., and Horton, N. C. (2018) The run-on oligomer filament enzyme mechanism of SgrAI: Part 1. Assembly kinetics of the run-on oligomer

- filament, *J Biol Chem* 293, 14585-14598.
- [25] Park, C. K., Sanchez, J. L., Barahona, C., Basantes, L. E., Sanchez, J., Hernandez, C., and Horton, N. C. (2018) The run-on oligomer filament enzyme mechanism of SgrAI: Part 2. Kinetic modeling of the full DNA cleavage pathway, *J Biol Chem* 293, 14599-14615.
- [26] Blaisdell, B. E., Campbell, A. M., and Karlin, S. (1996) Similarities and dissimilarities of phage genomes, *Proc Natl Acad Sci U S A* 93, 5854-5859.
- [27] Karlin, S., Mrazek, J., and Campbell, A. M. (1997) Compositional biases of bacterial genomes and evolutionary implications, *Journal of bacteriology* 179, 3899-3913.
- [28] Gelfand, M. S., and Koonin, E. V. (1997) Avoidance of palindromic words in bacterial and archaeal genomes: a close connection with restriction enzymes, *Nucleic Acids Res* 25, 2430-2439.
- [29] Sharp, P. M. (1986) Molecular evolution of bacteriophages: evidence of selection against the recognition sites of host restriction enzymes, *Mol Biol Evol* 3, 75-83.
- [30] Moineau, S., Pandian, S., and Klaenhammer, T. R. (1993) Restriction/Modification systems and restriction endonucleases are more effective on lactococcal bacteriophages that have emerged recently in the dairy industry, *Appl Environ Microbiol* 59, 197-202.
- [31] Lee, S., Ward, T. J., Siletzky, R. M., and Kathariou, S. (2012) Two novel type II restriction-modification systems occupying genomically equivalent locations on the chromosomes of *Listeria monocytogenes* strains, *Appl Environ Microbiol* 78, 2623-2630.
- [32] Kasarjian, J. K., Iida, M., and Ryu, J. (2003) New restriction enzymes discovered from *Escherichia coli* clinical strains using a plasmid transformation method, *Nucleic Acids Res* 31, e22.
- [33] Michalowski, C. B., and Little, J. W. (2005) Positive autoregulation of *cI* is a dispensable feature of the phage lambda gene regulatory circuitry, *Journal of bacteriology* 187, 6430-6442.
- [34] Kong, H., Wenham, L. S., and Dalton, M. (2000) Method for cloning and producing the SgrAI restriction endonuclease, (States, U., Ed.), New England Biolabs, Inc., USA.
- [35] Lin, P. M., Lee, C. H., and Roberts, R. J. (1989) Cloning and characterization of the genes encoding the MspI restriction modification system, *Nucleic Acids Res* 17, 3001-3011.
- [36] Rippe, K., von Hippel, P. H., and Langowski, J. (1995) Action at a distance: DNA-looping and initiation of transcription, *Trends in biochemical sciences* 20, 500-506.
- [37] Handa, N., Ichige, A., Kusano, K., and Kobayashi, I. (2000) Cellular responses to postsegregational killing by restriction-modification genes, *Journal of bacteriology* 182, 2218-2229.
- [38] Pleska, M., Qian, L., Okura, R., Bergmiller, T., Wakamoto, Y., Kussell, E., and Guet, C. C. (2016) Bacterial Autoimmunity Due to a Restriction-Modification System, *Curr Biol* 26, 404-409.
- [39] Darmon, E., Eykelenboom, J. K., Lopez-Vernaza, M. A., White, M. A., and Leach, D. R. (2014) Repair on the go: *E. coli* maintains a high proliferation rate while repairing a chronic DNA double-strand break, *PLoS One* 9, e110784.
- [40] Sargentini, N. J., Diver, W. P., and Smith, K. C. (1983) The effect of growth conditions on inducible, *recA*-dependent resistance to X rays in *Escherichia coli*, *Radiat Res* 93, 364-380.
- [41] Sam, M. D., and Perona, J. J. (1999) Mn²⁺-dependent catalysis by restriction enzymes: pre-steady-state analysis of EcoRV endonuclease reveals burst kinetics and the origins of reduced activity, *J. Am. Chem. Soc.* 121, 1444-1447.
- [42] Lesser, D. R., Kurpiewski, M. R., and Jen-Jacobson, L. (1990) The energetic basis of specificity in the Eco RI endonuclease--DNA interaction, *Science* 250, 776-786.
- [43] Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., and Horvath, P. (2007) CRISPR provides acquired resistance against viruses in prokaryotes, *Science* 315, 1709-1712.
- [44] Jiang, F., and Doudna, J. A. (2017) CRISPR-Cas9 Structures and Mechanisms, *Annu Rev Biophys* 46, 505-529.
- [45] Gong, S., Yu, H. H., Johnson, K. A., and Taylor, D. W. (2018) DNA Unwinding Is the Primary Determinant of CRISPR-Cas9 Activity, *Cell Rep* 22, 359-371.
- [46] Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C., and Doudna, J. A. (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9, *Nature* 507, 62-67.
- [47] Tock, M. R., and Dryden, D. T. (2005) The biology of restriction and anti-restriction, *Current opinion in microbiology* 8, 466-472.
- [48] Rosenberg, A. H., Lade, B. N., Chui, D. S., Lin, S. W., Dunn, J. J., and Studier, F. W. (1987) Vectors for selective expression of cloned DNAs by T7 RNA polymerase, *Gene* 56, 125-135.
- [49] Mkrtumian, N. M., and Lomovskaya, N. D. (1972) A mutation affecting the ability of temperature sensitive actinophage PhiC31 of *Streptomyces coelicolor* to lyse and lysogenise, *Genetika* 8, 135-141.
- [50] Chater, K. F., and Carter, A. T. (1979) A new, wide host-range, temperate bacteriophage (R4) of *Streptomyces* and

- its interaction with some restriction-modification systems, *115*, 431-442.
- [51] Barry, R. M., Bitbol, A. F., Lorestani, A., Charles, E. J., Habrian, C. H., Hansen, J. M., Li, H. J., Baldwin, E. P., Wingreen, N. S., Kollman, J. M., and Gitai, Z. (2014) Large-scale filament formation inhibits the activity of CTP synthetase, *Elife* 3, e03638.
 - [52] Lynch, E. M., Hicks, D. R., Shepherd, M., Endrizzi, J. A., Maker, A., Hansen, J. M., Barry, R. M., Gitai, Z., Baldwin, E. P., and Kollman, J. M. (2017) Human CTP synthase filament structure reveals the active enzyme conformation, *Nat Struct Mol Biol* 24, 507-514.
 - [53] Petrovska, I., Nuske, E., Munder, M. C., Kulasegaran, G., Malinovska, L., Kroschwald, S., Richter, D., Fahmy, K., Gibson, K., Verbavatz, J. M., and Alberti, S. (2014) Filament formation by metabolic enzymes is a specific adaptation to an advanced state of cellular starvation, *Elife* 3, e02409.
 - [54] Webb, B. A., Dosey, A. M., Wittmann, T., Kollman, J. M., and Barber, D. L. (2017) The glycolytic enzyme phosphofructokinase-1 assembles into filaments, *The Journal of cell biology* 216, 2305-2313.
 - [55] Hunkeler, M., Hagmann, A., Stutfeld, E., Chami, M., Guri, Y., Stahlberg, H., and Maier, T. (2018) Structural basis for regulation of human acetyl-CoA carboxylase, *Nature* 558, 470-474.
 - [56] Dunten, P. W., Little, E. J., Gregory, M. T., Manohar, V. M., Dalton, M., Hough, D., Bitinaite, J., and Horton, N. C. (2008) The structure of SgrAI bound to DNA; recognition of an 8 base pair target, *Nucleic Acids Res* 36, 5405-5416.
 - [57] Aggarwal, A. K. (1990) Crystallization of DNA binding proteins with oligodeoxynucleotides., *Methods: A Companion to Methods in Enzymology* 1, 83-90.
 - [58] Fasman, G. D. (1975) *CRC Handbook of Biochemistry and Molecular Biology*, 3rd ed., CRC, Cleveland, OH.
 - [59] Johnson, K. A., Simpson, Z.B., Blom, T. (2009) Global Kinetic Explorer: A new computer program for dynamic simulation and fitting of kinetic data, *Analytical Biochemistry* 387, 20-29.
 - [60] Johnson, K. A. (2009) Fitting enzyme kinetic data with KinTek Global Kinetic Explorer, *Methods Enzymol* 467, 601-626.
 - [61] Park, C. K., Joshi, H. K., Agrawal, A., Ghare, M. I., Little, E. J., Dunten, P. W., Bitinaite, J., and Horton, N. C. (2010) Domain swapping in allosteric modulation of DNA specificity, *PLoS biology* 8, e1000554.

Supplemental Material

for

The need for speed: run-on oligomer filament formation provides maximum speed with maximum sequestration of activity

Claudia Barahona, L. Emilia Basantes, Kassidy J. Tompkins, Desirae M. Heitman, Barbara I. Chukwu,

Juan Sanchez, Jonathan L. Sanchez, Niloofar Ghadirian, Chad K. Park & N. C. Horton*

Department of Molecular and Cellular Biology, University of Arizona, Tucson, Arizona 85721

Table S1. Western data to quantitate expression levels

SgrAI Enzyme	Protein Level (%)²
WT	100%
T4D	38%
S6D	70%
I7E	57%
R24E	40%
N25E	30%
P27W	82%
P27G	74%
Q34D	80%
I51E	75%
S56E	68%
S56Q	29%
A57E	62%
A57Q	108%
I59E	59%
M62E	33%
R84E	33%
R127A	100%
R131A	38%
R134A	63%

Table S2. Calculation of Local Concentration of Sites on Phage DNA

Parameter	Number	Units
Estimated molecules SgrAI per cell	100	molecules
Molecules phage per cell (at the minimum)	1	molecules
size of cell in microns	1	microns
volume of a spherical cell	5×10^{-13}	cm^3
	5×10^{-16}	L
Concentration of SgrAI in the cell^a	3×10^{-7}	M
	300	nM
Calculation of the concentration of DNA and SgrAI recognition sites in the cell		
Concentration of phage DNA in cell	3×10^{-9}	M
	3	nM
Calculation of the local concentration of sites on the same DNA molecule^b		
size of phage DNA	50,000	bp
Distance between sites on phage (if 5 sites per phage) ^c	8,333	bp
Distance in Ångstrom	28,333	Å
Radius of gyration	2,173	Å
Volume occupied by 2 sites in phage using radius=Rg	4×10^{10}	Å^3
	4×10^{-14}	cm^3
	4×10^{-17}	L
Local concentration of 2 SgrAI/site on phage using Rg	8×10^{-8}	M
	80	nM

^a**Calculation of [SgrAI] in the cell:** first, the size of the cell is used to calculate volume, and with 100 molecules of SgrAI estimated, a concentration can be calculated as $C = 100 / (\text{Avogadro's number} \times \text{volume of the cell})$, and is calculated to be 300 nM. With 1 copy of DNA per cell, a similar calculation gives 3 nM.

^b**To calculate the local concentration of sites on the same phage DNA to each other:** with 3.4 Å/bp, the 8,333 bp distance between sites gives 28,333 Å between sites on linear DNA if fully extended (and in B form). The radius of gyration was calculated according to the equation $R_g = ((2 \times P \times N) / 6)^{1/2}$ in Å, P is persistence length (500 Å)^l, N is the linear distance (in Å) between sites. From this radius, a volume can be calculated ($V = 4/3\pi r^3$), and the concentration calculated as 2 molecules (the two DNA bound SgrAI) in a volume (calculated as 4×10^{-17} L), which is ~80 nM.

^cThe number of primary sites will vary with each phage genome, for example, λ phage contains 6 primary sites, though 3 are predicted based on the statistics in a 50,000 kb genome.

Table S3. Equilibria used in simulations of the ROO filament and Binary mechanisms

Reaction Step	Forward Rate Constant	Reverse Rate Constant
Equations for Run-on Oligomer Filament Mechanism		
<i>Binding of SgrAI binding to its recognition site in DNA (site) to create the SgrAI/DNA complex R</i>		
$\text{SgrAI} + \text{site} \rightleftharpoons \text{R}$	k_1	k_{-1}
<i>Self-association of a SgrAI/DNA complexes (i.e. R)</i>		
$\text{R} + \text{R} \rightleftharpoons \text{RR}$	k_2	k_{-2}
$\text{R} + \text{R} \rightleftharpoons \text{RR}$	k_2	k_{-2}
$\text{RR} + \text{R} \rightleftharpoons \text{RRR}$	k_2	k_{-2}
$\text{R} + \text{RR} \rightleftharpoons \text{RRR}$	k_2	k_{-2}
<i>DNA cleavage within the ROO filament (X denotes SgrAI bound to cleaved DNA)</i>		
$\text{RR} \rightleftharpoons \text{XR}$	k_3	k_{-3}
$\text{RR} \rightleftharpoons \text{RX}$	k_3	k_{-3}
$\text{XR} \rightleftharpoons \text{XX}$	k_3	k_{-3}
$\text{RX} \rightleftharpoons \text{XX}$	k_3	k_{-3}
$\text{RRR} \rightleftharpoons \text{XRR}$	k_3	k_{-3}
$\text{RRR} \rightleftharpoons \text{RXR}$	k_3	k_{-3}
$\text{RRR} \rightleftharpoons \text{RRX}$	k_3	k_{-3}
$\text{XRR} \rightleftharpoons \text{XXR}$	k_3	k_{-3}
$\text{XRR} \rightleftharpoons \text{XRX}$	k_3	k_{-3}
$\text{XXR} \rightleftharpoons \text{XXX}$	k_3	k_{-3}
$\text{XRX} \rightleftharpoons \text{XXX}$	k_3	k_{-3}
$\text{RXR} \rightleftharpoons \text{XXR}$	k_3	k_{-3}
$\text{RXR} \rightleftharpoons \text{RXX}$	k_3	k_{-3}
$\text{XXR} \rightleftharpoons \text{XXX}$	k_3	k_{-3}
$\text{RXX} \rightleftharpoons \text{XXX}$	k_3	k_{-3}
$\text{RRX} \rightleftharpoons \text{XRX}$	k_3	k_{-3}
$\text{RRX} \rightleftharpoons \text{RXX}$	k_3	k_{-3}
$\text{XRX} \rightleftharpoons \text{XXX}$	k_3	k_{-3}
$\text{RXX} \rightleftharpoons \text{XXX}$	k_3	k_{-3}
<i>Dissociation of run-on oligomer filaments that contain some SgrAI/DNA complexes with cleaved DNA (i.e. X) and some with uncleaved (i.e. R)</i>		
$\text{XR} \rightleftharpoons \text{R} + \text{X}$	k_{-2}	k_2
$\text{RX} \rightleftharpoons \text{R} + \text{X}$	k_{-2}	k_2
$\text{XX} \rightleftharpoons \text{X} + \text{X}$	k_{-2}	k_2
$\text{XRR} \rightleftharpoons \text{X} + \text{RR}$	k_{-2}	k_2
$\text{XRR} \rightleftharpoons \text{XR} + \text{R}$	k_{-2}	k_2
$\text{RXR} \rightleftharpoons \text{R} + \text{XR}$	k_{-2}	k_2
$\text{RXR} \rightleftharpoons \text{RX} + \text{R}$	k_{-2}	k_2
$\text{RRX} \rightleftharpoons \text{R} + \text{RX}$	k_{-2}	k_2
$\text{RRX} \rightleftharpoons \text{RR} + \text{X}$	k_{-2}	k_2
$\text{XXR} \rightleftharpoons \text{X} + \text{XR}$	k_{-2}	k_2
$\text{XXR} \rightleftharpoons \text{XX} + \text{R}$	k_{-2}	k_2

$\text{XRX} \rightleftharpoons \text{X} + \text{RX}$	k_{-2}	k_2
$\text{XRX} \rightleftharpoons \text{XR} + \text{X}$	k_{-2}	k_2
$\text{RXX} \rightleftharpoons \text{R} + \text{XX}$	k_{-2}	k_2
$\text{RXX} \rightleftharpoons \text{RX} + \text{X}$	k_{-2}	k_2
$\text{XXX} \rightleftharpoons \text{X} + \text{XX}$	k_{-2}	k_2
$\text{XXX} \rightleftharpoons \text{XX} + \text{X}$	k_{-2}	k_2
<i>Dissociation of SgrAI/DNA complexes which contain cleaved DNA (i.e. X) to SgrAI and cleaved DNA</i>		
$\text{X} \rightleftharpoons \text{SgrAI} + \text{cleaved DNA}$	k_4	k_{-4}
Equations for Binary Mechanism		
<i>Binding of Binary enzyme E to its recognition site in DNA (site) to create the E/DNA complex R</i>		
$\text{E} + \text{site} \rightleftharpoons \text{R}$	k_1	k_{-1}
<i>Self-association of a E/DNA complexes (i.e. R)</i>		
$\text{R} + \text{R} \rightleftharpoons \text{RR}$	k_2	k_{-2}
<i>DNA cleavage within the Binary complex (X denotes enzyme E bound to cleaved DNA)</i>		
$\text{RR} \rightleftharpoons \text{XR}$	k_3	k_{-3}
$\text{RR} \rightleftharpoons \text{RX}$	k_3	k_{-3}
$\text{XR} \rightleftharpoons \text{XX}$	k_3	k_{-3}
$\text{RX} \rightleftharpoons \text{XX}$	k_3	k_{-3}
<i>Dissociation of Binary complexes that contain some E/DNA complexes with cleaved DNA (i.e. X) and some with uncleaved DNA (i.e. R)</i>		
$\text{XR} \rightleftharpoons \text{X} + \text{R}$	k_{-2}	k_2
$\text{RX} \rightleftharpoons \text{R} + \text{X}$	k_{-2}	k_2
$\text{XX} \rightleftharpoons \text{X} + \text{X}$	k_{-2}	k_2
<i>Dissociation of E/DNA complexes which contain cleaved DNA (i.e. X) to enzyme E and cleaved DNA</i>		
$\text{X} \rightleftharpoons \text{E} + \text{cleaved DNA}$	k_4	k_{-4}

Table S4. Rate constants using in simulations

Reaction	Forward Rate Constant	Reverse Rate Constant
DNA recognition site (site) binding by SgrAI (ROO mechanism) or enzyme E (Binary mechanism)	$k_1 = 10^9 \text{ M}^{-1} \text{ s}^{-1}$	$k_{-1} = 0.06 \text{ s}^{-1}$
Association and self-association of enzyme-substrate (R) and enzyme-product complexes (X)	$k_2 = 1.3 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$	$k_{-2} = 0.03 \text{ s}^{-1}$
DNA cleavage by SgrAI (ROO mechanism) or E (Binary mechanism)	$k_3 = 0.8 \text{ s}^{-1}$	$k_3 = 0$ (considered irreversible)
Product release (release of cleaved DNA from SgrAI or E)	$k_4 = 0.4 \text{ s}^{-1}$	$k_{-4} = 3 \times 10^7 \text{ M}^{-1} \text{ s}^{-1}$ (Fig. 3A-B, D) $k_{-4} = 3 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ (Fig. 3C)

Table S5. Equations for simulating reactions with secondary site DNA

Reaction Step	Forward Rate Constant	Reverse Rate Constant
<i>SgrAI binding to primary site DNA or secondary site DNA into SgrAI/DNA complexes A and R, respectively</i>		
SgrAI+PRIMARY $\rightleftharpoons$ P	k_1	k_{-1}
SgrAI+SECONDARY $\rightleftharpoons$ S	k_2	k_{-2}
<i>Self-association of a SgrAI/PRIMARY complex (i.e. P) with another SgrAI/PRIMARY complex (i.e. P)</i>		
P + P $\rightleftharpoons$ PP	k_3	k_{-3}
P + P $\rightleftharpoons$ PP	k_3	k_{-3}
SP + P $\rightleftharpoons$ SPP	k_3	k_{-3}
P + PS $\rightleftharpoons$ PPS	k_3	k_{-3}
PP + P $\rightleftharpoons$ PPP	k_3	k_{-3}
P + PP $\rightleftharpoons$ PPP	k_3	k_{-3}
PPP + P $\rightleftharpoons$ PPPP	k_3	k_{-3}
P + PPP $\rightleftharpoons$ PPPP	k_3	k_{-3}
P + PPS $\rightleftharpoons$ PPPS	k_3	k_{-3}
P + PSP $\rightleftharpoons$ PPSP	k_3	k_{-3}
PSP + P $\rightleftharpoons$ PSPP	k_3	k_{-3}
SPP + P $\rightleftharpoons$ SPPP	k_3	k_{-3}
SP+PP $\rightleftharpoons$ SPPP	k_3	k_{-3}
SP+PS $\rightleftharpoons$ SPPS	k_3	k_{-3}
<i>Association of a SgrAI/PRIMARY complex (i.e. P) with a SgrAI/SECONDARY complex (i.e. S)</i>		
P + S $\rightleftharpoons$ PS	k_4	k_{-4}
S + P $\rightleftharpoons$ SP	k_4	k_{-4}
PS + P $\rightleftharpoons$ PSP	k_4	k_{-4}
P + SP $\rightleftharpoons$ PSP	k_4	k_{-4}
PP + S $\rightleftharpoons$ PPS	k_4	k_{-4}
S + PP $\rightleftharpoons$ SPP	k_4	k_{-4}
S + PS $\rightleftharpoons$ SPS	k_4	k_{-4}
SP + S $\rightleftharpoons$ SPS	k_4	k_{-4}
S + PPP $\rightleftharpoons$ SPPP	k_4	k_{-4}
S + PPS $\rightleftharpoons$ SPPS	k_4	k_{-4}
SPP + S $\rightleftharpoons$ SPPS	k_4	k_{-4}
S + PSP $\rightleftharpoons$ SPSP	k_4	k_{-4}
SPS + P $\rightleftharpoons$ SPSP	k_4	k_{-4}
P + SPS $\rightleftharpoons$ PSPS	k_4	k_{-4}
PSP + S $\rightleftharpoons$ PSPS	k_4	k_{-4}
PPP + S $\rightleftharpoons$ PPPS	k_4	k_{-4}
PPS + P $\rightleftharpoons$ PPSP	k_4	k_{-4}
P + SPP $\rightleftharpoons$ PSPP	k_4	k_{-4}
PP+SP $\rightleftharpoons$ PPSP	k_4	k_{-4}
PS+PP $\rightleftharpoons$ PSPP	k_4	k_{-4}

$\text{PS} + \text{PS} \rightleftharpoons \text{PSPS}$	k_4	k_{-4}
$\text{SP} + \text{SP} \rightleftharpoons \text{SPSP}$	k_4	k_{-4}
$\text{PP} + \text{PS} \rightleftharpoons \text{PPPS}$	k_4	k_{-4}
<i>Cleavage of secondary site DNA within the ROO filament (S becomes X)</i>		
$\text{PS} \rightleftharpoons \text{PX}$	k_5	k_{-5}
$\text{SP} \rightleftharpoons \text{XP}$	k_5	k_{-5}
$\text{PSP} \rightleftharpoons \text{PXP}$	k_5	k_{-5}
$\text{PPS} \rightleftharpoons \text{PPX}$	k_5	k_{-5}
$\text{SPP} \rightleftharpoons \text{XPP}$	k_5	k_{-5}
$\text{SPS} \rightleftharpoons \text{XPX}$	k_5	k_{-5}
$\text{PPPS} \rightleftharpoons \text{PPPX}$	k_5	k_{-5}
$\text{PPSP} \rightleftharpoons \text{PPXP}$	k_5	k_{-5}
$\text{PSPP} \rightleftharpoons \text{PXPP}$	k_5	k_{-5}
$\text{SPPP} \rightleftharpoons \text{XPPP}$	k_5	k_{-5}
$\text{PSPS} \rightleftharpoons \text{PSPX}$	k_5	k_{-5}
$\text{PSPS} \rightleftharpoons \text{PXPS}$	k_5	k_{-5}
$\text{PXPS} \rightleftharpoons \text{PXPX}$	k_5	k_{-5}
$\text{PSPX} \rightleftharpoons \text{PXPX}$	k_5	k_{-5}
$\text{SPPS} \rightleftharpoons \text{XPPX}$	k_5	k_{-5}
$\text{SPSP} \rightleftharpoons \text{XPPX}$	k_5	k_{-5}
<i>Dissociation of a SgrAI/PRIMARY complex (i.e. P) with another SgrAI/PRIMARY complex (i.e. P)</i>		
$\text{XPP} \rightleftharpoons \text{XP} + \text{P}$	k_{-3}	k_3
$\text{PPX} \rightleftharpoons \text{P} + \text{PX}$	k_{-3}	k_3
$\text{PPP} \rightleftharpoons \text{PP} + \text{P}$	k_{-3}	k_3
$\text{PPP} \rightleftharpoons \text{P} + \text{PP}$	k_{-3}	k_3
$\text{PPPX} \rightleftharpoons \text{P} + \text{PPX}$	k_{-3}	k_3
$\text{PPXP} \rightleftharpoons \text{P} + \text{PXP}$	k_{-3}	k_3
$\text{PXPP} \rightleftharpoons \text{PXP} + \text{P}$	k_{-3}	k_3
$\text{XPPP} \rightleftharpoons \text{XPP} + \text{P}$	k_{-3}	k_3
$\text{PPPP} \rightleftharpoons \text{PP} + \text{PP}$	k_{-3}	k_3
$\text{PPPX} \rightleftharpoons \text{PP} + \text{PX}$	k_{-3}	k_3
$\text{XPPP} \rightleftharpoons \text{XP} + \text{PP}$	k_{-3}	k_3
$\text{XPPX} \rightleftharpoons \text{XP} + \text{PX}$	k_{-3}	k_3
<i>Dissociation of SgrAI/SECONDARY complex with cleaved secondary site DNA (i.e. X) from a SgrAI/PRIMARY complex (i.e. P)</i>		
$\text{PX} \rightleftharpoons \text{P} + \text{X}$	k_{-4}	k_4
$\text{SP} \rightleftharpoons \text{X} + \text{P}$	k_{-4}	k_4
$\text{PXP} \rightleftharpoons \text{PX} + \text{P}$	k_{-4}	k_4
$\text{PXP} \rightleftharpoons \text{P} + \text{XP}$	k_{-4}	k_4
$\text{PPX} \rightleftharpoons \text{PP} + \text{X}$	k_{-4}	k_4
$\text{XPP} \rightleftharpoons \text{X} + \text{PP}$	k_{-4}	k_4
$\text{XPS} \rightleftharpoons \text{X} + \text{PS}$	k_{-4}	k_4
$\text{XPS} \rightleftharpoons \text{XP} + \text{S}$	k_{-4}	k_4

$SPX \rightleftharpoons S + PX$	k_{-4}	k_4
$SPX \rightleftharpoons SP + X$	k_{-4}	k_4
$XPX \rightleftharpoons X + PX$	k_{-4}	k_4
$XPX \rightleftharpoons XP + X$	k_{-4}	k_4
$PPPX \rightleftharpoons PPP + X$	k_{-4}	k_4
$PPXP \rightleftharpoons PPX + P$	k_{-4}	k_4
$PXPP \rightleftharpoons P + XPP$	k_{-4}	k_4
$XPPP \rightleftharpoons X + PPP$	k_{-4}	k_4
$PXPX \rightleftharpoons P + XPX$	k_{-4}	k_4
$PXPX \rightleftharpoons PXP + X$	k_{-4}	k_4
$XPPX \rightleftharpoons X + PPX$	k_{-4}	k_4
$XPPX \rightleftharpoons XPP + X$	k_{-4}	k_4
$XPXP \rightleftharpoons X + PXP$	k_{-4}	k_4
$XPXP \rightleftharpoons XPX + P$	k_{-4}	k_4
$PSPX \rightleftharpoons P + SPX$	k_{-4}	k_4
$PSPX \rightleftharpoons PSP + X$	k_{-4}	k_4
$PXPS \rightleftharpoons P + XPS$	k_{-4}	k_4
$PXPS \rightleftharpoons PXP + S$	k_{-4}	k_4
$SPPX \rightleftharpoons S + PPX$	k_{-4}	k_4
$SPPX \rightleftharpoons SPP + X$	k_{-4}	k_4
$XPPS \rightleftharpoons X + PPS$	k_{-4}	k_4
$XPPS \rightleftharpoons XPP + S$	k_{-4}	k_4
$SPXP \rightleftharpoons S + PXP$	k_{-4}	k_4
$SPXP \rightleftharpoons SPX + P$	k_{-4}	k_4
$XPSP \rightleftharpoons X + PSP$	k_{-4}	k_4
$XPSP \rightleftharpoons XPS + P$	k_{-4}	k_4
$PPXP \rightleftharpoons PP + XP$	k_{-4}	k_4
$PXPP \rightleftharpoons PX + PP$	k_{-4}	k_4
$PXPX \rightleftharpoons PX + PX$	k_{-4}	k_4
$XPXP \rightleftharpoons XP + XP$	k_{-4}	k_4
<i>Dissociation of cleaved SECONDARY site DNA from SgrAI</i>		
$X \rightleftharpoons \text{SgrAI} + \text{SECONDARY}^{\text{cleaved}}$	k_6	k_{-6}
Binary Mechanism		
<i>Binding of PRIMARY and SECONDARY site DNA to enzyme (E) to give complexes P and S, respectively</i>		
$E + \text{PRIMARY} \rightleftharpoons P$	k_1	k_{-1}
$E + \text{SECONDARY} \rightleftharpoons S$	k_2	k_{-2}
<i>Association of two enzyme-substrate complexes-enzyme bound to primary (P) may self-associate, or associate with enzyme bound to secondary (S) but enzyme bound to secondary does not self-associate</i>		
$P + P \rightleftharpoons PP$	k_3	k_{-3}
$P + S \rightleftharpoons PS$	k_4	k_{-4}
<i>Conversion of substrate to product (S complexes only)</i>		
$PS \rightleftharpoons PX$	k_5	k_{-5}
<i>Dissociation of Binary complex with cleaved SECONDARY site DNA</i>		

$\text{PX} \rightleftharpoons \text{P} + \text{X}$	k_{-4}	k_4
<i>Dissociation of cleaved SECONDARY site DNA from enzyme (E)</i>		
$\text{X} \rightleftharpoons \text{E} + \text{SECONDARY}^{\text{cleaved}}$	k_6	k_{-6}

Table S6. Rate Constants used in simulating secondary site cleavage (equations of Table S5)

Reaction	Forward Rate Constant	Reverse Rate Constant
Binding of SgrAI (ROO filament mechanism) or enzyme E (Binary mechanism) to PRIMARY site DNA	$k_1 = 10^8 \text{ M}^{-1}\text{s}^{-1}$	$k_{-1} = 0.006 \text{ s}^{-1}$
Binding of SgrAI (ROO mechanism) or enzyme E (Binary mechanism) to SECONDARY site DNA	$k_2 = 10^8 \text{ M}^{-1}\text{s}^{-1}$	$k_{-2} = 0.06 \text{ s}^{-1}$
Association of two SgrAI or two enzyme E complexes containing PRIMARY site DNA (P)	ROO: $k_3 = 1.3 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ Binary: $k_3 = 1.3 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ (Fig. 5A) $k_3 = 6.0 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ (Fig. 5B)	$k_{-3} = 0.03 \text{ s}^{-1}$
Association of enzyme-DNA complexes where one contains PRIMARY site DNA (P) and the other contains SECONDARY site DNA, cleaved or uncleaved (S or X , respectively)	ROO: $k_4 = 6.5 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ Binary: $k_4 = 6.5 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ (Fig. 5A) $k_4 = 3.0 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ (Fig. 5B)	$k_{-4} = 0.03 \text{ s}^{-1}$
Cleavage of SECONDARY site DNA within complex	$k_5 = 0.8 \text{ s}^{-1}$	$k_{-5} = 0$ (set to 0 to be irreversible)
Dissociation of cleaved SECONDARY site DNA from SgrAI (ROO) or enzyme E (Binary)	$k_6 = 0.4 \text{ s}^{-1}$	$k_{-6} = 0$ (set to 0 to be irreversible)

Table S7. Initial Concentrations used in the simulation of secondary site cleavage (see Tables S5-S6)

Species	Initial Concentration (nM)
SgrAI or E	4000 (excess to ensure complete DNA binding)
PRIMARY	3 (dotted lines) or 1000 (solid lines)
SECONDARY	3 (dotted lines) or 1000 (solid lines)

- [1] Rippe, K., von Hippel, P. H., and Langowski, J. (1995) Action at a distance: DNA-looping and initiation of transcription, *Trends in biochemical sciences* 20, 500-506.