Sequential eviction of crowded nucleoprotein complexes by the exonuclease RecBCD molecular motor

Tsuyoshi Terakawa, Sy Redding, Timothy D. Silverstein, and Eric C. Greene

In physiological settings, all nucleic acids motor proteins must travel along substrates that are crowded with other proteins. However, the physical basis for how motor proteins behave in these highly crowded environments remains unknown. Here, we use real-time single-molecule imaging to determine how the ATP-dependent translocase RecBCD travels along DNA occupied by tandem arrays of high-affinity DNA binding proteins. We show that RecBCD forces each protein into its nearest adjacent neighbor, causing rapid disruption of the protein-nucleic acid interaction. This mechanism is not the same way that RecBCD disrupts isolated nucleoprotein complexes on otherwise naked DNA. Instead, molecular crowding itself completely alters the mechanism by which RecBCD removes tightly bound protein obstacles from DNA.


Author contributions: T.T., S.R., and T.D.S. designed research; T.T. and S.R. performed research; T.D.S. contributed new reagents/analytic tools; T.T., S.R., and E.C.G. analyzed data; and T.T. and E.C.G. wrote the paper.

The authors declare no conflict of interest.

Significance

Chromosomes are crowded places, and any nucleic acid motor proteins that act on DNA must function within these crowded environments. How crowded environments affect motor protein behaviors remains largely unexplored. Here, we use single-molecule fluorescence microscopy to visualize the ATP-dependent motor protein RecBCD as it travels along crowded DNA molecules bearing long tandem arrays of DNA binding proteins. Our findings show that RecBCD can push through highly crowded protein arrays while evicting the proteins from DNA. This behavior on crowded DNA is distinct from a previously described mechanism by which RecBCD disrupts single isolated nucleoprotein complexes. These findings may provide insights into how other types of motor proteins travel along crowded nucleic acids.
provide insights into how molecular motors behave while traveling along nucleic acids in crowded physiological settings.

Results

Models for Protein Eviction in Crowded Environments. We have previously shown that RecBCD can push a single EcoRI<sup>E111Q</sup> dimer or single RNAP complexes for very long distances along DNA before they are eventually displaced into solution (Fig. 1A) (19). The question that naturally arises from this result is what happens when RecBCD encounters DNA-bound proteins in highly crowded environments (Fig. 1B). Does RecBCD stall when traveling on crowded DNA? Is RecBCD able to continue processive translocation in crowded conditions, and if so, how? To help answer these questions, we first constructed a general kinetic model that, in principle, might be applied to any motor protein that must travel on DNA in crowded condition. Then, we experimentally tested the hypotheses predicted from these models for the specific case of RecBCD.

As an initial step toward understanding how motor proteins might behave in crowded settings, we first considered three generalized scenarios describing potential outcomes of RecBCD collisions with protein arrays (Fig. 1C). In the (i) accumulation model, RecBCD pushes each protein into its nearest neighbor without dislodging any of the proteins from the DNA, resulting in greater resistance as proteins continue to accumulate in front of the translocase. For the (ii) sequential model, RecBCD actively evicts each protein as it is encountered. In the (iii) spontaneous model, the proteins spontaneously dissociate according to their intrinsic dissociation rate constants, and RecBCD must wait for these dissociation events before moving forward. We include the spontaneous model as a formal possibility, although we note that this model is unlikely to be correct for RecBCD, because previous experiments have shown that RecBCD can quickly push EcoRI<sup>E111Q</sup>, RNAP, and lac repressor off of their respective cognate binding sites with no evidence of either slowing or pausing (19). Importantly, the accumulation and sequential models are not mutually exclusive but rather, may be considered to reflect a continuum of models. Indeed, RecBCD readily pushes isolated proteins for extended distances along DNA (19), suggesting that it might also push proteins into one another on crowded DNA. These observations suggest that some variation of the accumulation model could apply for RecBCD acting in crowded environments. To account for this possibility, we also considered an alternative variation of the sequential model, in which small numbers of proteins can accumulate in front of RecBCD before the accrued resistance leads to sequential dissociation (Fig. S1A). For clarity, we refer to this variation of the sequential model as “sequential eviction after collision model” hereafter.

Monte Carlo Simulations of Protein Eviction by DNA Translocases. We next performed kinetic Monte Carlo (KMC) simulations as a means to predict potential experimental outcomes for each of the different models. For these KMC simulations, we modeled the behavior of RecBCD on DNA substrates bearing 5–50x tandem arrays of the high-affinity DNA binding protein EcoRI<sup>E111Q</sup>. This protein is a catalytically inactive version of the EcoRI restriction endonuclease that binds tightly to DNA, but it is unable to cleave its cognate target site (21). We chose EcoRI<sup>E111Q</sup> for our studies, because it is one of the highest-affinity DNA binding proteins known to exist, with site-specific and nonspecific dissociation constants (K<sub>d</sub>) of ~2.5 μM and ~4.8 pM (21), respectively. EcoRI<sup>E111Q</sup> is also a highly potent block to both the transcription (22–24) and DNA replication machineries (25, 26). In addition, WT EcoRI can withstand up to ~20–40 pN applied force (27). EcoRI<sup>E111Q</sup> binds to its cognate target ~3,000-fold more tightly than WT EcoRI. Therefore, we infer that EcoRI<sup>E111Q</sup> can resist at least as much force as the WT protein.

Within each KMC simulation, the DNA-bound proteins must either slide or dissociate on collision with RecBCD. The accumulation model is realized by prohibiting dissociation of the obstacle proteins, which are instead always pushed by RecBCD. The sequential model requires RecBCD to provoke protein dissociation before moving forward. In the spontaneous model, RecBCD must wait until proteins dissociate according to their intrinsic dissociation rate constants. Each model predicts that RecBCD will slow or stall on encountering the array; these events should be revealed as an experimentally observable pause coinciding with the location of protein array (Fig. S2A). Importantly, the accumulation, sequential, and spontaneous dissociation models all yield distinct predictions for the relationship between apparent pause duration (Δt) and array size (Fig. 1D and Fig. S2F). The accumulation model predicts an exponential increase in pause duration with increasingly large protein arrays, the sequential model predicts a linear relationship between pause duration and array size, and the spontaneous model predicts a logarithmic variation in pause duration for the different array sizes. The spontaneous dissociation model

![Fig. 1. Models for translocase behavior in crowded environments.](https://example.com/fig1)

---

**A** Schematic illustration of RecBCD pushing a single EcoRI<sup>E111Q</sup> based on our previously published findings (19). **B** Illustration of the experimental question being addressed in this work, namely what takes place when RecBCD encounters a crowded array of EcoRI<sup>E111Q</sup> on DNA. **C** Schematic depictions of three generalized models (accumulation, sequential, and spontaneous) for RecBCD movement through protein arrays. Details of each model are presented in the text. **D** Results from KMC simulations for each different model showing the predicted relationships between RecBCD pause durations (Δt) and protein array size. The durations (Δt) represent the time in which RecBCD remains within the region of the array. The array size represents how many EcoRI binding sites are within the array. **E** Predicted postarray RecBCD velocities for each model. The postarray velocity represents the predicted velocity of RecBCD after passage through the array.

*Terakawa et al.*
also predicts that RecBCD will traverse the array orders of magnitude more slowly than the other models, because RecBCD must wait for each protein to dissociate (Fig. 1D). In addition, the accumulation model predicts that RecBCD will experience a persistent reduction in velocity after traversing the array because of the accumulated resistance of the proteins that it must push as it continues to move along the DNA (Fig. 1E). Finally, if small numbers of EcoRI_E111Q can build up in front of RecBCD before dissociation, then the pause duration will scale approximately linearly with array size as up to few proteins accumulate in front of RecBCD, similar to expectations for a purely sequential model (Fig. S1B). However, pause duration begins to scale exponentially if larger numbers of proteins can accumulate in front of RecBCD before dissociation, yielding results that would be more similar to the pure accumulation model (Fig. S1B).

**Visualizing Removal of Unlabeled EcoRI_E111Q from Crowded DNA.** We next sought to establish an experimental approach for directly testing the predictions for each of the different models. To accomplish this aim, we engineered λ-phage DNA molecules bearing tandem arrays of 5, 10, 30, or 50 EcoRI binding sites (Fig. S3). We then used single-molecule DNA curtain assays to experimentally determine whether quantum dot (Qdot)-tagged RecBCD could traverse these protein arrays (Fig. 2A) and if so, determine which of the models presented above might most closely reflect the experimental data. In all of the experiments, RecBCD molecules are injected and incubated with the DNA curtain for 20 min in the absence of ATP. After washing out the free RecBCD molecules in solution, we initiated the translocation and DNA digestion reaction by supplementing the buffer with 1 mM ATP. Thus, any RecBCD that dissociates from the DNA ends is not replenished.

In the absence of EcoRI_E111Q, Qdot-tagged RecBCD displayed high processivity and monotonic translocation with two peaks in the velocity distribution, corresponding to 745 ± 37 (13.6%) and 1,368 ± 18 bp s⁻¹ (86.4%) (Fig. S4 A and B). This result is in good agreement with reports for the properties of unlabeled RecBCD on naked DNA (18, 28, 29).

We next asked whether the presence of EcoRI_E111Q affected the translocation behavior of RecBCD. Remarkably, RecBCD was still able to processively translocate along the DNA in the presence of saturating EcoRI_E111Q concentrations (Fig. 2B and Fig. S5). However, as predicted by the KMC simulations, RecBCD exhibited a noticeable pause on encountering each of the protein arrays (Fig. 2B and C), and the average pause duration (Δt) scaled linearly with array size (Fig. 2D). Control experiments with unlabeled RecBCD and unlabeled EcoRI_E111Q were in close agreement with results from Qdot-tagged RecBCD, arguing against the possibility that the Qdot might alter the outcomes of the collisions (Fig. S6). Together, these results are in closest agreement with expectations for the sequential eviction model or a variation of the sequential model involving the accumulation of a small number of EcoRI_E111Q dimers in front of RecBCD (Fig. 1D and E and Fig. S1B). We do not yet know precisely how many EcoRI_E111Q dimers might accumulate in front of RecBCD before they start dissociating. However, DNA curtain experiments have revealed that RecBCD can concurrently push at least two EcoRI_E111Q dimers (Fig. S1C), whereas the tandem arrays of five EcoRI_E111Q dimers cause a noticeable pause in RecBCD translocation (Fig. 2). Therefore, we conclude that between two and five EcoRI_E111Q dimers might accumulate in front of RecBCD before the combined resistance leads to their sequential eviction.

RecBCD slowed to an average apparent velocity of 27.3 ± 3.4 bp s⁻¹ while traversing the EcoRI_E111Q arrays, but the velocity drastically increased once beyond the array. We measured the RecBCD velocities from pre- and postarray trajectories in the presence (Fig. 2E) and absence of EcoRI_E111Q (Fig. S4C). Comparison of these two scatter plots reveals no statistically significant difference in the RecBCD velocities that could be attributed to its passage through an occupied EcoRI_E111Q array. Instead, any variation in pre- and postcollision velocities seems to arise from the limited precision of the velocity measurement or normal variation in RecBCD translocation rates as opposed a persistent, systematic reduction in translocation velocity. It should be noted that there can be up to two EcoRI_E111Q molecules pushed in front of RecBCD in the prearray trajectories because of the native EcoRI binding sites located upstream from the cloned arrays (Fig. 2). Thus, the velocity does not change even if a small number of EcoRI_E111Q molecules have accumulated in front of

![Fig. 2](https://www.pnas.org/cgi/doi/10.1073/pnas.1701368114)
RecBCD in the postarray trajectories. Together, these results are also in closest agreement with expectations for the sequential eviction after collision model.

**Sequential Disruption of Fluorescent EcoRI\textsubscript{E111Q} Arrays.** The KMC simulations and experimental results presented above together with the observation that RecBCD readily pushes isolated molecules of EcoRI\textsubscript{E111Q} (19) all suggest that RecBCD sequentially removes proteins from crowded DNA and that it may do so by pushing the proteins into one another. This model makes at least two important predictions that can be experimentally tested using DNA curtain assays. The sequential eviction after collision models predicts that passage of RecBCD through a fluorescently tagged protein array should coincide with a linear decrease in the array signal as the proteins are evicted, but the last protein(s) within the array should be pushed for long distances along the naked DNA, because they will encounter no resistance from more distal obstacles (19) (Fig. 3A). To test this first prediction, we labeled a λ-DNA substrate bearing a 50× EcoRI binding site array with Qdot-tagged EcoRI\textsubscript{E111Q} and then asked whether and how unlabeled RecBCD passed through these arrays. As anticipated, eviction of the fluorescent proteins from the 50× array was initially observed as a linear decrease in the overall fluorescence signal intensity as unlabeled RecBCD moved slowly through the array (Fig. 3B). Also, as predicted, on reaching the end of the array, RecBCD resumed its normal velocity and pushed the remaining protein(s) toward the end of the DNA molecule (Fig. 3B).

The sequential eviction after collision models also predicts that, if RecBCD pushes a single proximal EcoRI\textsubscript{E111Q} into a more distal EcoRI\textsubscript{E111Q} array, then the resulting collision should coincide with the preferential eviction of the proximal protein as it is driven into the larger array (Fig. 4A). To test this prediction, we performed two-color single-molecule experiments, in which separate aliquots of EcoRI\textsubscript{E111Q} and then asked whether and how unlabeled RecBCD passed through these arrays. As anticipated, eviction of the fluorescent proteins from the 50× array was initially observed as a linear decrease in the overall fluorescence signal intensity as unlabeled RecBCD moved slowly through the array (Fig. 3B). Also, as predicted, on reaching the end of the array, RecBCD resumed its normal velocity and pushed the remaining protein(s) toward the end of the DNA molecule (Fig. 3B).

To determine whether and how RecBCD might disrupt RNAP complexes, including core RNAP, RNAP holoenzyme, stalled elongation complexes, and actively transcribing RNApol complexes, we performed two-color single-molecule experiments, in which separate aliquots of EcoRI\textsubscript{E111Q} and then asked whether and how unlabeled RecBCD passed through these arrays. As anticipated, eviction of the fluorescent proteins from the 50× array was initially observed as a linear decrease in the overall fluorescence signal intensity as unlabeled RecBCD moved slowly through the array (Fig. 3B). Also, as predicted, on reaching the end of the array, RecBCD resumed its normal velocity and pushed the remaining protein(s) toward the end of the DNA molecule (Fig. 3B).

**RNAP Rapidly Dissociates When Forced into EcoRI\textsubscript{E111Q}** Our results show that RecBCD sequentially removes EcoRI\textsubscript{E111Q} from DNA by pushing the proteins into one another. We next sought to determine whether similar principles apply to other nucleoprotein roadblocks. RNAP is perhaps the most abundant and formidable nucleoprotein roadblock that will be encountered by RecBCD in living cells. A single *E. coli* cell contains ∼2,000 molecules of RNAP, and under typical growth conditions, ≥65% of these polymerases are bound to the bacterial chromosome (30). RNAP is also of particular interest, because it is a high-affinity DNA binding protein (K\textsubscript{d} ≈ 10−100 pM) and a powerful translocase capable of moving under an applied load of up to ∼14–25 pN (31). RNAPs can survive encounters with replication forks and stall fork progression in head-on collisions (32–36). Indeed, the highly transcribed rRNA genes are a potent blockade to DNA replication (37–39). We have previously shown that RecBCD can disrupt individual *E. coli* RNAP complexes, including core RNAP, RNAP holoenzyme, stalled elongation complexes, and actively transcribing polymerases in either head-to-head or head-to-tail orientation (19). RecBCD pushes isolated RNAP for long distances over naked DNA, and dissociation takes place as RecBCD forces the polymerase to step from one nonspecific binding site to the next (19).

To determine whether and how RecBCD might disrupt RNAP on crowded DNA, we first sought to establish what happens when RecBCD pushes RNAP into tandem 5× arrays of EcoRI\textsubscript{E111Q}...
Sequential Eviction of Tandem RNA Polymerases. Interestingly, when EcoRI\textsuperscript{E111Q} was absent from the reactions, a second population of RNAP dissociated from the DNA when forced into one another by RecBCD. For these experiments, Qdot-tagged RNAP holoenzyme was bound to the native phage promoters (19, 40), and unlabeled RecBCD was loaded onto the free ends of the DNA molecules. RecBCD translocation was then initiated by the addition of ATP. Remarkably, RNAP dissociates from the DNA almost immediately on being pushed by RecBCD into the EcoRI array position was entirely dependent on the presence of EcoRI\textsuperscript{E111Q}, and when EcoRI\textsuperscript{E111Q} was absent, many of the polymerases were instead pushed to the ends of the naked DNA molecules (Fig. 5A). We conclude that RNAP rapidly dissociates from the DNA when pushed into other high-affinity DNA binding proteins by RecBCD, in good agreement with the sequential eviction after collision model.

Sequential Eviction of Tandem RNA Polymerases. Interestingly, when EcoRI\textsuperscript{E111Q} was absent from the reactions, a second population of RNAP dissociated at a position coinciding with the location of the \(\lambda_{P_{BL}}\) promoter (Fig. 5A) (40). One possible explanation for this observation is that the Qdot-tagged RNAP might be encountering unlabeled RNAP bound to the \(\lambda_{P_{BL}}\) promoter as it is pushed along the DNA by RecBCD, and the resulting collisions with the unlabeled proteins may have provoked rapid eviction of the Qdot-tagged protein from the DNA. Therefore, we next sought to directly examine what happens when RecBCD pushed two RNAPs into one another. To accomplish this aim, we relied on the eight native promoters present in the \(\lambda\)-phage genome, which allows multiple RNAP complexes to be loaded onto the same DNA molecule (40). We first performed DNA curtain experiments using unlabeled RecBCD and promoter-bound RNAP open complexes, which were labeled with a single-color Qdot (Qdot 705) (Fig. 5B). Remarkably, analysis of the cumulative fluorescence intensity of the DNA-bound polymerases suggested that only one of the two polymerases remained on the DNA when RecBCD pushed them into one another (Fig. 5B). We conclude that, although RecBCD readily pushes single RNAP complexes along DNA, it does not appear to push two RNAPs at the same time. Instead, as predicted by the sequential eviction after collision model, one of the two polymerases quickly falls off the DNA when they collide with one another.

The results described above provide evidence that the sequential eviction after collision model may apply to RecBCD encounters with RNAPs in crowded settings. Importantly, the sequential eviction after collision model specifically predicts that the proximal polymerase should be preferentially evicted when pushed by RecBCD into the distal polymerase. We next sought to verify this prediction by determining which of two polymerases dissociated from the DNA when forced into one another by RecBCD. We, therefore, conducted two-color DNA curtain assays, in which separate aliquots of RNAP were labeled with either green (Qdot 605) or red (Qdot 705) Qdots. The differentially labeled polymerases were then mixed together and bound to the native phage promoters, unlabeled RecBCD was bound to the free DNA ends, and translocation was initiated by the injection of ATP. These experiments revealed that, for \(\sim 96\%\) of observed collisions (\(n = 50/52\)) involving RecBCD and two tandem molecules of RNAP, the proximal polymerase almost immediately dissociated from the DNA on being pushed into the distal polymerase (Fig. 5C). We conclude that RecBCD can rapidly and sequentially evict RNAP from crowded DNA and that it does so specifically by forcing the polymerases into one another.
Discussion

Our results show that RecBCD rapidly and sequentially removes crowded nucleoprotein complexes from crowded DNA by pushing them into one another. These findings reveal that molecular crowding itself can have a crucial and unanticipated influence on how molecular motor proteins clear nucleic acids of bound obstacles.

Molecular Crowding Alters the Mechanism of Protein Eviction by RecBCD.

The mechanism by which RecBCD removes high-affinity nucleoprotein complexes from DNA is dependent on molecular crowding (Fig. 6). When RecBCD encounters isolated molecules of either EcoRIE111Q or RNAP on otherwise naked DNA, which tightly binds on specific binding sites before the encounter, it can push these proteins over average distances of 13,000 ± 9,100 and 10,460 ± 7,690 bp, respectively (19). The proteins eventually dissociate as they are forced to step between successive nonspecific binding sites, and the probability \( P \) of dissociation is directly proportional to the number of steps \( n \) that EcoRIE111Q or RNAP is forced to take while being pushed along the DNA (Fig. 6A) (19). However, the probability of dissociating during any given step is low, and as a consequence, RecBCD can push isolated proteins for extended distances on naked DNA (Fig. 6A). This outcome is in marked contrast to what takes place in crowded environments, where EcoRIE111Q and RNAP both dissociate almost immediately when pushed into more distal proteins or protein arrays (Fig. 6B).

Fig. 6. A generalized model for protein eviction in crowded environments. (A) RecBCD can push isolated proteins for extended distances along the otherwise naked DNA, and the probability \( P \) of dissociation scales proportionally with the number of steps \( n \) that the protein is forced to take as it is pushed by RecBCD (19). (B) In contrast, when RecBCD encounters proteins in crowded environments, it pushes each protein into its nearest adjacent neighbor, resulting in rapid dissociation of the proximal protein.

Fig. 5. RecBCD sequentially disrupts tandem RNAPs. (A) Experimental schematic (Upper) and resulting dissociation position distribution data (Lower) for Qdot-tagged RNAP pushed into 5x EcoRI arrays by RecBCD. (B) Representative one-color kymograph showing two Qdot-tagged RNAP complexes (promoter-bound holoenzyme) being pushed into one another by unlabeled RecBCD and the corresponding graph showing the cumulative fluorescence intensity of the Qdot-tagged proteins. (C) Representative two-color kymograph showing two Qdot-tagged RNAP complexes (promoter-bound holoenzyme) being pushed into one another by unlabeled RecBCD; similar outcomes were observed in most \( n = 50/52 \) two-color RNAP collisions, and in the remaining cases, RecBCD pushed both proteins \( n = 2/52 \). p and d represent proximal and distal RNAP, respectively. In B and C, reactions were initiated by the addition of 1 mM ATP into the RecBCD buffer (40 mM Tris-HCl, pH 7.5, 2 mM MgCl\(_2\), 0.2 mg mL\(^{-1}\) Pluronic), and gaps in the RNAP trajectories result from Qdot blinking.
This much more rapid eviction takes place on crowded DNA specifically, because RecBCD pushes the proteins into one another. The striking differences between the outcomes of isolated collisions involving single-nucleoprotein complexes and the outcomes of collisions on crowded DNA highlight the dramatic and unexpected impact that molecular crowding has on the mechanism by which RecBCD interacts with nucleoprotein obstacles that it encounters while traveling along DNA. This sequential mechanism for protein eviction may likely reflect what takes place in vivo, where long tracts of naked DNA are unlikely to exist (41, 42). The general applicability of the sequential evaporation after collision model for RecBCD is supported by the similarities between the experimental findings for EcoRI\textsubscript{E111Q} and \textit{E. coli} RNAP. The significance of this mechanism is that it could serve to prevent RecBCD from being stalled by tightly bound obstacles that might otherwise accumulate in front of the helicase as they translocate along DNA.

**Sequential Removal of RNAP from DNA by RecBCD.** Interestingly, our work suggests that a small number of EcoRI\textsubscript{E111Q} complexes can accumulate in front of RecBCD before they start dissociating from the DNA. In contrast to EcoRI\textsubscript{E111Q}, it does not appear as though multiple molecules of RNAp can accumulate in front of RecBCD. Instead, the proximal polymerase dissociates from the DNA almost immediately on being pushed into a distal protein. These observations indicate that RecBCD removes RNAP from crowded DNA much more easily than it removes EcoRI\textsubscript{E111Q}. This difference may reflect the fact that RNAP is a naturally occurring obstacle that will likely be encountered whenever RecBCD acts on DNA in a cellular environment. We speculate that coevolution of RNAP and RecBCD may have tuned to relative binding strengths of these two nucleoprotein complexes to ensure that RNAP cannot impede the movement of RecBCD. Alternatively, the collision of RecBCD may induce some more specific structural transition within the RNAp complex, allowing it to be rapidly evicted from DNA by the action of RecBCD. Future work will be necessary to further define the molecular mechanisms that contribute to RNAP dissociation from DNA by RecBCD.

**Mechanisms of Protein Dissociation.** Through-DNA allosteric communication can influence the dissociation of stationary proteins that are bound in close spatial proximity to one another (43). Previous studies have shown that protein pairs, including glucocorticoid receptor and BamHI or lac repressor together with either EcoRV or T7 RNAp, exhibited up to fivefold changes in dissociation rates when the corresponding partner was bound to a nearby DNA site (43). These experimental findings have been attributed to through-DNA allosteric communication based on the long-range oscillatory changes in DNA major and minor groove widths observed in molecular dynamics simulations (43–45). Similarly, we find that a static RecBCD complex positioned immediately adjacent to an EcoRI site causes approximately a twofold reduction in DNA cleavage by EcoRI (Fig. S7). This reduction in cleavage is comparable in magnitude to the effects previously ascribed to the through-DNA allosteric model. However, this twofold effect is substantially less than the rapid dissociation rates observed as RecBCD traverses an array of EcoRI\textsubscript{E111Q}, suggesting that through-DNA allostericity may not be a predominant factor affecting protein displacement by RecBCD. In addition, our experimental data show that, under crowded conditions, RecBCD causes rapid dissociation of the most proximal protein, but only when it is pushed into a more distal obstacle, indicating that crowded environments enhance protein dissociation by RecBCD relative to isolated collisions. The substantial increase in rate enhancement when RecBCD is moving through a protein array together with the dependence of proximal protein dissociation on the presence of a more distal obstacle suggest that the through-DNA allosteric model cannot account for RecBCD-mediated protein displacement under crowded conditions. Future work will be essential to further evaluate more detailed molecular insights into precisely how protein eviction from crowded environments and determine whether other types of DNA translocases may act similarly to RecBCD.

**Nucleoprotein Obstacles and Genome Integrity.** Nucleoprotein complexes are the primary source of replication fork stalling (46), and their presence represents a major challenge to genome integrity (1, 37, 38, 47–49). Indeed, prokaryotic and eukaryotic replisomes both require accessory helicases to clear tightly bound proteins from DNA (25, 50–54). For instance, the \textit{E. coli} replisome requires the accessory helicases Rep and UvrD to prevent replication fork collapse on encountering RNAp and other types of high-affinity nucleoprotein complexes (25, 46). Similarly, \textit{Bacillus subtilis} requires the accessory helicase PcrA to facilitate replication fork progression through highly transcribed genes (52), and the helicases Pif1 and Rrm3 are required for efficient fork progression through difficult to replicate regions in \textit{Saccharomyces cerevisiae} (55). In addition, Rep, UvrD, and Pif1 can all push isolated ssDNA binding proteins along ssDNA (56) using a mechanism that is similar in many respects to what takes place during RecBCD collisions with isolated proteins (19). However, the physical basis by which these accessory helicases assist the replisome remains unknown. One possibility is that they may strip proteins from crowded DNA through the mechanism similar to that used by RecBCD. Although many helicases including them have much lower translocation rates and processivity than RecBCD while acting on their own, the replisome itself is a highly processive molecular machine, and these helicases may act differently while working together with the replisome. Future work will be essential for further establishing how these replication accessory helicases (in both the presence and the absence of the replisome) and other types of motor proteins disrupt tightly bound nucleoprotein complexes on crowded nucleic acids.

**Conclusion**

We have presented a model describing the ability of RecBCD to sequentially clear crowded DNA of nucleoprotein complexes. The key feature of the sequential evaporation after collision model is that RecBCD provokes rapid disruption of crowded nucleoprotein complexes by pushing these obstacles into one another. This model suggests that molecular crowding itself alters the mechanism by which RecBCD removes proteins from DNA. The sequential evaporation after collision model suggests that ATP-dependent nucleic acid motor proteins can respond differently to encounters with isolated nucleoprotein complexes compared with encounters involving multiple nucleoprotein complexes. The general principles revealed from our studies with RecBCD may also apply to the behavior of other types of motor proteins as they travel along crowded nucleic acids.

**Materials and Methods**

**KMC Simulations.** The KMC simulations included three kinds of molecules: the DNA, the translocase, and the roadblock protein. The DNA was a 10 track with 49 bpk in length. Within the simulations, the translocase moves along DNA at a fixed rate of 1,500 bp s\textsuperscript{-1} (k\textsubscript{trans}). The roadblock proteins bind to and dissociate from specific binding sites with a rate constant k\textsuperscript{rns} and dissociate from non-specific binding sites with a rate constant k\textsuperscript{rns}. When the translocase encounters a roadblock protein, they make a complex (tr). The tr complex moves along DNA with a rate constant k\textsuperscript{tr} = k\textsuperscript{trans} exp(-F, A, X/k T), where F, A, and X is the direction of movement, k is the Boltzmann constant, and T is temperature (310 K). Likewise, when the tr encounters the next roadblock protein, they make a larger complex (tr, n). In this case, F is dependent on the number of roadblock proteins bound to specific vs. non-specific sites, such that F = F\textsuperscript{trans} + n F\textsuperscript{rns}, where n and F\textsuperscript{rns} are the numbers of proteins bound to non-specific and specific sites, respectively. F\textsuperscript{tr} exp = F\textsuperscript{trans} exp - F\textsuperscript{rns}. The numbers of proteins bound to non-specific and specific sites, respectively. When the translocase encounters a
roadblock protein, it can also induce the dissociation of the protein with the dissociation rate constant written as $k_{sp} = k_{sp}^{\Delta x} + k_{sp}^{\Delta y}$. Then, we calculate the time when a particular event takes place ($t = 1$) as the probability of each particular event ($P_{emission} = k_{emission} / k_{total}$). The current time and state of the system are updated accordingly, and the procedure is repeated until the translocase reaches the opposite DNA end. The three different scenarios illustrated in Fig. 1A were realized by altering the parameter values for $F_{i}^{\Delta x}, F_{i}^{\Delta y}, k_{in}, k_{out}$. For the dissociation model, we set $k_{sp}^{\Delta x} = \infty$ and $F_{i}^{\Delta y} = \infty$ to prohibit additional movement of the translocase complex on encountering a protein until the protein spontaneously dissociates from the translocase-DNA complex. For the dissociation plus storage buffer dissociation model, the dissociation can be analytically derived, and $<t_t> = \frac{k_{sp}}{k_{total}} \log \langle N_{k}\rangle$, where $k_{sp}$ is the rate constant for dissociation, and $N_{k}$ and $N_{0}$ are the numbers of proteins bound to the array initially and at time $t$, respectively, in accord with the simulation result.

For the alternative variation of the accumulation model, we varied the parameter describing resistance of a roadblock protein on specific binding sites ($F_{i}^{\Delta x} = 0.05, 0.1, 0.2, 0.3, 0.5, 0.6, 0.8$ for each model parameter at different values $r_{sp}^{\Delta x} = 0.01$, $F_{i}^{\Delta y} = 0.4, k_{sp}^{\Delta x} = 4.5 \times 10^{-6}$, $k_{sp}^{\Delta y} = 4.5 \times 10^{-3}$). This setting allows us to change the relative rates of EcoRI$^{\Delta y}$ dissociation and sliding, resulting in the accumulation of varying numbers of EcoRI$^{\Delta y}$ in front of RecBCD before dissociation (Fig. S1A). Each different value of $F_{i}^{\Delta x}$ leads to the accumulation of a different number ($N$) of EcoRI$^{\Delta x}$ proteins in front of RecBCD (Fig. S1B).

Proteins and DNA. Biotinylated RecBCD was purified from E. coli IM109(DE3) cells co-transformed with plasmids for expression of Rec and RecBD with an avidity tag on the C terminus of RecD. The cells were grown in 2YT broth (16 g/L tryptone, 10 g/L yeast extract, and 5 g/L NaCl) to OD$_{600}$ = 0.6 in the presence of 25 μg mL$^{-1}$ Chloramphenicol and 100 μg mL$^{-1}$ Carbenicillin. The media were then supplemented with 0.2 mM biotin, protein expression was induced by addition of 0.5 mM isopropyl-$\beta$-d-thiogalactopyranoside (IPTG), and cells were grown for an additional 16 h at 37 °C. Cells were then harvested by centrifugation, resuspended in lysis buffer (50 mM Tris HCl, pH 7.5, 0.5 mM PMSF, 10% sucrose), and lysed by sonication. The lysate was clarified by centrifugation, and the supernatant was loaded onto the chitin column (New England Biolabs). The column was washed with buffer A (40 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM EDTA, 5% glycerol), and RecBCD was eluted with buffer B (40 mM Tris-HCl, pH 7.5, 1 mM EDTA, 5% glycerol, 1 M NaCl, 50 mM DTT). After the purification, RecBCD was dialyzed into storage buffer containing 40 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM EDTA, and 35% glycerol, frozen on liquid nitrogen, and stored at −80 °C.

Catalytically inactive EcoRI$^{\Delta y}$ bearing a flag epitope tag was expressed described (40). All of the restriction enzymes and ligases were purchased from New England Biolabs.

E. coli RNAS. Single-molecule experiments were conducted using a custom-built total internal reflection microscope and DNA curtains as previously described (19, 58). In brief, a biotinylated λ-DNA molecules were anchored to a supported lipid bilayer on the surface of a microfluidic sample chamber through a biotin-streptavidin linkage as previously described (19, 58). The DNA molecules were then purified by microfluidic chamber surface. The resulting ligation products were packaged into phage particles using Gigapack III packaging extracts (catalog no. 200201; Agilent Technologies), and the resulting phage DNA was purified as previously described (40). All of the restriction enzymes and ligases were purchased from New England Biolabs.
**Experiments with RNAP.** For assays using Qdot RecBCD and unlabeled EcoRIE111Q (Fig. 2), the λ-DNA substrates (150 pM) bearing different size EcoRI arrays were preincubated with EcoRIE111Q before injecting the DNA into the sample chambers. The amount of EcoRIE111Q used in the assays was sufficient to saturate each array, and the amount of protein used in each experiment scaled with the size of the array as follows: 5x array, 5 nM EcoRIE111Q; 10x array, 10 nM EcoRIE111Q; 30x array, 30 nM EcoRIE111Q; and 50x array, 50 nM EcoRIE111Q. All preincubations were conducted in 100 μL of EcoRI buffer (40 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10 mM MgCl2, 0.2 mg mL−1 Pluronic). The binding reactions were then diluted into 1 mL of RecBCD buffer, injected into the flow cell, and incubated for an additional 30 min. Free DNA and excess EcoRIE111Q were then flushed from the sample chambers using RecBCD buffer. Qdot-tagged RecBCD was then injected into the sample chambers, and translocation was initiated by the addition of 1 mM ATP as described above.

**Single-Color EcoRIE111Q Experiments.** For DNA curtain assays using unlabeled RecBCD and Qdot-tagged EcoRIE111Q (Figs. 3 and 4), the unlabeled EcoRIE111Q was preincubated with the λ-DNA and then injected into the flow cell as described above. The DNA-bound EcoRIE111Q molecules were then labeled with anti-FLAG antibody-conjugated Qdots by injecting the Qdots (10 nM) followed by a 5-min incubation before flushing the unbound Qdots from the sample chamber. The antibody-conjugated Qdots were prepared using the SiteClick Qdot Antibody Labeling Kit (catalog no. S10469 for Qdot 605 and catalog No. S10454 for Qdot 705; Thermo Fisher Scientific).

**Two-Color EcoRIE111Q Experiments.** λ-DNA (150 pM) with a 5x EcoRI binding site array was preincubated with 5 nM EcoRIE111Q in 100 μL of EcoRI buffer (40 mM Tris-HCl, pH 7.5, 10 mM MgCl2, 0.2 mg mL−1 Pluronic). The reaction was then diluted to a total volume of 1 mL in RecBCD buffer, injected into a flow cell, and incubated for an additional 30 min. Free DNA and EcoRIE111Q were then flushed out of the sample chamber. The DNA-bound EcoRIE111Q was then labeled by injecting anti-FLAG antibody-conjugated Qdots (5 nM) with different emission maxima (605 or 705 nm; colored green and magenta in all kymographs) followed by a 5-min incubation. RecBCD was then loaded onto the DNA, translocated by the addition of 1 mM ATP, and data were collected as described above. Note that the binding distributions of “magenta” and “green” EcoRIE111Q were random and that reaction trajectories, in which the two colors were appropriately segregated between the 5x arrays and 1x native binding sites, were selected by visual inspection (Fig. 4).

**Experiments with RNAP.** WT λ-DNA (150 pM) was diluted to a total volume of 1 mL in RNAP buffer (40 mM Tris-HCl, pH 7.5, 100 mM KCl, 2 mM MgCl2, 0.2 mg mL−1 Pluronic). The reaction was then diluted to a total volume of 1 mL in RecBCD buffer, injected into a flow cell, and incubated for an additional 30 min. Free DNA and EcoRIE111Q were then flushed out of the sample chamber. The DNA-bound EcoRIE111Q was then labeled by injecting anti-FLAG antibody-conjugated Qdots (5 nM) with different emission maxima (605 or 705 nm; colored green and magenta in all kymographs) followed by a 5-min incubation. RecBCD was then loaded onto the DNA, translocated by the addition of 1 mM ATP, and data were collected as described above. Note that the binding distributions of “magenta” and “green” EcoRIE111Q were random and that reaction trajectories, in which the two colors were appropriately segregated between the 5x arrays and 1x native binding sites, were selected by visual inspection (Fig. 4).

**Through-DNA Allosteric Inhibition by RecBCD.** Previous studies have shown that through-DNA allosteric communication can reduce the affinity of site-specific DNA binding proteins by up to a factor of five (43). Therefore, we devised a bulk biochemical measurement to determine whether (static) RecBCD might inhibit the binding of EcoRI to a nearby site, which would be consistent with the through-DNA allosteric model (Fig. 57). For this aim, we designed hairpin oligonucleotides labeled with Alexa488 for detection. The oligonucleotides contained a free blunt end for binding RecBCD and a single EcoRI cleavage located either immediately adjacent to RecBCD (proximal) or separated by a 10-bp spacer from the RecBCD binding site (distal). The oligonucleotides were incubated at 95 °C for 10 min and then cooled at room temperature for 1 h. The annealed oligonucleotides were then preincubated for 10 min with RecBCD in 18 μL of buffer containing 40 mM Tris-HCl, pH 7.5, and 2 mM MgCl2. Cleavage reactions were initiated by the addition of 2.0 μL of 100 mM WT EcoRI. Aliquots (2 μL) were then transferred to new tubes containing termination buffer (90% formamide, 0.5% EDTA, 0.1% orange G) at the indicated time intervals. Samples were resolved on 10% denaturing urea polyacrylamide gels at 200 V for 50 min. The Alexa488 signal was detected with a Typhoon FLA 7000 (GE Healthcare), and images were analyzed using ImageJ (https://imagej.nih.gov/ij/). Data were analyzed by plotting the normalized signal intensity of the 30-min time point for each sample and subtracting the Signal of RecBCD and that the effect is greater for the proximal vs. distal EcoRI substrate, and the relative magnitude of these effects are consistent with previous studies of through-DNA allosteric inhibition (43).

**ACKNOWLEDGMENTS.** We thank members of the laboratory of E.C.G. for discussion and assistance throughout this work, and we thank Shige Takada, Jayil Lee, Daniel Dudzевич, Fabian Erdel, Corentin Moers, Kyle Kaniecki, Chu Jian Ma, Johannes Stigler, Justin Steinfeld, Luisina De Tullio, and Mayu S. Terakawa for comments on the manuscript. The data described in this manuscript are archived in the laboratory of E.C.G. This work was supported by a Japan Society for the Promotion of Science fellowship (to T.T.), NIH Fellowship F32GM101819 (to T.D.S.), and NIH Grant R35GM118026 (to E.C.G.).
Fig. S1. KMC simulations of the alternative sequential dissociation model. (A) Schematic illustration of a hybrid model invoking components of both accumulation and sequential eviction. In this example, three proteins accumulate \((N = 3)\) before dissociation. Additional details of this model are presented in Materials and Methods. (B) Pause duration distributions \((\Delta t_i)\) for simulations involving differential numbers of accumulated proteins \((N)\) for each different array size as indicated. We fit the data for \(n = 5\) with the linear and exponential equations. The sums of squared residuals for best fit curves are 0.2 (seconds\(^2\)) for the linear equation and 221.9 (seconds\(^2\)) for the exponential equation, suggesting that the linear equation fits better than the exponential equation. (C) Experimental schematic and kymograph showing that RecBCD (unlabeled) can push two molecules of differentially labeled EcoRI\(E_{111}Q\).
Fig. S2. KMC simulations to predict translocase movement through protein arrays. (A) Representative kymographs derived from the KMC simulations of the accumulation, sequential, and spontaneous models. The white lines represent the movement of the translocase along a DNA molecule, pausing on encountering the protein array. (B) Distributions of pause durations (Δt). Colored lines represent the distribution derived from simulations using the 5–50× protein arrays as indicated. Each distribution was calculated from a total of 1,000 simulation trajectories.
**Fig. S3.** Engineered 5x, 10x, 30x, and 50x arrays of EcoRI binding sites on \( \lambda \)-DNA. (A) A schematic of the \( \lambda \)-DNA and a schematic of the expected fragment sizes for the different DNA array lengths when digested with SalI and EagI. (B) A 1% agarose gel showing that the length of engineered sites is as expected. In this assay, 5-μL aliquots of the engineered \( \lambda \)-DNA were incubated with 1 μL of SalI-HF (R3138S; New England BioLabs) and EagI-HF (R3505S; New England BioLabs) in 38 μL water supplemented with 5 μL of 5× CutSmart buffer (B7204S; New England BioLabs) for 2 h. The reaction was mixed with 10 μL of 6× loading dye (B7024S; New England BioLabs) and run on the gel for 1 h at 100 V; a 1-kb DNA ladder (N3232S; New England BioLabs) was also run. The magenta arrow highlights the 3.3-kb native DNA fragment [SalI (33.4 kb) to EagI (36.7 kb)] present before cloning in the array fragments. The green arrows highlight the array-containing fragments for each of the different substrates. MW, molecular weight.
Fig. S4. Translocation properties of Qdot-tagged RecBCD in the absence of EcoRI\textsuperscript{E111Q}. (A) An example of a kymograph showing the movement of Qdot-tagged RecBCD along an unlabeled DNA molecule in the absence of EcoRI\textsuperscript{E111Q}. RecBCD was prebound to the DNA, and translocation was initiated by the injection of ATP (white arrow). (B) Velocity distribution for Qdot-tagged RecBCD ($n = 269$). (C) Scatter plot of RecBCD velocities before and after encountering the EcoRI binding site array for control experiments conducted in the absence of any EcoRI\textsuperscript{E111Q} ($n = 269$). The $R$ represents the Pearson correlation coefficient.
Fig. S5. The saturation of the λ-DNA substrate with EcoRI<sup>E111Q</sup>. (A) An example of the bulk biochemical assay used to establish conditions for saturating the EcoRI binding site arrays with EcoRI<sup>E111Q</sup>. In this example, the λ-DNA substrate bearing a 5× binding site array was preincubated with varying concentrations of EcoRI<sup>E111Q</sup> for 30 min at room temperature. The reactions were then challenged by the addition of WT EcoRI for 15 min at 37 °C. DNA products were then resolved on a 1% agarose gel. (B) Position distribution histogram from a DNA curtain assay with Qdot-tagged EcoRI<sup>E111Q</sup> bound to a λ-DNA harboring a 5× array. The location of the array is indicated along with the four peaks corresponding to the native EcoRI binding sites. The height of the 5× peak is similar to the 1× peak heights, because the data are not corrected for signal intensity.
Fig. S6. The collision of unlabeled RecBCD with an unlabeled 5× EcoRI array. (A) Representative kymographs from experiments where unlabeled RecBCD translocates on YoYo-1–stained DNA containing the 5× EcoRI binding site array (position indicated by red arrows) in the absence (Upper) and presence (Lower) of unlabeled EcoRI(E111Q). The yellow arrows in Lower indicate the beginning and ending of the RecBCD pause when EcoRI(E111Q) is bound to the DNA. Note that the protein-bound EcoRI arrays routinely fluoresce more brightly with YoYo-1 compared with the naked flanking DNA. (B) Pause time distribution for unlabeled in the presence of unlabeled EcoRI(E111Q). The dataset is fitted by a Gaussian distribution to derive the average pause duration (9.5 ± 0.4 s).
Fig. S7. Digestion of DNA oligonucleotides by WT EcoRI in the presence and absence of RecBCD. (A) Schematic of the oligonucleotide substrates containing proximal and distal EcoRI binding sites. The EcoRI binding sites are shown in blue, and the RecBCD binding site is shown in red. (B) An example denaturing gel showing a time course of EcoRI digestion in the presence of (static) RecBCD. (C) The normalized signal intensity of digested fraction at 30 min is plotted for each substrate in the presence and absence of RecBCD. Error bars represent SD calculated from three independent experiments. In the cartoon depiction, the DNA is black, EcoRI binding sites are red, EcoRI is magenta, and the RecBCD molecules are green.