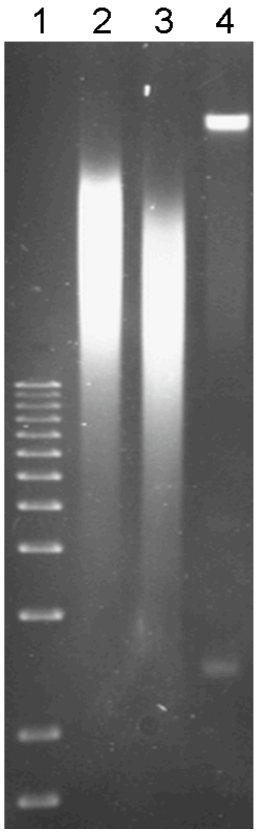
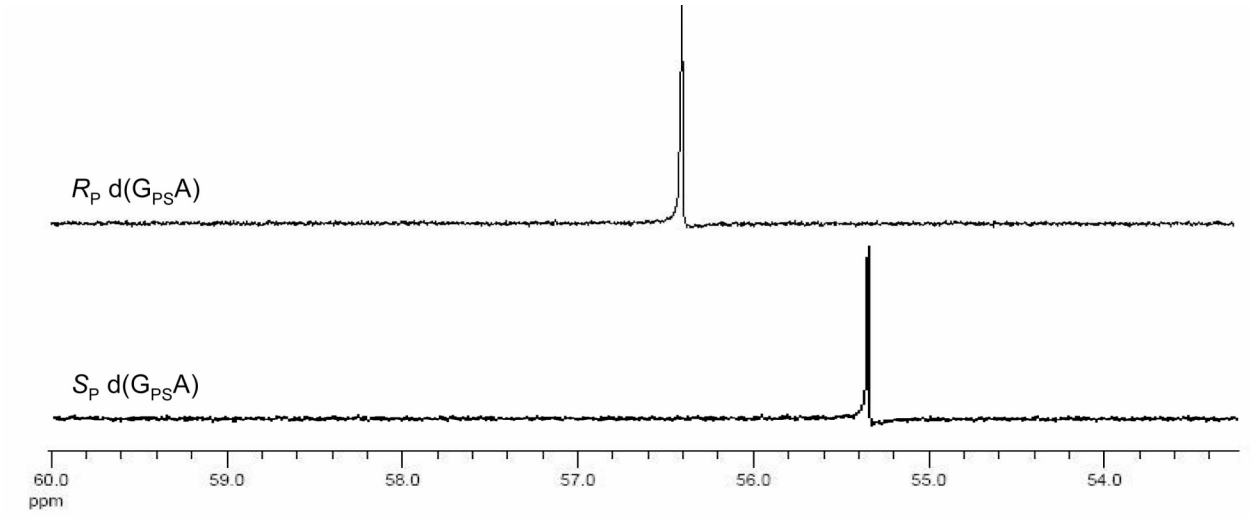
Supplementary Information for
"Phosphorothioation of DNA in bacteria by *dnd* genes"

Lianrong Wang, Shi Chen, Tiegang Xu, Koli Taghizadeh, John Wishnok, Xiufen Zhou, Delin You, Zixin Deng, Peter C. Dedon

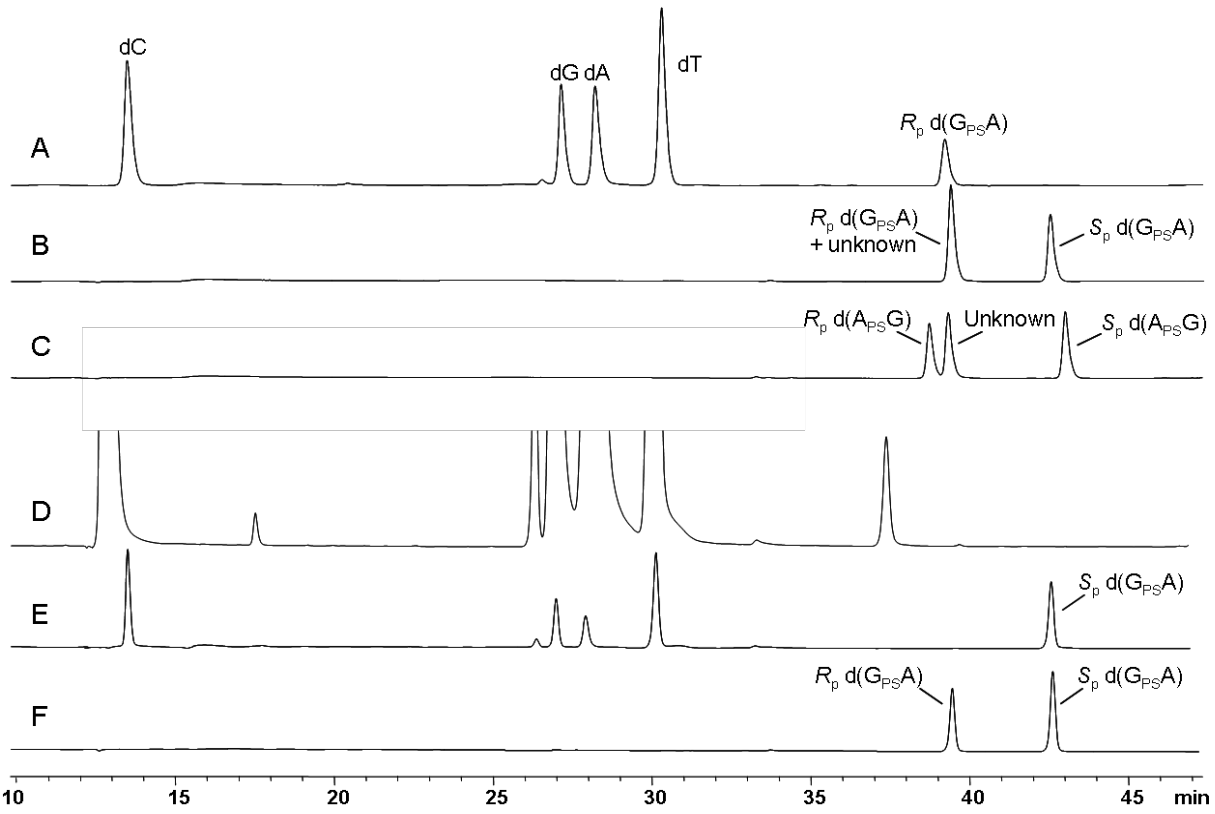
SUPPLEMENTARY FIGURES



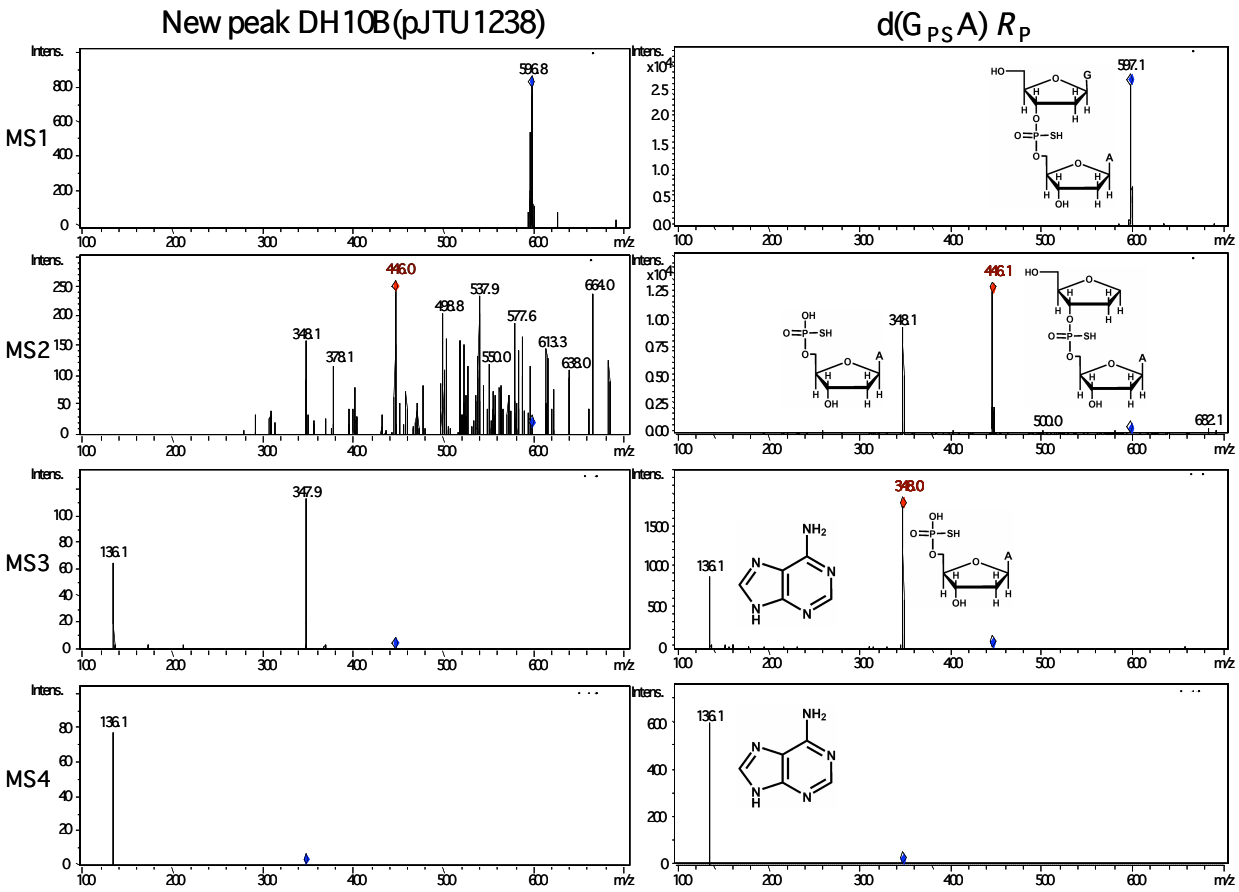
Supplementary Figure 1. Dnd phenotypes of B7A and DH10B(pJTU1238) and DH10B(SK$^+$). Genomic DNA from *E. coli* strains B7A (lane 2) and DH10B(pJTU1238) (lane 3) both undergo degradation (streaking in the lane) during pulsed field gel electrophoresis in Tris-containing buffer, while DNA from the DH10B(SK$^+$) strain lacking *dnd* genes remains intact (lane 4). Lane1: 1 kb DNA ladder.



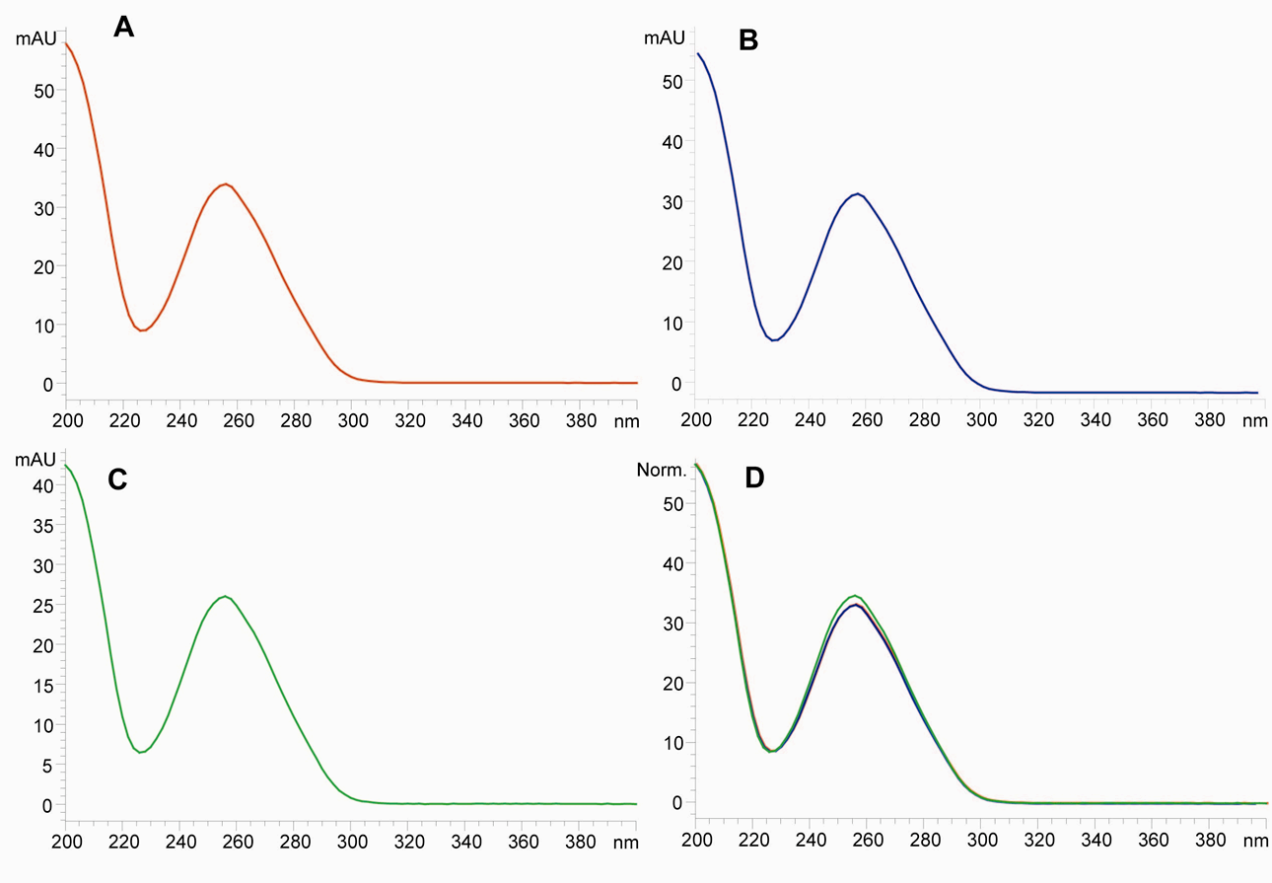
Supplementary Figure 2. ^{31}P NMR confirmation of the phosphorothioate stereochemistry in synthetic oligonucleotides d(G_{PS}A) (R_p) and d(G_{PS}A) (S_p). The values of d(G_{PS}A) (R_P) and d(G_{PS}A) (S_P) were determined to be 56.418 ppm and 55.355 ppm, respectively, which is consistent with published studies.[1,2]



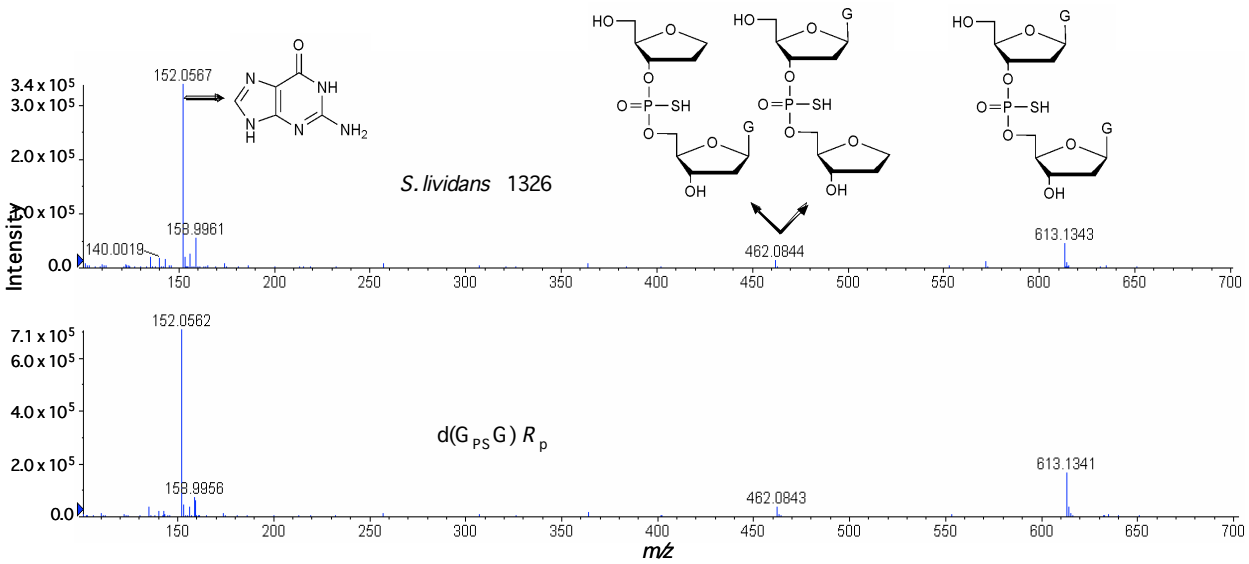
Supplementary Figure 3. HPLC corroboration of phosphorothioate stereochemistry using synthetic oligodeoxynucleotides. Further confirmation of the d(G_{PS}A) structure of the putative phosphorothioate-containing dinucleotide and assignment of phosphorothioate stereochemistry were achieved using synthetic oligodeoxynucleotides resolved by HPLC. **(A)** d(TTG_{PS}ACC) was digested with nuclease P1, snake venom phosphodiesterase and alkaline phosphatase and resolved by HPLC. **(B, C)** The putative d(G_{PS}A) (unknown) was mixed with synthetic R_p and S_p isomers of d(G_{PS}A) and d(A_{PS}G) and the mixtures resolved by HPLC. **(D)** DNA from *E. coli* strain DH10B(pJTU1238) subjected to exhaustive digestion with snake venom phosphodiesterase and alkaline phosphatase. **(E)** d(TTG_{PS}ACC) containing a mixture of R_p and S_p phosphorothioate bonds was subjected to exhaustive digestion with snake venom phosphodiesterase (specific to R_p; does not cut S_p) and alkaline phosphatase. **(F)** Mixture of d(G_{PS}A) (R_p) and d(G_{PS}A) (S_p) as reference standards.



Supplementary Figure 4. Mass fragmentation of putative d(G_{PS}A) (R_p) from DH10B(pJTU1238) in comparison with synthetic d(G_{PS}A) (R_p). The HPLC fraction containing the putative d(G_{PS}A) (R_p) isolated from *E. coli* strain DH10B(pJTU1238) and the synthetic d(G_{PS}A) (R_p) were analyzed by mass fragmentation. The putative and synthetic d(G_{PS}A) (R_p) (M+H *m/z* 597) were both sequentially fragmented to characteristic ions with *m/z* 446, 348 and 136. Corresponding fragments are shown adjacent to each ion. Identical results were obtained with *E. coli* strain B7A.



Supplementary Figure 5. UV spectra of putative d(G_{PS}A) (R_p) from B7A and DH10B(pJTU1238) in comparison with synthetic d(G_{PS}A) (R_p). According to individual (A, B and C) and overlaid (D) UV spectra, the putative phosphorothioate-containing peaks from B7A (A) and DH10B(pJTU1238) (B) have UV spectral properties identical to the synthesized d(G_{PS}A) (R_p) (C) with $_{max}$ 256 nm. This is similar to d(GA) with normal phosphodiester linkage and the absence of absorbance at ~340 nm excludes sulphur modifications of the base, such as 6-thioinosine and 6-thioguanosine.



Supplementary Figure 6. Analysis of the putative d(G_{PS}G) species isolated from *Streptomyces lividans* 1326. Purified putative d(G_{PS}G) isolated by HPLC resolution of nuclease-digested genomic DNA from *Streptomyces lividans* 1326 and synthetic d(G_{PS}G) (R_p) displayed the same characteristic *m/z* of 613.134 (calculated *m/z* 613.134) and accompanying fragment ions at *m/z* 462.084 and 152.056. Corresponding fragments were shown on the top of each ion (*m/z* 462.084 could be loss of either guanine). No species corresponding to d(G_{PS}G) (R_p) was observed in the *dnd* gene cluster deletion mutant ZX1.

SUPPLEMENTARY TABLE 1

Scintillation counting of ^{35}S in HPLC fractions from digested genomic DNA.

	Radioactivity (cpm) in HPLC fractions		
Bacterial strains	***39-40 min***	***46-47 min***	***47-48 min***
E. coli B7A	539	38	105
DH10B(pJTU1238)	377	65	65
DH10B(SK^+)	21	16	22
S. lividans 1326	383	56	129
ZX1	29	73	59

The highest levels of radioactivity were located in the 39-40 min fraction from *E. coli* strains DH10B(pJTU1238) and B7A, and from *Streptomyces lividans* 1326; background levels of radioactive material were isolated from strain DH10B(SK^+). Low levels of radioactivity were detected in the 7-8 min fraction from digested samples of all five strains. The negative control ZX1 showed a weak radioactive signal between 46 and 48 min.

SUPPLEMENTARY DATA

Characterization of stereochemistry of phosphorothioate bond. An oligodeoxynucleotide d($TTG_{PS}ACC$) bearing a racemic phosphorothioate bond was hydrolyzed with nuclease P1, snake venom phosphodiesterase and alkaline phosphatase as employed with the genomic DNA. The resulting nucleosides and nucleotides were analyzed by LC-MS. As shown in Supplementary Fig. 3A, the hydrolysate contained four canonical 2-deoxynucleoside peaks (dC, dG, dA, and dT) and a single d($G_{PS}A$) (R_p) peak with *m/z* 597, identical to the results from genomic DNA from B7A and DH10B(pJTU1238). The absence of a signal for the S_p configuration of the isomeric phosphorothioate mixture in the synthetic oligodeoxynucleotide was determined to be due to the greater sensitivity of the S_p isomer to hydrolysis by nuclease P1 compared to the R_p isomer.[3] In contrast, snake venom phosphodiesterase selectively hydrolyzes the R_p but not the S_p isomer,[4] as shown in Supplementary Fig. 3E for the d($TTG_{PS}ACC$) bearing a racemic phosphorothioate bond. We noted that, under the buffer conditions employed for the hydrolysis of DNA with nuclease P1,

snake venom phosphodiesterase and alkaline phosphatase,[5] the activity of the snake venom phosphodiesterase was fortuitously inhibited such that, while the R_p phosphorothioate bond was hydrolyzed to a small extent, the R_p isomer in $G_{PS}A$-containing sequences was preserved (Fig. 1 and Supplementary Fig. 3A). When samples of DH10B(pJTU1238) and B7A genomic DNA were hydrolyzed with only snake venom phosphodiesterase, which reacts with the R_p but not the S_p isomer, and alkaline phosphatase under the two-enzyme conditions (see Supplementary METHODS), we could not detect either the R_p or S_p isomers of d($G_{PS}A$) by LC-MS (Supplementary Fig. 3D). On the other hand, only the d($G_{PS}A$) (R_p) was observed in genomic DNA digested with nuclease P1, which is active against the S_p but not the R_p isomer,[6] and alkaline phosphatase. In aggregate, these results suggest that phosphorothioation of DNA by the *dnd* modification system is at least stereo-selective for the R_p configuration of the phosphorothioate.

SUPPLEMENTARY METHODS

Bacterial strains, plasmids and oligodeoxynucleotides. *E. coli* strain B7A (O148:H28:CS6:LT^+:ST^+) harboring a set of *dnd* homologous genes (*dndA,* GenBank accession number ZP 00714230; *dndB,* ZP 00714765; *dndC,* ZP 00714764; and *dndD,* ZP 00714763; *dndE,* 22596..22258 in NZ AAJT01000066) is an enterrotoxigenic *E. coli* strain[7]. *E. coli* DH10B strain (Invitrogen) was employed as a host for plasmids pJTU1238 and pBluescript II SK^+. *Streptomyces lividans* ZX1 is a derivative of the wild-type 1326 with deletion of a ~93-kb genomic island carrying the *dnd* gene cluster.[8] Recombinant plasmid pJTU1238 is a derivative of pBluescript II SK^+ carrying the *dnd* gene cluster from *Salmonella enterica* serovar Cerro 87 (T. Xu, unpublished data). Standard R_p and S_p phosphorothioate-modified d($G_{PS}A$), d($A_{PS}G$), d($G_{PS}G$), and unmodified d(GA) were chemically synthesized and purified by IBA BioTagnology. The configuration of phosphorothioate-modified d($G_{PS}A$), d($A_{PS}G$), d($G_{PS}G$) were all characterized by digestion with nuclease P1 (specific digestion of S_p) or snake venom phosphodiesterase (specific digestion of R_p). Furthermore, ^{31}P NMR studies confirmed the configuration of d($G_{PS}A$) (R_p) and d($G_{PS}A$) (S_p) (Supplementary Fig. 2). The 6-mer

oligodeoxynucleotide d($TTG_{PS}ACC$) with a mixture of R_p and S_p phosphorothioates was synthesized by Integrated DNA Technologies.

^{35}S incorporation experiments. Luria-Bertani medium (100 mL) supplemented with 1 mCi of L-[^{35}S]-cysteine (American Radiolabeled Chemicals, Inc.) was inoculated with 1 mL of an overnight culture of B7A, DH10B(pJTU1238), DH10B(SK^+). The cells were grown overnight at 37 °C and harvested by centrifugation with three washes with phosphate-buffered saline (PBS; 10 mM potassium phosphate, 137 mM NaCl, pH 7.4) and RNA-free genomic DNA was isolated using the QIAGEN Genomic-tip 500/G kit (QIAGEN). For *Streptomyces lividans* 1326 and ZX1, DNA were isolated from mycelia after growth in 100 mL of minimal medium[9] supplemented with 1 mCi of L-[^{35}S]-cysteine for 3 d at 30 °C.

Enzymatic hydrolysis of DNA. Samples (50 μg in 200 μL) of isolated DNA or oligodeoxynucleotide were hydrolyzed with 4 U nuclease P1 (USBiological) in 0.3 M sodium acetate, pH 5.6, 0.5 mM $ZnCl_2$, 0.5 mM deferoxamine mesylate (Sigma) at 37 °C for 2 h, as described in previous studies of DNA damage analysis.[5] Following adjustment of pH (200 μL of 30 mM sodium acetate, pH 8.1), 40 U of alkaline phosphatase (New England Biolabs) and 2 U of snake venom phosphodiesterase (*Crotalus adamanteus*, USB) were added and the mixture was incubated at 37 °C for 6 h before centrifugal ultrafiltration using Microcon YM-10 (Millipore). We discovered that, in this three-enzyme system, the activity of snake venom phosphodiesterase against phosphorothioate linkages was largely inhibited, which fortuitously preserved the R_P phosphorothioate-containing dinucleotide product present in DNA and oligonucleotide samples as confirmed by digestion of d($TTG_{PS}ACC$) and d($G_{PS}A$) (R_p). However, the activity of snake venom phosphodiesterase against normal phosphodiester linkages was largely preserved and compensated for the incomplete hydrolysis provided by nuclease P1 alone. For studies involving hydrolysis with only snake venom phosphodiesterase and alkaline phosphatase, hydrolysis of genomic DNA (50 μg) was carried out at 37 °C in 100 mM Tris-HCl, pH 8.75, 2 mM $MgCl_2$, 3 U snake venom phosphodiesterase and 40 U alkaline phosphatase in a total volume of 100 μL. After 6 h of incubation, the mixture was analyzed by LC-MS following the centrifugal

ultrafiltration. In this condition, the snake venom phosphodiesterase selectively digests the R_p but not S_p configuration.[4]

High-performance liquid chromatography (HPLC) analysis of DNA hydrolysates. Aliquots of the hydrolyzed DNA samples (50 µg) were loaded onto a Phenomenex C18 reversed phase column (250 x 4.6 mm, Synergi 4µ Hydro-RP 80A). Elution was carried out using two-step program starting with 99% buffer A (0.1% acetic acid in water) and 1% buffer B (0.1% acetic acid in acetonitrile) to 13% buffer B over 35.5 min and with 13% buffer B to 30% buffer B at a flow rate of 0.4 mL/min during another 20 min. For ^{35}S-labeled samples (corresponding to 20 µg of genomic DNA), HPLC fractions were collected at one-minute intervals and radioactivity was quantified on a Beckman LS6000SC scintillation counter.

Liquid chromatography mass spectrometry (LC-MS). LC-MS was performed on an Agilent LC/MSD Trap XCT Ultra mass spectrometer with an electrospray ionization source in positive ion mode. The LC was operated using the same column, buffer system and gradient as described earlier. Drying gas flow: 13 L/min; nebulizer pressure: 30 psi; drying gas temperature: 350 °C; capillary voltage: 3,200 V. High resolution mass spectrometry was performed on an Agilent G1969A LC-MSD TOF system with an electrospray ionization source in positive ion mode; drying gas flow, 12 L/min; nebulizer pressure, 25 psi; drying gas temperature, 325 °C; capillary voltage, 3,200 V. The samples were injected onto an Agilent ZORBAX C18 column (100 x 2.1 mm, 3.5 µm), with elution performed starting with 95% buffer A (0.25% acetic acid in water) and 5% buffer B (95% acetonitrile containing 0.25% acetic acid) to 30% buffer B over 16 min at a flow rate of 0.2 mL/min.

Pulse-field gel electrophoresis (PFGE). PFGE was performed in a contour-clamped homogeneous electric field system (Bio-Rad).[10] A 0.8% agarose gel was run in an electrophoresis buffer of 0.5 x TBE (50 mM Tris-borate buffer, pH 8.0, 0.1 mM EDTA) at 6 V/cm and 14 °C with switch times of 1-6 s for 9 h. The gel was stained with 0.5 mg/mL ethidium bromide for 30 min before digital analysis (GeneGenius).

***In vitro* Tris-dependent DNA cleavage assay.** An *in vitro* Tris-dependent DNA cleavage assay was carried out as described by Ray *et al*.[11] Briefly, 250 mL of Tris buffer (40 mM Tris, 20 mM sodium acetate, 0.8 mM EDTA, adjusted to pH 7.5 with acetic acid) maintained at 37 °C was activated for DNA cleavage in a gel chamber by applying a constant voltage of 80 V for 20 min. A 500 µL aliquot of the buffer was removed ~5 mm from the anode and was added to DNA samples with incubation at 37 °C for 2 h. The DNA samples were then dried under vacuum and dissolved in water prior to LC-MS analysis.

Mass Fragmentation. Mass fragmentation was performed on an Agilent 1100 series LC/MSD Trap XCT Ultra system in positive ion mode, with drying gas flow, 5 L/min; nebulizer pressure, 15 psi; drying gas temperature, 325 °C. The samples were injected onto an Agilent Eclipse XDB-C18 column (50 x 1.0 mm, 3.5 µm), with elution performed starting with 95% buffer A (0.25% acetic acid in water) and 5% buffer B (95% acetonitrile containing 0.25% acetic acid) to 80% buffer B over 15 min at a flow rate of 10 µL/min. The fragmentation amplitude was varied between 1.0 and 1.8 V.

REFERENCES

1. Bartlett, P.A. & Eckstein, F. Stereochemical course of polymerization catalyzed by avian myeloblastosis virus reverse transcriptase. *J. Biol. Chem.* **257**, 8879-8884 (1982).
2. Romaniuk, P.J. & Eckstein, F. A study of the mechanism of T4 DNA polymerase with diastereomeric phosphorothioate analogues of deoxyadenosine triphosphate. *J. Biol. Chem.* **257**, 7684-7688 (1982).
3. Potter, B.V. & Eckstein, F. Cleavage of phosphorothioate-substituted DNA by restriction endonucleases. *J. Biol. Chem.* **259**, 14243-14248 (1984).
4. Burgers, P.M. & Eckstein, F. Diastereomers of 5'-O-adenosyl 3'-O-uridyl phosphorothioate: chemical synthesis and enzymatic properties. *Biochemistry* **18**, 592-596 (1979).
5. Pang, B. et al. Lipid peroxidation dominates the chemistry of DNA adduct formation in a mouse model of inflammation. *Carcinogenesis* **28**, 1807-1813 (2007).

6. Potter, B.V., Romaniuk, P.J. & Eckstein, F. Stereochemical course of DNA hydrolysis by nuclease S1. *J. Biol. Chem.* **258**, 1758-1760 (1983).
7. DuPont, H.L. et al. Pathogenesis of Escherichia coli diarrhea. *N. Engl. J. Med..* **285**, 1-9 (1971).
8. Zhou, X. et al. Streptomyces coelicolor A3(2) lacks a genomic island present in the chromosome of Streptomyces lividans 66. *Appl. Environ. Microbiol.* **70**, 7110-7118 (2004).
9. Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F. & Hopwood, D.A. *Practical Streptomyces Genetics*, (John Innes Foundation, Norwich, 2000).
10. Chu, G., Vollrath, D. & Davis, R.W. Separation of large DNA molecules by contour-clamped homogeneous electric fields. *Science* **234**, 1582-1585 (1986).
11. Ray, T., Weaden, J. & Dyson, P. Tris-dependent site-specific cleavage of Streptomyces lividans DNA. *FEMS Microbiol. Lett.* **75**, 247-252 (1992).