Identification of oligonucleotide sequences that direct the movement of the *Escherichia coli* FtsK translocase

Oren Levy‡*, Jerod L. Ptacin†*, Paul J. Pease‡*, Jeff Gore§, Michael B. Eisen‡*, Carlos Bustamante‡§*, and Nicholas R. Cozzarelli‡**

*Biophysics Graduate Group, ‡Department of Molecular and Cell Biology, ¶Genomics Division, Ernest Orlando Lawrence Berkeley National Laboratory, †Department of Physics, and §Howard Hughes Medical Institute, University of California, Berkeley, CA 94720

Contributed by Nicholas R. Cozzarelli, October 13, 2005

FtsK from *Escherichia coli* is a fast and sequence-directed DNA translocase with roles in chromosome dimer resolution, segregation, and decatenation. From the movement of single FtsK particles on defined DNA substrates and an analysis of skewed DNA sequences in bacteria, we identify GNGNAGGG, its complement, or both as a sequence motif that controls translocation directionality. GNGNAGGG is skewed so that it is predominantly on the leading strand of chromosomal replication. Translocation across this octamer from the 3′ side of the G-rich strand causes FtsK to pause, turn around, and translocate in the opposite direction. Only 39 ± 4% of the encounters between FtsK and the octamer result in a turnaround, congruent with our optimum turnaround probability prediction of 30%. The probability that the observed skew of GNGNAGGG within 1 megabase of *dif* occurred by chance in *E. coli* is 1.7 × 10⁻⁵², and similarly dramatic skews are found in the five other bacterial genomes we examined. The fact that FtsK acts only in the terminus region and the octamer skew extends from origin to terminus implies that this skew is also important in other basic cellular processes that are common among bacteria. Finally, we show that the FtsK translocase is a powerful motor that is able to displace a triplex-forming oligo from a DNA substrate.

The DNA sequence of an organism encodes information on multiple levels. Beyond directly encoding proteins, specific DNA sequences recruit and direct the proteins required for processes such as DNA replication, transcription, and repair. DNA sequence-dependent proteins typically locate their binding site by diffusion where they then carry out a specific action (1). In contrast, the *Escherichia coli* DNA translocase FtsK uses an ATP-powered search mechanism to reach its target (*dif*), the site of chromosome dimer resolution (2–5). During replication, crossing over by homologous recombination can lead to the formation of a chromosome dimer that must be resolved into monomers before segregation (2). FtsK is required to activate the site-specific recombinases XerC and XerD to resolve these dimers at the *dif* site, which is near the terminus of replication (2, 6). FtsK is anchored in the membrane at the division septum and actively translocates the *dif* sites into spatial proximity, rather than waiting for *dif* to diffuse into the septal region (7–9). The overall efficiency of reaching *dif* by translocation is greatly affected by two factors. First, FtsK must be able to bypass or displace bound proteins or other potential roadblocks it may encounter on the DNA. Second, FtsK must maintain its overall direction of translocation toward *dif*.

Perhaps the simplest mechanism for ensuring that FtsK always translocates toward *dif* would involve FtsK loading at sequences that point FtsK in the proper direction. However, single-molecule observations of the FtsK motor domain (FtsK<sub>50C</sub>) translocating on DNA showed that FtsK could spontaneously reverse direction while maintaining an overall directionality (5, 10). Our prior study implies that recurring DNA sequences regulate the direction of FtsK translocation (5). These results substantiate genetic evidence that the misorientation of polar sequences near *dif*, such as lambda insertions, can disrupt proper segregation of the terminus region (11, 12). The polarity of a sequence can be quantified as a skew, which is defined as the percentage of occurrences of the sequence on the leading strand. These observations led to a model for FtsK directionality in which highly strand-biased sequences that switch skew at *dif* direct FtsK translocation (12). One such skewed motif, RGNAGGGGS (RAG), was initially proposed to direct FtsK based on its high skew in the *dif* region (12); however, no direct evidence linked the RAG sequence to FtsK activity.

Skewed sequence motifs such as RAG are widespread in prokaryotic genomes, yet their functions remain largely undetermined (13, 14). Here, we demonstrate that control of FtsK directionality is a unique function of a family of skewed octamers in the *E. coli* genome. We directly visualized FtsK reversal sites on the chromosomal region surrounding *dif*, and in conjunction with an informatics analysis, we identified the GNGNAGGG motif, its complement, or both as the best candidate to specify FtsK directionality. Using a single-molecule assay, we find that a GNGNAGGG sequence efficiently reverses FtsK translocation. FtsK reverses direction only when encountering the sequence from the 3′ end of the G-rich strand. FtsK paused at the turnaround sites and reversed at a frequency of 39 ± 4%. This frequency agrees well with a theoretical analysis of the optimal probability for FtsK sequence recognition based on typical in vivo parameters. Finally, we show that FtsK is a powerful motor that is capable of stripping a triplex-forming oligo from DNA in a sequence-dependent fashion, thus demonstrating that FtsK can indeed remove roadblocks as it translocates toward *dif*.

Materials and Methods

**FtsK<sub>50C</sub> Purification.** All experiments used the C-terminal motor domain of FtsK fused to 50 residues from the N terminus (FtsK<sub>50C</sub>) (4). FtsK<sub>50C</sub> was purified as described in ref. 1 and also in Supporting Text (which is published as supporting information on the PNAS web site).

**Single-Molecule DNA Substrate Preparation.** Single-molecule DNA substrates, or tethers, were created by ligating biotin- or digoxigenin-modified PCR products to the ends of DNA sequences of interest (5). DNA tethers from the *E. coli* *dif* region (c-tether) were prepared as described in ref. 5. For test sequence analysis, ~48-kb DNA tethers were constructed by the ligation of a 41-kb fragment of lambda DNA (3) to a 6-kb fragment containing the sequence to be tested. The parent vector was a variant of pTYB1 (NEB, Beverly, MA) in which the sequence GTGCAGGG was removed by PCR inclusion.

Conflict of interest statement: No conflicts declared.

Freely available online through the PNAS open access option.

Abbreviations: FRS, FtsK recognition sequence; Mb, megabase.

1 O.L., J.L.P., and P.J.P. contributed equally to this work.
2 To whom correspondence may be addressed: Department of Molecular and Cell Biology, University of California, 16 Barker Hall, Berkeley, CA 94720. E-mail: carlos@alice.berkeley.edu or ncozzare@berkeley.edu.

© 2005 by The National Academy of Sciences of the USA

PNAS | December 6, 2005 | vol. 102 | no. 49
amplification, digestion with XmaI, and ligation with T4 DNA ligase. The test sequence was cloned into the XbaI site of the resultant plasmid (pJP1) as a duplex oligonucleotide containing 5-mer repeats of the sequence GGGCAGGGG (anti), CCCCT-GGCC (iso), or GGAGGGCGGG (scramble), spaced with four intervening A, C, or T nucleotides and verified by sequencing (see Supporting Text for exact sequence). To generate clones with a single test sequence, we amplified pJP1 by using tailed PCR primers containing an XbaI site and the sequence GGGCAGGGG or CCCCTG GCC.

To produce fragments for tether assembly, plasmids described above were digested with BamHI, dephosphorylated, and purified. Products were then digested with KpnI and gel-purified. The 41-kb lambda, 6-kb plasmid, and two molecular handle fragments were then ligated together by using T4 DNA ligase.

Single-Molecule Experimental Procedures and Data Analysis. Translocations were carried out with 10 μg/ml FtsK50C in 50 mM Tris, pH 7.5/5 mM MgCl2/50 mM NaCl/1 mM DTT/3 mM ATP/100 μg/ml BSA. The DNA tethers were held at ~40 pN of tension between a 2.7-μm anti-digoxigenin antibody-coated polystyrene bead in an optical trap and a 3.2-μm streptavidin-coated polystyrene bead immobilized on a glass pipette. Translocation of FtsK50C particles was recorded on digital video at 30 Hz, and image analysis was used to ascertain position as a function of time. Noise was smoothed by using an eight-point moving average. A translocation event is defined as the period of activity in which FtsK50C moves for at least 0.5 μm in one direction and at a minimum speed of 0.25 μm/sec. A turnaround occurs when a translocation event terminates, and FtsK50C translocates in the opposite direction. A pause is the time between two consecutive translocation events. Pause duration analysis was done on turnaround zones II or I. All errors are given as standard errors.

Informatics. We compiled a list of all octamers (words) occurring on the leading strand of E. coli within 15 kb to the left of dif. For each octamer, we identified and separately grouped all octamers found in the same 15-kb region with 1, 2, or 3 mismatches (mismatch words). We generated a list of candidates by comparing each octamer to its corresponding mismatch words list. We first grouped all of the mismatch words that had mismatches in the same positions. Next, we determined all of the combinations of the degenerate bases (N, R, S, W, etc.) for these mismatch positions. The result was a list of octamer words with 0, 1, 2, or 3 degenerate bases. The above process was repeated on the opposite strand of the 15-kb region to the right of dif. The list was filtered by requiring that on the 15-kb regions to the left and right of dif, (i) the octamers have a skew >50, (ii) the octamers occur within 1,000 bp of the turnaround peak center (Fig. 1A, zones I, II, and IV), and (iii) the observed sequence distribution was not present by chance alone (P value < 0.01). A nucleic acids substitution-scoring matrix (see Supporting Text) was used to cluster similar octamers into distinct motifs. Two octamers were considered to be part of the same group if any alignment between them gave a substitution score that was above the arbitrarily selected threshold of 22. The motifs were determined by manual observations for any group that had more than five members.

Triplex Substrate Preparation and Displacement Assays. DNA substrates for triplex displacement were prepared by first cloning a DNA triplex-promoting site (15) into the EcoRV and PstI sites of a derivative of pBS-SK(+) ). Duplex oligos containing 5-mer repeats of the anti, iso, or scramble sequences (described in Single-Molecule DNA Substrate Preparation) or derivatives were cloned into the SpeI site, and the final products were verified by sequencing. DNA triplexes were assembled as described in ref. 15 and Supporting Text. Triplex displacement reactions were carried out at 25°C in 50 mM Tris, pH 7.5/5 mM MgCl2/3 mM ATP/0.1 mg/ml BSA/1 mM DTT/0–152 nM FtsK50C/5 nM triplex substrate.

Results

Identification of FtsK Turnaround Regions. Previously, we used optical tweezers to track the movement of FtsK50C particles on single DNA molecules. We demonstrated that only a single motor is active in the FtsK particle (ref. 5 and Supporting Text). We showed that FtsK50C moves unidirectionally overall, punctuated by occasional reversals, both on lambda DNA and on the 30-kb chromosomal regions immediately to the left or right of dif (5). On a chromosomal DNA molecule with dif at the center (c-tether), FtsK oscillated around dif. We visually tracked FtsK50C particles as they oscillated about dif on the c-tether to identify possible turnaround zones (Movie 1, which is published as supporting information on the PNAS web site). To improve the resolution of the assay, we kept the DNA tether at a high force of ~40 pN to reduce the size and frequency of FtsK50C-induced loops (5, 10). The location of FtsK50C as a function of DNA sequence could be measured within ±750 bp, with resolution limited by diffraction, the size of the FtsK50C particle, and FtsK50C-induced looping of the DNA.

A histogram of FtsK50C turnaround locations shows that 88% of turnarounds (486/555 over three tethers) occur when FtsK50C is translocating away from dif (Fig. 1A). This behavior is expected, given FtsK’s biological role. There are several zones where FtsK50C frequently reverses direction (Fig. 1A). Two of these zones are located at 2,500 ± 750 bp (I) and 500 ± 750 bp (II) to the left of dif. To the right of dif, the most prominent turnaround zone is 2,400 ± 750 bp from dif (IV). A fourth peak (III) is located between dif and zone IV. However, its location and magnitude varies among the three tethers analyzed (Fig. 6, which is published as supporting information on the PNAS web site), and therefore we excluded it from further analysis. We also found that there is a sequence dependence of the dwell time of FtsK50C as FtsK50C pauses at the turnaround zones (compare Fig. 1A and B). The mean pause time at zone I was ~0.9 sec.
is 1.0 ± 0.1 sec, and the pause duration appears to be random (Fig. 1C). We conclude that there are specific DNA sequences in the observed turnaround zones, the FtsK recognition sequences (FRSs), which cause FtsK to reverse direction when encountering the sequence in one orientation. A significant skew in the orientation of the FRS toward dif can efficiently bias FtsK translocation in the direction of dif. We discuss in the next section the implication of these results for the identity of the FRS.

Bioinformatics Identifies Candidate FRSs. To develop a short list of candidate FRSs, we used the experimentally identified turnaround zones to constrain an informatics analysis of skewed sequences near dif. Salzberg et al. (14) identified 150 octamers in the E. coli chromosome that are significantly skewed and whose skew switches strand at the origin and terminus of replication. We reasoned that the direct observation of FtsK turnaround zones would help us focus our informatics analysis by concentrating on sequences present in these zones; however, we did not limit our analysis to candidates that were identified by Salzberg’s study. An FRS should display a significant skew on the E. coli chromosome and switch strands at dif. The algorithm was initially executed against the 15-kb regions to the left and right of dif; because these regions are covered by the dif-centered DNA tether on which we observed directional movement of FtsK (5). We modeled the FRS as an octamer to make the search computationally more tractable (see Supporting Text). Degerenercy in up to three positions was allowed; we presumed that the FRS is a motif rather than a specific sequence.

We initially identified eight motifs that satisfied these criteria (Table 1, which is published as supporting information on the PNAS web site). We anticipated that FtsK activity extends beyond the immediate dif proximal region, as was confirmed by Corre and Louarn (16). Therefore, we extended the analysis to longer regions. Only the GNGNAGGG candidate maintained a strong skew and statistical significance within 100 kb and 1 megabase (Mb) from dif (Fig. 24; see also Fig. 7A, which is published as supporting information on the PNAS web site). This motif has a skew of 86 (P = 4.4 × 10–11) at 100 kb and a skew of 79 (P = 1.7 × 10–53) at 1 Mb away from dif. RAG was not identified as a candidate motif in this analysis, because there are no RAGs within turnaround zone II. The other two major turnaround zones (I and IV) contain both RAG and GNGNAGGG sequences (Fig. 1A). One of the major reasons that RAG was proposed as an FRS is its very high skew of 1.0 (76 vs. 70, respectively). Finally, we verified that the skew is 83 at 15 kb and 77 at 100 kb and 1 Mb away from dif. This probability is extremely small (P = 1.7 × 10–53) at 1 Mb away from dif. It is remarkable that statistically significant and similar skew levels are found in organisms that diverged so long ago. In summary, the observation of FtsK turnaround zones near dif, in combination with the informatics analysis, strongly implicates the GNGNAGGG motif as an FRS.

Direct Observation of FtsK Directionality Switch by an FRS. We used single DNA molecules in an optical tweezers to directly test the effect of our candidate sequences on FtsK translocation. We constructed DNA tethers consisting of three segments: a 41-kb fragment of lambda phage DNA, the test sequence, and a 6-kb plasmid spacer (Fig. 3A). The lambda DNA promotes FtsK50C translocation toward the sequence to be tested (5). Each test sequence was engineered in two different orientations at the junction between the lambda and plasmid DNA. In the anti orientation, FtsK approaches the test sequence from the 5’ end of the G-rich strand, we expected the FRS to reverse the FtsK50C particles until an FRS in the lambda DNA portion caused FtsK to reverse direction again. Therefore, the test sequence should act as a gatekeeper between the lambda portion of the tether and the plasmid. In the iso orientation, the test sequence is approached from the 5’ end of the G-rich strand, and we expected that the FtsK50C particles would not recognize the sequence and translocate past it.

We tested the candidate FRS (GGGCAGGGG), which is both a GNGNAGGG and a RAG sequence, as a single copy and as a multimer (5-mer) to enhance the turnaround effect. Individual DNA molecules were scored by the percentage of FtsK50C particles that reversed direction at the test sequence. The results show a dramatic and orientation-dependent effect of the test sequence on FtsK translocation. On both the iso 1-mer and 5-mer substrates,
FtsK<sub>50C</sub> reversed direction in only ∼10% of cases (Fig. 3B), suggesting that this reversal frequency is a background signal inherent in our assay. In contrast, FtsK<sub>50C</sub> displayed a dramatic increase in turnaround frequency on the anti substrates that depended on the number of sequences. The anti 1-mer and 5-mer turnaround frequencies were 47 ± 4% (77/162 on 13 tethers) and 91 ± 2% (150/165 in 7 tethers), respectively (Fig. 3B). Fig. 3C shows typical traces of single FtsK<sub>50C</sub> particles on anti 1-mer and anti 5-mer substrates (Movies 2 and 3, which are published as supporting information on the PNAS web site). FtsK efficiently bounced off the anti 5-mer substrate (compare Fig. 3 C and D). As predicted, FtsK<sub>50C</sub> particles often oscillated between the anti orientation test sequence and naturally occurring directionality sequences in the lambda portion of the tether, just as FtsK<sub>50C</sub> oscillates around dif on the E. coli chromosome (Movie 1). FtsK<sub>50C</sub> also fairly consistently bounced off the 5-mer GGGCAGGG derivative (68 ± 5%; 50/73 on 11 tethers), which satisfies the GNGNAGGG motif but not the RAG motif. The 5-mer sequence of GGCAGGGG, which is a RAG but not a GNGNAGGG motif, was 2.2 times less efficient at reversing FtsK direction (31 ± 8%; 10/32 on five tethers).

To demonstrate that a specific sequence is responsible for reversing FtsK<sub>50C</sub>, rather than base composition or G content, we used a 5-mer substrate in the anti orientation with the test sequence partially scrambled (GGCGGGAGG). The scrambling is only partial because the test sequence and the scrambled sequence share the same base (G) in five positions (GGCGGGAGG). On the partially scrambled 5-mer tether, FtsK reversed in 30 ± 5% (23/76 in 18 tethers) of observations, down from 91% on the anti 5-mer (Fig. 3B). Even one copy of the anti 1-mer is significantly better at turning FtsK<sub>50C</sub> around than five scrambled sequences (47% vs. 30%, respectively). Thus, we rule out the hypothesis that base content or G content is the dominating factor. It is clear that a number of octamers have a directing effect, although, thus far, GNGNAGGG shows the strongest effect.

Our ability to track the translocase in real time as it moves over an FRS allows us to measure the probability of recognition for the FtsK translocase. When FtsK<sub>50C</sub> approached an FRS from the 3′ end of the G-rich strand, the turnaround percentage was 39 ± 4% (192 observations) on the c-tether and 47 ± 4% (162 observations) on the anti 1-mer tether. Assuming that each test sequence is acting independently on the anti 5-mer tether, the turnaround percentage is 38 ± 5% per sequence. Thus, all our estimates of turnaround probability are in remarkably good agreement.

Although it may seem surprising that the directional signal is not deterministic, an optimal physiological search strategy using recognition sites with an imperfect skew requires a fractional turnaround probability. Consider a DNA substrate with n directional sequences per kb with fractional skewness s (defined as skew/100). If an FtsK motor binds at distance D kb from dif and has a velocity v, then what is the turnaround probability, p, that minimizes the expected time to reach dif? p = 1 is the worst possible value, because the motor will oscillate forever between two opposing sequences. However, p = 0 is clearly not optimal, because it is equivalent to not having any directional sequences. Assuming that the motor moves in a random direction after binding the DNA, then the expected time to reach dif can be expressed as (see Supporting Text):

$$T(D) = \frac{D}{v} \left[ 1 + \frac{2(1-s)}{(1-p)s - (1-s)} \right] + \frac{1-p}{p(1-p)s - (1-s)} \frac{1}{nv^2}$$

[1]

The expected time to reach dif is plotted in Fig. 4A for four skew values. Exceedingly high or low turnaround probabilities sharply increase the time required for the motor to reach dif. The expected time to reach dif is minimized at the optimal turnaround probability, which for a highly skewed sequence (s = 1) simplifies to:

Levy et al.
Here, \( N_+ = Dn(1 - s) \) is the expected number of directional sites between the motor and \( \text{dif} \) that are pointing in the opposing direction. For reasonable \textit{in vivo} values of \( D = 100 \text{ kb}, n = (1/3) \text{ kb}^{-1} \), and \( s = 0.90 \), the optimal probability is \( p = 0.30 \), a surprisingly low value but very close to the measured turnaround probability of 0.39 ± 0.04 for GNGNAGGG on its natural substrate (Fig. 4A). The optimal turnaround probability as a function of the expected number of FRSs in the anti orientation (pointing away from \( \text{dif} \)) is provided in Fig. 9, which is published as supporting information on the PNAS web site.

**FtsK Acts as a Molecular Wirestripper.** As FtsK draws DNA through its stationary position in the closing septum, it must pass potential roadblocks, such as bound RNA and proteins. To examine this “wirestripping” capability, we measured FtsK-mediated displacement of a radiolabeled 21-nucleotide, parallel pyrimidine–purine–pyrimidine triplex (Fig. 5A), which is a method that was used in the past to study DNA translocation (15, 25). These structures form only at acidic pH, because they require the protonation of cytosine for formation but, after assembly, remain relatively stable at neutral pH (26). Upon displacement at neutral pH, cytosines are rapidly deprotonated, preventing reassembly of the triplex (15) and constituting an irreversible signal for wirestripping and translocation past a given sequence. Using this substrate, we found that FtsK indeed disrupts a DNA triplex helix in an ATP-dependent manner (Fig. 5B).

We next showed that a 5-mer FRS in the anti orientation near the DNA triplex protects the structure from FtsK-mediated wirestripping. A 5-mer repeat of the sequence GGCCAGGGG was placed in either the iso or anti orientation between an \( \approx 3 \text{-kb “antenna” sequence and a DNA triplex.} \) FtsK should most often load within the antenna and therefore must pass the GGCCAGGGG sequences to encounter the triplex. A 5-mer repeat of the partially scrambled sequence, GGAGGCGGGG, was also tested in the anti orientation. On the iso substrate, FtsK displaced the triplex approximately twice as rapidly as on the anti substrate. On the scrambled substrate, the rate of displacement was approximately equal to that of the iso substrate (Fig. 5C). These results confirm in bulk our single-molecule observations that GGCCAGGGGG reverses FtsK in an orientation- and sequence-specific manner. The fairly modest difference in triplex displacement rate between iso and anti substrates is perhaps not surprising. In \textit{singulo}, FtsK almost always passes the sequence eventually, which would lead to a displacement signal in the bulk assay. Furthermore, the natural polarity of FRSs in the antenna region points toward the directionality patch and may continuously bounce FtsK back toward the triplex.

To examine further the specificity of the FRS, we repeated the assay with several test sequences containing point mutations. Mutations of A to T in GGCGAAGGGG completely abrogated recognition by FtsK (Fig. 10A, which is published as supporting information on the PNAS web site). To evaluate GNGNAGGGG and RAG motifs separately, we changed either the first or the last G in the GGCCAGGGG sequence to T. A 5-mer repeat of the sequence GGCCAGGGGT, a GNGNAGGG but not a RAG sequence, behaved similarly to the complete GGCCAGGGG, as evidenced by the fact that displacement rates were similar to those using the full GGCCAGGGG substrates (Fig. 10B). The mutation of the first G to a T tested a RAG sequence (GGCAGGGG) that does not satisfy the GNGNAGGGG motif. A 5-mer repeat of this sequence was only slightly better than the iso 5-mer substrate, suggesting less efficient recognition (Fig. 10C).

### Discussion

FtsK is a fast and powerful DNA translocase that is responsible for critical tasks in cell division: clearing DNA from the closing septum and promotion of chromosome dimer resolution. FtsK must operate not just quickly but in the right direction; FtsK activity would be counterproductive if it pushed the two \( \text{dif} \) sites away from each other. Our previous single-molecule work with purified FtsK50C demonstrated that the DNA sequence directs the translocase (5). Here, we demonstrate that a specific instance of the GNGNAGGGG motif, its complement, or both are effective in specifying FtsK’s directionality and that the skew of GNGNAGGGG well explains FtsK’s action \textit{in vivo}. We found that FtsK also recognizes some related sequences but not as efficiently as the GNGNAGGGG sequence we tested. Remarkably, despite moving at \( \approx 5 \text{ kb/sec} \), FtsK50C can read the DNA sequence at a base pair level, as the transversion of A to T abolishes recognition, indicating a major groove interaction. To our knowledge, no other DNA translocase

\[
P_{\text{opt}} = \frac{1}{1 + \sqrt{2Dn(1 - s)}} = \frac{1}{1 + \sqrt{2N_+}}
\]

---

**Fig. 5.** FtsK50C is a molecular wirestripper. (A) Schematic depicting FtsK50C displacing a DNA triplex substrate. FtsK50C (blue circle) likely binds within the 3-kb duplex region because of relative size, and translocation into the DNA triplex (jagged line) leads to triplex displacement. (B) Triplex displacement by FtsK requires ATP hydrolysis. Varying amounts (0–152 nM) of FtsK50C were incubated for 15 min at 25°C with 3 mM ATP (lanes 1–4), no ATP or 3 mM ATP5 with 152 nM FtsK50C controls (lanes 5 and 6), or no ATP and no FtsK control (lane 7). Lane 8 is identical to lane 4, but the reaction mixture was heat-denatured at 75°C for 5 min before loading. (C) Molecular wirestripping by FtsK50C is controlled by its directionality sequence. Schematic diagram of DNA triplex substrates indicating the location of directionality sequences (top). DNA triplex displacement reactions were performed as described in Materials and Methods, using substrates with 5-mer repeats of the sequence GGCCAGGGG in iso orientation (black), anti orientation (red), or the partially scrambled sequence GGAGGCGGGG in the anti orientation (blue).
is known to change direction when encountering a specific DNA sequence.

We propose that FtsK’s directionality-sensing mechanism monitors DNA sequence during translocation and only recognizes the FRS when approaching it from the 3’ end of the G-rich strand (anti orientation). When FtsK encounters an FRS from the 5’ end of the G-rich strand (iso orientation), FtsK travels uninhibited, but in the anti orientation, the motor is momentarily stalled, as evidenced by the abrupt stop and pause observed at turnaround sites. The stalling could be due to the inactivation of the ATPase domain or another component that is important to power stroke generation. FtsK is made up of at least two motors, and only one of the motors is active at any one time (5). A turnaround occurs when translocation resumes in the opposite direction by activating the reverse motor. The pause time may reflect the time necessary to activate the motor that carries FtsK in the opposite direction.

On the anti 5-mer tether, FtsK reversed 91% of the time, as opposed to 11% on the iso 5-mer tether. The FRS acts as a one-way door that preferentially allows FtsK to pass in one direction. However, the turnaround process is not 100% efficient. On the chromosomal region surrounding dif, we determined that the reversal percentage is 39 ± 4% when approaching from the 3’ end of the G-rich strand. Interestingly, a similar high frequency is observed for RecBCD. RecBCD requires Chi to generate a 3’ ssDNA tail for homologous recombination and not for directionality (27). RecBCD recognizes Chi ~30–40% of the time based on bulk phase experiments that measure the attenuation of nuclease activity. This estimate may be a lower bound, because it is based on detection of Chi-containing ssDNA products, and, despite attenuation of nuclease activity, the remaining level of activity may noncovalently degrade some of this ssDNA, leading to an underestimation of recognition efficiency (S. Kowalczykowski, personal communication).

A number of experimental and theoretical studies have been performed to elucidate the optimal search strategy for a diffusional search process such as that performed by restriction enzymes (1, 28). Less well understood is the process relevant for FtsK: namely, a directional, ATP-powered search guided by skewed DNA sequences. We derived a mathematical expression describing the time sequence of values, as evidenced by the wide minima region in Fig. 4A. Nevertheless, a very low or high turnaround probability is clearly undesirable, because it would greatly increase the time to reach dif. DNA motifs with a skewed orientation toward dif are abundant and conserved in many bacterial species (14). Despite many efforts, their roles have largely remained unclear. The only skewed sequence whose function has been described thus far is the RecBCD recognition sequence Chi. In this work, we have demonstrated a unique function for skewed sequences: the control of FtsK directionality. Establishing and maintaining the observed GNGNAGGG skew level is energetically prohibitive. The comparison of the GNGNAGGG skew levels of many bacterial genomes suggests that their common ancestor had already established the GNGNAGGG skew and that there might have been a selective pressure to maintain it after divergence.

Given that the GNGNAGGG skew also switches at the origin of replication in these diverse organisms, it seems likely that the skew was not established for directing FtsK, whose only known roles are near the termination of replication. The presence of the skew in many organisms suggests that it is important in other basic cellular processes that are common among bacteria. For example, Viollier et al. (29) recently demonstrated in Caulobacter crescentus that a nascent replicated chromosomal region is quickly condensed and transported to a specific destination in the daughter cell. This process occurs before septum formation, and thus FtsK activity can be ruled out. The process that moves these chromosome loci must be directional, because the DNA segments travel in a specific direction and to a specific location. The exact proteins involved and the mechanism of this process are poorly understood, but it is possible that the FRS skew is involved. It has been speculated that the skew could be acting as a barrier to horizontal gene transfer between organisms that don’t share this skew, possibly due to FtsK usage of the skew for directionality (30).

In addition to the challenge of determining in which direction to translocate, FtsK must be able to bypass any bound proteins or other roadblocks along the DNA. We show that FtsK can displace a DNA triplex during translocation. FtsK can generate >60 pN of force (5), and therefore it is likely that this writestripping activity can dislodge a wide variety of roadblocks in vivo. However, just as the termination protein Tus specifically stops the activity of helicases, there may be obstacles that specifically stop FtsK. The most obvious candidate is XerD, with which FtsK interacts to activate chromosome donor resolution.

We thank Ed Green and Steven E. Brenner for helpful discussions and time on the multiprocessor cluster. Thomas H. Massey and David Sherratt (both of University of Oxford, Oxford) kindly provided the FtsK from his tagged construct. This work was supported by National Institutes of Health Grants GM31655 (to N.R.C.), GM32543 (to C.B.), GM31657 (to J.L.P.), and GM31655 (to P.J.P.); the Program in Mathematics and Molecular Biology; a Burroughs Wellcome Graduate Fellowship (to O.L.); the Fannie and John Hertz Foundation (J.G.); and U.S. Department of Energy Grants DE-AC03-76SF00098, G1T2LBN, and SNANOB (to C.B.).