A reverse transcriptase-related protein mediates phage resistance and polymerizes untemplated DNA in vitro

Chen Wang1, Manuela Villion2, Cameron Semper1, Colin Coros1, Sylvain Moineau2,3 and Steven Zimmerly1,*

1Department of Biological Sciences, University of Calgary, Calgary, Alberta T2N 1N4, 2Département de Biochimie, Microbiologie et de Bioinformatique, Faculté des Sciences et de Génie, Université Laval Quebec City, Quebec G1V 0A6 and 3Groupe de Recherche en Ecologie Buccale (GREB) and Félix d’Hérelle Reference Center for Bacterial Viruses, Faculté de Médecine Dentaire, Université Laval, Quebec City, Quebec G1V 0A6, Canada

Received March 12, 2011; Revised April 30, 2011; Accepted May 3, 2011

ABSTRACT

Reverse transcriptases (RTs) are RNA-dependent DNA polymerases that usually function in the replication of selfish DNAs such as retrotansposons and retroviruses. Here, we have biochemically characterized a RT-related protein, AbiK, which is required for abortive phage infection in the Gram-positive bacterium Lactococcus lactis. In vitro, AbiK does not exhibit the properties expected for an RT, but polymerizes long DNAs of ‘random’ sequence, analogous to a terminal transferase. Moreover, the polymerized DNAs appear to be covalently attached to the AbiK protein, presumably because an amino acid serves as a primer. Mutagenesis experiments indicate that the polymerase activity resides in the RT motifs and is essential for phage resistance in vivo. These results establish a novel biochemical property and a non-replicative biological role for a polymerase.

INTRODUCTION

Reverse transcriptases (RTs) play many roles in biology, including in retroviruses, retrotansposons, hepadnaviruses, telomerase and organellar retroplasmids (1–5). Most eukaryotic RTs are components of selfish DNAs, and their polymerization activities allow the elements to replicate and spread without necessarily contributing benefits to their host organism. In spite of this general pattern of selfishness, a few eukaryotic RTs have evolved useful functions, such as telomerase and HetA-TART retroelements, both of which protect chromosomal ends from sequence loss during DNA replication (4,6).

RTs in bacteria are less prevalent and less well studied than those in eukaryotes. In fact, we have an incomplete knowledge of the types of bacterial RTs that exist in nature. Two recent compilations have identified a surprising number of uncharacterized RTs and RT-like proteins in bacterial genomes, including at least 20 phylogenetically based groupings with 11 domain architectures (7,8). Of these, only three types have been substantially characterized: group II introns, diversity-generating retroelements (DGRs) and retrons.

Group II introns are the most abundant RTs in bacteria by far (8), and have the classical properties of retrotansposons, being capable of efficient retromobility (9,10). Their mobility mechanism is known as target-primed reverse transcription (TPRT), and uses the cooperative activities of a catalytic, self-splicing intron and an intron-encoded RT. Group II introns insert at high efficiencies into homing sites of defined sequences, and at lower frequencies into ectopic sites.

DGRs in contrast are not selfish, but contribute a useful function to their hosts through the generation of sequence diversity in target genes. The Bordetella phage DGR, for example, uses a template-dependent, RT-mediated process to introduce nucleotide substitutions into the variable region of the phage gene mtd. The mtd gene encodes a tail protein responsible for host recognition, and the variability allows the phage to adapt its tropism to the dynamic Bordetella cell surface, which changes between in vivo and ex vivo phases (11). Other DGRs are not phage-associated, and may diversify cellular genes (12). The third characterized bacterial retroelement, the retron, carries out a specialized reverse transcription reaction to
produce branched RNA–DNA molecules called multi-copy single-stranded DNAs (msDNAs) (13). Retrons are not independently mobile, nor has a biological function been identified to date (14).

Among the remaining putative RTs in bacteria, there is little indication for active retromobility, as would be suggested by multiple RT copies in a genome. Instead, it has been predicted that the RTs contribute beneficial functions, analogous to telomerase and DGRs (7,8). Supporting this hypothesis, two uncharacterized RT classes are associated with bacterial CRISPR/Cas defense systems, and presumably help protect against invading phages and plasmids (7,8,15,16). Three other uncharacterized RTs (AbiK, AbiA, Abi-P2) have been experimentally associated with abortive phage infection (17,18) or phage exclusion mechanisms (19), but none of these has been studied extensively.

Abortive infection (Abi) is a type of bacterial defense system in which a virulent phage DNA is injected into bacteria, but phage maturation is blocked and most infected cells die (20,21). Abi systems are prevalent across bacteria, and can act specifically against a single phage group, or more generally against multiple groups of phages. Abi systems are typically mediated by a single host gene, which is often plasmid-encoded. In many cases the exact modes of action are incompletely understood.

The AbiK system of Lactococcus lactis is encoded by the native plasmid pSRQ800 of a lactococcal strain isolated from raw milk, an ecological niche known to contain many virulent phages (17). AbiK provides broad immunity against the most predominant lactococcal phage groups, typically reducing infectivity by 105. The system is encoded by a single gene, abiK, which is constitutively transcribed. It has been noted for some time that the N-terminal sequence of AbiK is related to RTs (22).

In the present study, we investigate the biochemical properties of the AbiK protein and show that it is an active polymerase; however, it does not possess the activities expected for an RT. Rather AbiK polymerizes ‘random’ DNA sequence analogous to a terminal transferase. Furthermore, the DNA appears to be attached covalently to AbiK, putatively through an amino acid that serves as primer for DNA polymerization. Our results expand the mechanistic repertoire of RT-related enzymes, and suggest that the AbiK system is a biochemically novel mechanism of phage resistance that is mediated by an RT-related polymerase.

MATERIALS AND METHODS

Plasmid constructs

The wild-type AbiK expression construct was made by PCR-amplifying AbiK sequence from the plasmid pSRQ818 (17) with the primers AbiK-S-B and AbiK-AS-K (Supplementary Table S1), and cloning into the BamHI and KpnI sites of pQE30 (Qiagen, Valencia, CA, USA) to generate pQE30-AbiK. The BamHI-Smal fragment of pQE30-AbiK was subcloned into pGEX4T1 (GE Healthcare, Piscataway, NJ, USA) to generate the plasmid pGEX4T1-AbiK. Mutations were introduced into AbiK sequence by recombinant PCR of pGEX4T1-AbiK using the primer pairs AbiK-m-S and AbiK(DI-DVII)AS, and AbiK-m-AS and AbiK(DI-DVII)S (Supplementary Table S1).

Constructs to test abortive infection in vivo were made by subcloning mutant sequences into the BglII and XhoI sites of the lactococcal plasmid pSRQ823. pSRQ823 is a pNZ123-based vector containing the EcoRI-XbaI abiK-fragment from pSRQ818 which originated from the natural L. lactis plasmid pSRQ800 (17). For the gst-abiK fusion construct in pNZ123, the gst-abiK fragments from pGEX4T1-AbiK wt were PCR-amplified with Pwo polymerase and primers rbs-gst-abiKFor and 3’_pGEX sequencing primer (Supplementary Table S1), thereby introducing a ribosome binding site. The PCR fragment was digested with XhoI and cloned into the XhoI site of pNZ123, with additional steps to remove Hsp70 (24).

Preparation of recombinant AbiK-GST

An overnight culture of Escherichia coli BL21 harboring the plasmid pGEX4T1-AbiK was grown in LB (23) at 37°C with ampicillin (100 μg/ml), and then diluted 100-fold into one litre of 2× YT medium (23), and grown to an OD600 of 0.5. Cells were induced with 1 mM IPTG for two hours at room temperature, harvested, and stored at −70°C. Protein purification was carried out according to the manufacturer’s manual for Glutathione Sepharose 4B beads, using the column binding protocol (GE Healthcare), with the following minor modifications and an additional step to remove Hsp70 (24). Frozen pellets were resuspended in 35 ml lysis buffer (16 mM Na2HPO4, 4 mM NaH2PO4, pH 7.3, 150 mM NaCl, 1% Triton X-100, 5 mM DTT, 1 mM PMSF) and broken with a French press. Then, 3.5 ml of 10× ATP stock (16 mM Na2HPO4, 4 mM NaH2PO4, pH 7.3, 150 mM NaCl, 1% Triton X-100, 5 mM DTT, 100 mM ATP, 200 mM MgCl2, 500 mM KCl) were added to the clarified lysate and incubated at 37°C for 10 min, followed by addition of 200 μl (7.6 mg) of denatured E. coli cellular proteins (24), and an additional incubation at 37°C for 20 min. The lysate was clarified by centrifugation, and applied to a 300 μl bed of Glutathione Sepharose 4B beads (GE Healthcare). Two washes were performed with 10 and 5 ml of lysate buffer containing either added 10 mM ATP, 20 mM MgCl2 and 50 mM KCl (wash buffer 1) or 5 mM ATP, 20 mM MgCl2 and 50 mM KCl (wash buffer 2). Purified protein was stored at −20°C in 25 mM Tris–HCl, pH 8.0, 50 mM NaCl, 5 mM reduced glutathione, 5 mM DTT and 50% glycerol. Mutant AbiK preparations were made in the same way, except that only 3.6 mg of denatured protein were used to remove Hsp70.

For thrombin treatment of AbiK-GST, 1 μg of the AbiK-GST preparation was incubated with 1 U of thrombin (GE Healthcare) at room temperature, either for 1.5 h in 1× RT buffer (50 mM Tris–HCl, pH 8.5, 100 mM KCl, 2 mM MgCl2, 5 mM DTT) (Figure 3A, Table 1), or for 1.5 or 17.5 h in 20 mM Tris–HCl, pH 8.5, 50 mM KCl, 5 mM MgCl2, and 5 mM DTT.
Table 1. In vitro polymerization assays with purified AbiK-GST protein

<table>
<thead>
<tr>
<th>Assay</th>
<th>32P incorporation (cpm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AbiK-GST + [α-32P]dTTP + poly rA/oligo(dT)$_{18}$</td>
<td>141 000 ± 1700</td>
</tr>
<tr>
<td>AbiK-GST + [α-32P]dTTP</td>
<td>141 000 ± 4000</td>
</tr>
<tr>
<td>AbiK-GST + [α-32P]dTTP + RNase A</td>
<td>157 000 ± 7800</td>
</tr>
<tr>
<td>AbiK-GST + [α-32P]dTTP + RNase A (10 min)</td>
<td>109 000 ± 2300</td>
</tr>
<tr>
<td>AbiK-GST + [α-32P]dTTP + 0.2 mM dATP, dCTP, dGTP</td>
<td>6500 ± 700</td>
</tr>
<tr>
<td>AbiK-GST + [α-32P]dTTP + 0.05 mM dATP, dCTP, dGTP</td>
<td>22 000 ± 15 000</td>
</tr>
<tr>
<td>AbiK-GST + [α-32P]dTTP + 0.2 mM dATP, dCTP, dGTP</td>
<td>5400 ± 300</td>
</tr>
<tr>
<td>AbiK-GST + [α-32P]dTTP + 0.2 mM dCTP, dGTP</td>
<td>5600 ± 200</td>
</tr>
<tr>
<td>AbiK-GST + [α-32P]dTTP + 0.2 mM dCTP, dGTP</td>
<td>4400 ± 800</td>
</tr>
<tr>
<td>AbiK-GST + RNase + [α-32P]dTTP + poly rA/oligo(dT)$_{18}$</td>
<td>215 000 ± 21 000</td>
</tr>
<tr>
<td>AbiK-GST + RNase + [α-32P]dTTP</td>
<td>230 000 ± 23 000</td>
</tr>
</tbody>
</table>

*See Supplementary Data for experimental description.

RESULTS

AbiK polymerization and protein labeling assays

Unless noted otherwise, 1 μg AbiK-GST preparation was incubated with 10 μCi [α-32P]dTTP in 1× RT buffer at 37°C for 10 min. This basic reaction was followed by different treatments. For chased reactions, the labeling reactions were followed by addition of either 0.2 mM each dNTP or 0.2 mM dTTP, and a further 10 min incubation. For some reactions, 2 μg proteinase K was added and incubated for 10 min at 37°C. For nuclease analyses, reactions were incubated at 37°C for 10 min with either 2 μg RNase A or with 10 U DNase I (Invitrogen, Carlsbad, CA, USA). For resolution by SDS–PAGE, samples were mixed with SDS loading dye, boiled and resolved on a 10% gel, and either dried and subjected to phosphorimaging or stained with Coomassie blue (23) or silver (25). For resolution on PAGE–urea gels, samples were extracted with phenol/chloroform/isooamyl alcohol (25:24:1) with 1% acrylamide carrier and precipitated with ethanol in the presence of 0.3 M NaOAc (pH 5.2), followed by heating to 80°C in 50% formamide. Samples were resolved on an 8% PAGE–urea gel, dried and phosphorimaged, or on a 20% PAGE–urea gel followed by phosphorimaging. Assays with poly rA/oligo (dT)$_{18}$ were carried out as previously described (26).

Cloning and sequencing of AbiK-synthesized DNAs

The cloning strategy was based on the observation that AbiK can use an oligonucleotide as a primer if provided in excess. Five micrograms of AbiK preparation were incubated with 0.5 μg oligonucleotide AD96 or PQEREV and 0.2 mM each dNTP (or 0.2 mM dATP, dCTP, dGTP, or 0.2 mM dATP, dGTP, dTTP) in 25 μl of 1× RT buffer at 37°C for 15 min. The reaction was stopped by heating to 95°C for 5 min, and digesting with 2 μg of proteinase K at 37°C for 1 h. After purification using a DNA Clean kit (Zymo Research, Orange, CA, USA), the eluted DNA was tagged with 0.5 mM dGTP and 20 U terminal transferase (New England Biolabs, Beverly, MA, USA) according to the manufacturer’s protocol. DNAs were amplified by PCR with primer pairs AD96 and AD98, or PQEREV and AD98, and the PCR products were gel-purified, cloned and sequenced.

In vivo assays for abortive infection

Clones containing AbiK mutants in pSRQ823 were transformed into L. lactis MG1363 by electroporation (27). The plasmid content of each transformant was confirmed by plasmid isolation (28) and sequencing. Some lactococcal plasmids were also extracted using QIAquick Spin Miniprep columns but with a slightly modified protocol. Cells were first treated by adding sucrose (25%) and lysozyme (30 mg/ml) into the P1 buffer, and incubated at 37°C for 15 min. L. lactis MG1363 strains were grown at 30°C in M17 medium (OXOID, Nepean, Ontario, Canada) supplemented with 0.5% glucose (GM17) and 5 μg/ml chloramphenicol.

To measure the phage resistance phenotype, the efficacy of plaquing (EOP) of lactococcal phage p2 (936 group, double-stranded DNA genome, Siphoviridae family, Caudovirales order) was estimated as described previously (22). Briefly, phage p2 was propagated on the wild-type phase-sensitive laboratory strain L. lactis MG1363 in GM17 medium supplemented with 10 mM CaCl$_2$ (29). The EOP was calculated by dividing the phage titer of an AbiK construct in MG1363 by the titer of the sensitive control (pNZ123 in MG1363).

RESULTS

The family of AbiK-related proteins contains RT domains

To gain further insight into the domain structure of the AbiK protein, its 599 amino acid sequence was aligned with its four closest relatives in GenBank, as well as representative group II intron RT sequences (Figure 1, Supplementary Figure S1). Only four relatives were chosen for the alignment because more distant sequences do not align across the entire length of AbiK sequence. The seven domains characteristic of RTs are identifiable in the AbiK-related sequences, although RT2 and RT7 align ambiguously. The identified RT motifs correspond to the finger and palm domains of the polymerase protein structure (by analogy with the HIV RT crystal structure), with the putative active site residues located in RT3 and RT5 (outlined asterisks in Figure 1, Supplementary Figure S1) (30). Thumb domains of RTs are typically located downstream of RT7, but are not highly conserved among RT groups. For AbiK, the region expected to be the thumb domain does not align with the X/thumb domain of group II introns. In fact, the ~200 amino acid C-terminal sequence of AbiK and its relatives does not have similarity to any other proteins in GenBank. The region is however inferred to have an Abi-specific role, in addition to being the polymerase thumb domain, because a 46 amino acid C-terminal deletion was previously shown to eliminate the phage resistance phenotype (22).
In light of the unusual biochemical properties found for AbiK in this study (below), it is of interest to reexamine the relationship between AbiK and other polymerase families. BLAST searches consistently identify AbiK as containing RT motifs, based on the Conserved Domain Database (CDD) of NCBI (not shown); however, AbiK sequence is only distantly related to characterized RTs [see ref. (8) for alignments among bacterial RTs]. Because RTs belong to the superfamily of RNA-dependent RNA polymerases (RdRPs) (31), AbiK was aligned to the RdRP1 family. Alignment is quite limited, with only domains 3, 4 and 5 aligning unambiguously (Figure 1, Supplementary Figure S1 and data not shown). Thus, sequence alignments do not suggest that AbiK is more closely related to RdRPs than to group II intron RTs. It can be concluded that AbiK belongs to the sequence family of RTs, regardless of its biochemical properties.

Expression and purification of AbiK in E. coli

The AbiK protein was expressed in E. coli as a fusion with a GST tag. In the initial affinity purification method, three bands were produced (Supplementary Figure S2A), the largest of which was the size expected for AbiK-GST, while the smallest band comigrated with GST, indicating that the majority of AbiK-GST is degraded, leaving the 26 kDa GST domain. The third band was determined by mass spectrometry to be Hsp70. Consequently, a modified purification protocol was used to eliminate the Hsp70 contamination along with substantial background biochemical activities (see ‘Materials and Methods’ section). In the resulting AbiK preparation, the GST tag could be cleaved from AbiK-GST with thrombin (Supplementary Figure S2B), while in the initial preparation it could not.

Purified AbiK has DNA polymerase activity

Purified AbiK was assayed for RT activity using conventional RT assays with poly rA/oligo(dT)18 substrate. High levels of [α-32P]dTTP incorporation were observed; however, incorporation occurred even without the poly rA/oligo(dT)18 substrate (Table 1). To test for a copurifying RNA in the AbiK preparation that might be a template, the preparation was treated with RNase A, either before or during the assay. RNase A does not degrade poly rA/oligo(dT)18 because it cleaves at pyrimidine residues. The fact that polymerization occurred regardless of the RNase treatment argues against the presence of an RNA template in the AbiK preparation. Interestingly, addition of non-radiolabeled dATP, dCTP and dGTP drastically reduced [α-32P]dTTP incorporation, which would not be expected for templated polymerization, but would occur for non-templated polymerization such as for a terminal deoxynucleotidyl transferase. Treatment of the AbiK-GST preparation with thrombin had little effect on [α-32P]dTTP incorporation, indicating that the GST tag minimally perturbs AbiK’s activity in vitro. Subsequently, the GST tag was found to also have a minimal effect on AbiK function in vivo (see below and Table 2).

The polymerized DNA appears to be attached to AbiK through a covalent bond

When AbiK reaction products were resolved on a polyacrylamide–urea gel, most products remained in the wells (Figure 2A, lanes 1 and 2). Treatment with protease allowed the resolution of heterogeneous products of <100 nt (Figure 2A, lane 4), which were extended to several hundred nucleotides after a chase with dNTPs (Figure 2A, lanes 5 and 6). Similar to the RT assay data (Table 1), the inclusion of dATP, dCTP and dGTP reduced the incorporation of the radiolabeled dTTP (Figure 2A, lanes 3 and 6). Phenol extraction of the products resulted in complete loss of signal in the gel (Figure 2A, lane 7), suggesting a covalent attachment between the radiolabeled DNA and AbiK. Together, the data indicate that the AbiK reaction product is a short radiolabeled DNA attached to the AbiK protein, which becomes elongated upon chasing with dTTPs.
To determine the shortest DNA linked to AbiK, AbiK-GST was incubated with [α-32P]dTTP for different time points, with some sample chased with dNTPs. Products were purified on a G-50 Sephadex column to remove unincorporated nucleotides, digested with proteinase K, and resolved on a 20% polyacrylamide–urea gel. The shortest detectable products migrated as heterogeneous bands of 18–25 nt (Figure 2B). Because protease digestion will not remove all amino acids from the DNA, and because it is not certain how the attached amino acids will affect gel migration, the shortest DNAs may be smaller than 18–25 nt. Other experiments supported the hypothesis of a short DNA produced in the labeling reaction. In initial cloning attempts, AbiK-generated DNAs were produced by incubating AbiK with a low concentration of dTTP, chasing with dNTPs, and then tailing with G’s. The resulting DNAs failed to PCR amplify with (dT)18 and (dC)18 primers, suggesting that the labeling reaction polymerizes too few T’s to anneal with the (dT)18 primer.

To test more directly for a covalent linkage between AbiK and DNA, a thrombin-treated preparation containing both AbiK-GST and AbiK was incubated with [α-32P]dTTP, boiled in SDS loading buffer, and resolved by SDS–PAGE (Figure 3). Radiolabeled bands were observed that comigrated with each form of AbiK, indicating that each polypeptide undergoes the same radiolabeling reaction, and confirming that the attached DNA is short and does not affect protein migration on the SDS–PAGE gel (Figure 3A, lanes 1–3). Chasing of the reaction with dNTPs resulted in greatly retarded migration for both protein forms (Figure 3A, lanes 4–7).

The reaction was further characterized by nuclease digestions. Consistent with previous data (Table 1), preincubation of AbiK with RNase A for 10 min did not abolish either labeled or chased reaction products (Figure 3B, lanes 1–4). In addition, the products of the AbiK reaction were resistant to RNase treatment (Figure 3B, lane 9). However, the reaction products had a differential susceptibility to DNase I. The chased products were sensitive to DNase I as expected (Figure 3B, lane 10), while the unchased products were resistant (Figure 3B, lane 7). One possible explanation for the difference is a conformational change of the AbiK protein, which exposes the attached DNA during the chasing reaction (see Discussion section).

The sequence of AbiK-polymerized DNAs indicates non-templated polymerization

Additional experiments failed to identify a candidate RNA or DNA template in the AbiK preparation. The preparation was treated with protease and/or phenol extraction followed by SDS–PAGE and silver staining. No nucleic acids were detectable (not shown). In another approach, the radiolabeled DNAs produced by AbiK were used as a probe in Southern and northern blot experiments, using blots of either DNA or RNA isolated from E. coli harboring the AbiK expression plasmid. No hybridization to E. coli cellular DNA or RNA was detectable (not shown), indicating that the AbiK-synthesized DNA sequence does not correspond to a DNA or RNA in the E. coli strain in which AbiK is expressed.

Consequently, the AbiK-synthesized DNAs were cloned and sequenced, taking advantage of the fact that AbiK can use an oligonucleotide as a primer to a small extent when the oligonucleotide is provided in excess (Figure 4A, Supplementary Figure S3). The sequences from 10 independent clones did not match either E. coli sequence or other sequence in GenBank, and were unique with respect to each other (Figure 4B, Supplementary Figure S4A), suggesting the absence of a template. To further test our hypothesis of non-templated, ‘random’ polymerization, AbiK was incubated with either dATP + dCTP + dGTP, or dATP + dGTP + dTTP. As hypothesized, the 10 cloned sequences from each reaction were composed of the corresponding nucleotides, and had no matches in GenBank (Figures 3D and 4C, Supplementary...
We conclude that AbiK polymerizes DNA without a template, and incorporates bases without sequence specificity, analogous to a terminal transferase. While there may be a modest bias for adenosines in the sequences (Supplementary Figure S4), nevertheless, all four bases are readily incorporated.

Mutations in the RT motifs abolish polymerization in vitro

To test whether the RT motifs of AbiK are responsible for the polymerization observed in vitro, a series of mutations was made across the seven RT domains (Figures 1 and 5, Supplementary Figure S1). Mutant proteins were expressed in E. coli, affinity purified, and tested for polymerization activity using the protein-labeling assay. Protein amounts were calibrated on a silver-stained gel to account for differing amounts of GST in the preparations (Supplementary Figure S5). All RT domain mutants were found to have significantly lower activity than wild-type AbiK (Figure 5B, Table 2), including Mut3 and Mut5, which lack the presumed catalytic residues. Of the RT mutants, only Mut2b and Mut7 had faint signals in long exposures; these mutants correspond to the least conserved of AbiK’s RT domain sequences (Figure 1, Supplementary Figure S1). Overall the loss of activity supports the hypothesis that polymerase activity of AbiK lies in its RT motifs. We cannot exclude the possibility that the AbiK mutations affected protein folding rather than polymerase activity per se; however, we note that eight different mutations showed a similar effect in greatly reducing or abolishing polymerase activity in vitro.

Mutations in the RT domains block abortive infection in vivo

AbiK mutants were also tested for the abortive phage infection phenotype in vivo by introducing their sequences into a high-copy plasmid (pNZ123) in L. lactis and challenging the cells with phages. Consistent with previous work (22), the wild-type AbiK sequence provided an efficacy of plaquing (EOP) of $2.6 \times 10^{-3}$ (Table 2) relative to a sensitive strain, while the AbiK-GST construct remained substantially active with an EOP of $3.7 \times 10^{-7}$. In contrast, mutations in the RT domains had levels of abortive infection that were reduced or undetectable, with the exception of Mut2b and Mut7 (Table 2). The less drastic impairment for these two mutants in vivo compared to in vitro (Table 2) might reflect the instability of AbiK activity in vitro or the availability of accessory proteins in vivo. Nonetheless, the assays provide good correlation between in vitro polymerization and in vivo abortive infection, and indicate that an activity of the RT domains (DNA polymerization) is critical to the abortive infection phenotype.

**DISCUSSION**

AbiK is a DNA polymerase with novel biochemical properties

This study establishes a novel type of DNA polymerase that synthesizes long DNAs of non-specific sequence (>500 nt) that are linked to the polymerase protein, likely through a covalent bond. These properties are not
typical of RTs; however, they reside in the RT sequence motifs, because mutations in the motifs abolish the self-labeling and DNA polymerization activities. To our knowledge, this is the first example of an enzyme possessing activities of both self-priming and random sequence synthesis.

Several lines of evidence support the unusual finding of non-templated polymerization by AbiK: (i) no nucleic acids were detectable in the AbiK preparation by silver staining or hybridization experiments; (ii) incorporation of \(^{32}\)PdTTP competes with incorporation of dATP, dCTP, and dGTP; (iii) long DNAs are produced by either incubation with dNTPs, or with dTTP alone; and (iv) the cloned sequences generated by AbiK do not correspond to \(E.\ coli\) sequence or any other sequence in GenBank. The most critical experiment, however, was the incubation of AbiK with only three nucleotides (either dATP+dCTP+dTTP or dATP+dGTP+dTTP), with the resulting cloned sequences consisting of multiple ‘random’ sequences of only those three nucleotides. Thus, AbiK appears to incorporate any available nucleotides into long polymers, similar to a terminal transferase.

Although non-templated incorporation was unanticipated, some eukaryotic RTs have been reported to carry out non-templated incorporation, albeit on a smaller scale and usually under artificial conditions. The RT of the non-long terminal repeat (non-LTR) element R2 of \(Bombyx\ mori\), for some mutant RNA templates, is capable of adding more than 50 non-templated nucleotides to its cleaved DNA target (i.e. its primer) before engaging its RNA template. These additions may be pseudotemplated rather than truly random (32,33). Another example is the RT of the mitochondrial retroplasmid Varkud, which can polymerize several non-templated nucleotides, either before engaging an RNA template in the absence of a primer, or when switching from one RNA template to another (34–36). Finally, yeast and human telomerase RT in the presence of manganese can add up to 20 non-templated nucleotides to the 3'-ends of DNA primers, which was noted to be analogous to terminal transferase activity (37). These different eukaryotic RT examples help to rationalize the ability of the bacterial AbiK to carry out non-templated polymerization, but they also suggest that like other RTs, AbiK may perform multiple reactions \(in\ vitro\), and be capable of reverse transcription under conditions not tested, or \(in\ vivo\). Experiments so far have failed to detect RNA-templated polymerization by AbiK \(in\ vitro\) (Table 1; data not shown), and so it appears that under our \(in\ vitro\) conditions, AbiK favors non-templated polymerization.

The self-priming property of lactococcal AbiK was likewise unexpected, but has been seen for the reverse transcriptase of the hepatitis B virus (HBRT) (3). Self-priming appears probable as well for the RTs of the fungal mitochondrial pFOXC retroplasmids (38). For HBRT, DNA polymerization is primed by a tyrosine residue located in the protein’s \(N/C24\) amino acid N-terminal domain. Initiation begins at a specific RNA structure and after reverse transcribing this priming motif, the RT relocates to the 3'-end of the viral RNA intermediate and reverse transcribes the entire genome. The synthesized DNA remains covalently linked to the RT throughout this reaction. For AbiK, we hypothesize that the priming domain is located in its C-terminal region, because this region does not yet have an identified function in abortive phage infection, yet is essential (17). The C-terminal sequence of AbiK contains multiple Y, S and T residues that might serve as priming residues (Figure 1, Supplementary Figure S1).

For the viral HBRT enzyme, there is a conformational change during the transition from priming to elongation phases of reverse transcription, which removes the N-terminal protein priming domain from the RT’s active site, and replaces it with the 3'-end of the RNA genome (39). An analogous rearrangement for AbiK may help
explain the nuclease susceptibility data and unusual kinetics of polymerization. Low concentrations of \([\alpha-32P]\)dTTP produce only a short DNA that is resistant to DNase I treatment (Figure 3B). This may correspond to a 'closed' priming conformation in which the DNA is protected in the active site by the priming domain. (As noted above, the initially polymerized DNA may be significantly shorter than 18–25 nt.) Upon chasing with high concentrations of dNTPs, the DNA rapidly extends to several hundred nucleotides (Figure 2) and is sensitive to DNase I; this may correspond to an 'open' conformation in which the DNA is not protected. Other explanations for the nuclease susceptibility data may involve differential steric effects caused by long and short DNAs.

A novel biological function for a polymerase

As a rule, polymerases function in genome replication; however, here we uncover a non-replicative role for a polymerase, in which an RT-related protein has evolved a function to provide resistance to phage infections (Figure 3B). This may correspond to a 'closed' priming conformation in which the DNA is protected in the active site by the priming domain. Upon chasing with high concentrations of dNTPs, the DNA rapidly extends to several hundred nucleotides (Figure 2) and is sensitive to DNase I; this may correspond to an 'open' conformation in which the DNA is not protected. Other explanations for the nuclease susceptibility data may involve differential steric effects caused by long and short DNAs.

Supplementary Data
Supplementary Data are available at NAR Online.
Table 2. In vivo abortive infection assays with AbiK mutants

<table>
<thead>
<tr>
<th>AbiK allele</th>
<th>Domain mutated</th>
<th>Efficacy of plaquing</th>
<th>32p incorporation (cpm)</th>
<th>Mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mut1</td>
<td>1</td>
<td>8.9 ± 0.7 × 10^-1</td>
<td>–</td>
<td>P89A, K90A</td>
</tr>
<tr>
<td>Mut2a</td>
<td>2</td>
<td>N.D.2±</td>
<td>±</td>
<td>R96A</td>
</tr>
<tr>
<td>Mut2b</td>
<td>2</td>
<td>2.6 ± 0.7 × 10^-8</td>
<td>+</td>
<td>P101A</td>
</tr>
<tr>
<td>Mut3</td>
<td>3</td>
<td>N.D.2±</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Mut4</td>
<td>4</td>
<td>8.8 ± 0.9 × 10^-1</td>
<td>–</td>
<td>P214A, G216A</td>
</tr>
<tr>
<td>Mut5</td>
<td>5</td>
<td>8.7 ± 1.1 × 10^-1</td>
<td>–</td>
<td>D247A, D248A</td>
</tr>
<tr>
<td>Mut6</td>
<td>6</td>
<td>9.0 ± 2.4 × 10^-1</td>
<td>–</td>
<td>K281A</td>
</tr>
<tr>
<td>Mut7</td>
<td>7</td>
<td>5.3 ± 2.2 × 10^-10</td>
<td>–</td>
<td>F321A</td>
</tr>
<tr>
<td>pSRQ823 WT</td>
<td></td>
<td>2.6 ± 1.0 × 10^-3</td>
<td>–</td>
<td>N.D.d</td>
</tr>
</tbody>
</table>

aMean ± standard deviation, n = 3.

bRT activities are summarized from Figure 5, with weak activities determined by overexposure.

cNo data. Construct could not be cloned in L. lactis.

dNo data. Wild-type protein was not purified without the GST tag.

FUNDING
This work was funded by a strategic grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada to S.M. and S.Z [03601-CG083160]. Salary support for S.Z. was from the Alberta Heritage Foundation for Medical Research. Funding for open access charge: Natural Sciences and Engineering Research Council (Canada).

Conflict of interest statement. None declared.

REFERENCES


