Published online 5 September 2011 *Nucleic Acids Research, 2011, Vol. 39, No. 22* ***9803–9819***
doi:10.1093/nar/gkr660

Structural and functional insight into the mechanism of an alkaline exonuclease from *Laribacter hongkongensis*

Wen Yang[1,2], Wen-yang Chen[3], Hui Wang[1,4], John W. S. Ho[5], Jian-Dong Huang[6], Patrick C. Y. Woo[7], Susanna K.P. Lau[7], Kwok-Yung Yuen[7], Qionglin Zhang[1,2], Weihong Zhou[1,2], Mark Bartlam[1,2,*], Rory M. Watt[3,*] and Zihe Rao[1,2,4]

[1]State Key Laboratory of Medicinal Chemical Biology, [2]College of Life Sciences, Nankai University, Tianjin 300071, [3]Oral Biosciences, Faculty of Dentistry, The University of Hong Kong, Prince Philip Dental Hospital, Sai Ying Pun, Hong Kong SAR, [4]Laboratory of Structural Biology, Tsinghua University, Beijing, 100084, [5]Department of Biochemistry, The Chinese University of Hong Kong, Shatin N.T., Hong Kong SAR, [6]Department of Biochemistry, Li Ka Shing Faculty of Medicine and [7]Department of Microbiology, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pok Fu Lam, Hong Kong SAR, China

Received May 12, 2011; Revised July 26, 2011; Accepted July 27, 2011

ABSTRACT

Alkaline exonuclease and single-strand DNA (ssDNA) annealing proteins (SSAPs) are key components of DNA recombination and repair systems within many prokaryotes, bacteriophages and virus-like genetic elements. The recently sequenced β-proteobacterium *Laribacter hongkongensis* (strain HLHK9) encodes putative homologs of alkaline exonuclease (LHK-Exo) and SSAP (LHK-Bet) proteins on its 3.17 Mb genome. Here, we report the biophysical, biochemical and structural characterization of recombinant LHK-Exo protein. LHK-Exo digests linear double-stranded DNA molecules from their 5′-termini in a highly processive manner. Exonuclease activities are optimum at pH 8.2 and essentially require Mg^{2+} or Mn^{2+} ions. 5′-phosphorylated DNA substrates are preferred over dephosphorylated ones. The crystal structure of LHK-Exo was resolved to 1.9 Å, revealing a 'doughnut-shaped' toroidal trimeric arrangement with a central tapered channel, analogous to that of λ-exonuclease (Exo) from bacteriophage-λ. Active sites containing two bound Mg^{2+} ions on each of the three monomers were located in clefts exposed to this central channel. Crystal structures of LHK-Exo in complex with dAMP and ssDNA were determined to elucidate the structural basis for substrate recognition and binding. Through structure-guided mutational analysis, we discuss the roles played by various active site residues. A conserved two metal ion catalytic mechanism is proposed for this class of alkaline exonucleases.

INTRODUCTION

The accurate repair of broken or damaged double-stranded DNA (dsDNA) molecules by homologous recombination (HR) is a process common to all living organisms (1,2). The process of HR involves genetic exchange (crossover) between two dsDNA molecules via stretches of DNA that share significant sequence homology. In prokaryotes, it plays pivotal roles in chromosome repair, replication and segregation (3,4). HR also promotes horizontal gene transfer between prokaryotes and resident plasmids, bacteriophages or assorted mobile genetic elements (5). Consequently, many prokaryotes have acquired HR-promoting proteins of viral/phage origin in addition to their 'endogenous' DNA recombinational-repair machinery (6).

Two-component HR-promoting proteins of viral/phage origin, comprising an alkaline exonuclease and a single-stranded DNA (ssDNA) annealing protein [SSAP, also referred to as a synaptase (7)] are widely distributed throughout the prokaryotic world (6,8–10). In these systems, the alkaline exonuclease protein fully digests one strand, or partially digests both strands of the linearized dsDNA substrate; generating ssDNA molecules, or long ssDNA 'tails' several hundred nucleotides

*To whom correspondence should be addressed. Tel/Fax: +86 22 23502351; Email: bartlam@nankai.edu.cn
Correspondence may also be addressed to Rory M. Watt. Tel: +852 2589 0482; Fax: +852 2547 6133; Email: rmwatt@hku.hk

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

in length. These are 'coated' with the partnering SSAP protein, which promotes their annealing with complementary regions of ssDNA on the chromosome or episome (11–13). This most likely occurs at the replication fork (14–16). DNA repair and replication proteins process the genetic intermediates formed, resulting in the creation of fully double-stranded recombinant DNA molecules (6,7,14). The best studied of the bacterial-based, two-component exonuclease and SSAP systems are the 'Red' proteins Exo (λ-exonuclease) and Bet from bacteriophage-λ (7,17); RecE and RecT from the *Escherichia coli Rac* prophage (7,12,18,19); G34.1P and G35P from bacteriophage SPP1 (13,20,21); gp60 and gp61 from Mycobacteriophage Che9c (22,23); and SXT-Exo (s066) and SXT-Bet (s065) from the SXT genetic element of *Vibrio cholerae* (24–26). Over recent years, various partnering pairs of exonuclease and SSAP proteins have been utilized in 'recombineering' procedures to create genetic alterations within a variety of prokaryotic organisms; including *E. coli*, *Salmonella*, *Pseudomonas* and *Mycobacteria* (22,27–29). In these systems, the exonuclease and SSAP proteins mediate HR-type events between PCR-generated dsDNA 'targeting cassettes' (encoding the desired genetic change) and the bacterial chromosome or resident episome (14).

First isolated and characterized in the 1960s (30,31) (and references therein), the activities and biophysical properties of the λ-exonuclease protein have been extensively studied. Phylogenetically related to exonucleases from the *Baculoviridae* and *Herpesviridae* families, it is the prototypical member of the alkaline exonuclease family (8), which belongs to the restriction endonuclease-like PD-(D/E)XK superfamily (32). Its activities are optimal at pH 9.2–9.5, and are dependent upon Mg^{2+} cations; although Mn^{2+} can substitute with a 3- to 4-fold reduction in catalytic efficiency (31). λ-exonuclease binds to the termini of linear dsDNA molecules, and processively digests the 5′-ended strand to produce long 3′-ssDNA tails, with the concomitant liberation of 5′-mononucleotides. It has previously been estimated that an average of 3000 (33) or 18 000 (34) nucleotides are digested without dissociation of the enzyme from the undigested strand. It has a strong preference for dsDNA molecules containing 5′-phosphorylated (5′-PO_4) ends, digesting them an estimated 2- to 100-fold more effectively than the corresponding dephosphorylated (5′-OH) substrates (35–37). Its average digestion rate has been estimated to be ∼4 nt/s (35), 9 nt/s (38), 15–20 nt/s (39), 32 nt/s (34) or ∼1000 nt/s (40), using a variety of bulk or single-molecule scale biophysical approaches.

The crystal structure of λ-exonuclease was reported in 1997 (41), revealing an unusual toroidal (doughnut-shaped) trimeric arrangement with a central tapered channel. This funnel-shaped channel had an inner diameter of ∼30 Å at one end, which is wide enough to accommodate a dsDNA substrate. However, it was only ∼15 Å wide at the narrow end, which is only capable of accommodating ssDNA. Each monomer contained a Mn^{2+} ion housed within a pocket lining the channel, which presumably equate to the active sites that normally accommodate catalytic Mg^{2+} ions. This led the authors to propose a mechanism whereby the dsDNA substrate enters at the wide opening, and the non-hydrolyzed 3′-strand exits at the opposite end.

The crystal structure of the catalytically active C-terminal domain of the functionally equivalent RecE exonuclease was recently reported (38). RecE formed an analogous toroidal tetrameric arrangement, with a central tapered channel that contained the catalytic residues. Although λ-exonuclease and RecE share little overall primary sequence similarity and have quite distinct multimeric arrangements, the central channels within these two proteins have a highly similar size and topology. This strongly suggests a conserved mode of action where the terminus of the dsDNA molecule enters at the wide end of the channel; the strands are separated, and the 5′-end is processively hydrolyzed as the protein encircles and translocates along the undigested 3′-strand (38,41).

As part of our ongoing efforts to characterize prokaryotic DNA recombination proteins, we identified two adjacent genes within a suspected prophage region on the genome of *Laribacter hongkongensis* HLHK9, which encode homologs of alkaline exonuclease (LHK-Exo, LHK_01497) and SSAP (LHK-Bet, LHK_01496) proteins (42). *Laribacter hongkongensis* is a recently identified species of β-proteobacteria that is a potential diarrheal pathogen (43,44). Here, we describe the functional and structural characterization of LHK-Exo, which shares 29% primary amino acid sequence identity with λ-exonuclease. The crystal structure of LHK-Exo was solved to 1.9 Å resolution, revealing a 'doughnut-like' toroidal structure, similar to that previously observed for λ-exonuclease (41). LHK-Exo functioned as a DNA exonuclease, digesting linear dsDNA with a clear preference for 5′-phosphorylated ends. Under the guidance of the structure of LHK-Exo in complex with DNA, a set of active site point mutants were constructed and functionally characterized to probe sequence–structure relationships. We compare the activities and properties of LHK-Exo with those of related nucleases; and discuss mechanisms for substrate recognition and binding, as well as a potential catalytic mechanism for this class of alkaline exonucleases.

MATERIALS AND METHODS

Materials

All buffer salts, metal salts, chemicals and biochemicals were purchased from Sigma-Aldrich or GE Biosciences. PicoGreen reagent was from Invitrogen (cat.# P7589). Unmodified and phosphorothioate modified oligonucleotides were purchased from TechDragon Ltd. (Shatin, Hong Kong). Where applicable, oligonucleotides were 5′-phosphorylated using T4 polynucleotide kinase (NEB), according to the manufacturer's instructions, then desalted on Auto Seq G50 columns (GE Biosciences). All restriction enzymes were purchased from New England Biolabs (NEB; Ipswich, MA, USA). pET28a and pET32a plasmid DNA were from Novagen; pUC18 was from Stratagene and pMAL-c2 was purchased

from New England Biolabs. Restricted (linearized) dsDNA was routinely purified using a Qiaquick Gel Extraction kit (Qiagen GmbH). When applicable, linearized dsDNA was 5′-dephosphorylated using calf intestinal phosphatase (NEB) then gel purified. dsDNA was resolved by electrophoresis on agarose/Tris–acetate EDTA (TAE) gels and stained with ethidium bromide (Sigma). ssDNA was analyzed by denaturing polyacrylamide gel electrophoresis on 7 M urea–TBE polyacrylamide (12%, 37.5:1) gels, which were stained using SYBR Gold Nucleic Acid Gel Stain (Invitrogen). Gels were visualized by UV-transillumination on a ChemiDoc XRS molecular imaging system equipped with Quantity One v4.6.6 software (BioRad). A 1 kb Plus DNA Ladder (Invitrogen) was used as a marker for dsDNA, and Oligo Length Standards 20/100 Ladder (Integrated DNA Technologies Inc., USA) was used for ssDNA gels. Graphs were prepared using SigmaPlot 10.0 (Systat Software, Inc).

Methods

Cloning, expression and purification of LHK-Exo. The *LHK-exo* gene (LHK_01497; GenBank: NC_012559.1, 1455161..1455772) was PCR amplified from *Laribacter hongkongensis* HLHK9 genomic DNA (Expand Polymerase, Roche) using the LHKexofor1 (ATATCCATGGAACAACGCAC) and LHKexorev2 (TTAAAAGCTTGGCGGCCTTTCGTTTCAG) primers. The amplified gene product was ligated via NcoI/HindIII (restriction sites underlined) into pET28a (Novagen) to create plasmid pEXNH1, which encodes a 13aa C-terminal hexa-histidine fusion (...*KLAAALEHHHHHH**). *Escherichia coli* BL21 (DE3) containing pEXNH1 was grown in Luria-Bertani (LB) medium (2 l) containing 50 μg/ml kanamycin (37°C), to an OD_{600} of 0.6–0.8, induced with IPTG (isopropyl-β-D-thiogalactopyranoside, 0.2 mM), then incubated at 25°C for 10 h. Bacterial cells were harvested by centrifugation (6 000 *g*, 30 min), washed with PBS buffer, then resuspended in PBS buffer containing protease inhibitors (Complete, Roche) and lysed by sonication (ice cooling). The supernatant obtained after centrifugation (15 000 *g*, 30 min) was filtered (0.45 μm syringe filter, Millipore) then applied to a Ni^{2+}-charged 5 ml Hitrap Chelating FF column (GE Healthcare) pre-equilibrated with PBS buffer. The column was washed (PBS buffer containing 10 mM imidazole, 20 ml) and the protein was eluted from the column using a linear gradient of imidazole (25–500 mM) in PBS buffer. Fractions containing the purest protein were pooled and concentrated (Amicon Ultra, Millipore) and buffer was exchanged using a G-25 Sephadex desalting column (GE Healthcare) pre-equilibrated with buffer A (20 mM Tris–HCl pH 8.0, 1 mM DTT). Protein was then purified by anion exchange chromatography (Resource Q, GE Healthcare), eluting with a linear gradient of 0–1 M NaCl in buffer A. Fractions containing the purest recombinant LHK-Exo protein were pooled then subjected to gel filtration chromatography (Superdex-200 HR, GE Healthcare) using buffer B (20 mM Tris pH 8.0, 150 mM NaCl, 1 mM DTT). The purity of resultant LHK-Exo was greater than ~98% as determined by SDS–PAGE analysis. Single residue mutants of LHK-Exo were constructed using the QuikChange Site-Directed Mutageneis kit (Stratagene) according to the manufacturer's instructions; using plasmid pEXNH1 as the template and the sets of oligonucleotides listed in Supplementary Table S1. Mutant proteins were expressed and purified as described above.

Qualitative determination of LHK-Exo activities

Mode of digestion. BamHI-linearized pET28a (1.8 μg, 0.54 pmol) was incubated with LHK-Exo protein (30 μg, 0.41 nmol of trimers) in 200 μl of Tris–HCl (pH 8.0, 50 mM), 50 mM NaCl, 7.5 mM $MgCl_2$ at 37°C. Aliquots (20 μl) were withdrawn at various time points (indicated in the text) and analyzed on a 1% agarose/TAE gel.

Polarity of digestion. The 712-bp dsDNA substrates ('unmodified') were synthesized by PCR using pET32a as template, with two different pairs of primers, one pair of T7 (TAATACGACTCACTATAGGG) and T7rev (GCTAGTTATTGCTCAGCGG) for amplification of 'unmodified' substrate, and another pair of the phosphorothioate modified primers T7forPT3 (TAATACGACTCACTATA$_S$G$_S$G$_S$G) and T7revPT3 (GCTAGTTATTGCTCAG$_S$C$_S$G$_S$G) for 'PT-modified' substrate (where subscript 'S' denotes a nuclease-resistant phosphorothioate linkage between 2 nt). Both PCR-amplified substrates were 5′-phosphorylated (T4 polynucleotide kinase, NEB) and purified (QIAquick PCR purification kit, QIAgen) prior to use. The two dsDNA substrates (0.1 μg, 0.23 pmol) were separately incubated with 6 μg of lambda-Exo (74 pmol of trimers) or LHK-Exo (82 pmol of trimers) in Tris–HCl (25 mM, pH 8.0), 7.5 mM $MgCl_2$, 1 mM DTT (total volume 40 μl) at 37°C. Aliquots (20 μl) were quenched (20 mM EDTA + 1% SDS) immediately and after 20 min, then analyzed on 1% agarose TAE gels.

Single strand DNA exonuclease activities. Phosphorylated [5′-PO_4-$(dT)_{50}$] or non-phosphorylated [5′-OH-$(dT)_{50}$] forms of a 50-mer or oligothymidine (0.4 nmol) were incubated with LHK-Exo protein (4.5 μg, 61.4 pmol of trimers) in 25 mM Tris–HCl, pH 8.0; 7.5 mM $MgCl_2$; 1 mM DTT at 37°C; total volume 80 μl. Aliquots (20 μl) were removed at 0, 0.5, 5 and 20 min, and immediately quenched (30% glycerol, 7 M urea-TBE; 80 μl), prior to analysis by denaturing polyacrylamide gel electrophoresis.

Quantitative determination of DNA exonuclease activities. Quenched PicoGreen DNA fluorescence assays were performed essentially as described previously (21,26,45) with minor modifications. Typical assays (60 μl) contained LHK-Exo protein (8 ng, 0.11 pmol of trimers) and 30 ng linear pUC18 DNA (0.018 pmol) or EcoRV-linearized pMAL-c2 (0.0073 pmol) in nuclease assay buffer [Tris–HCl (25 mM, pH 8.0), 7.5 mM $MgCl_2$, 1 mM DTT]. Assays were initiated by the addition of linear dsDNA substrate, incubated at 25°C for 20 min before being quenched by the addition of EDTA to 20 mM. The 250-fold diluted PicoGreen reagent

(Invitrogen) was added to quenched reaction mixtures and fluorescence levels (which are proportional to the quantities of dsDNA present) were measured at an excitation 485 nm/emission 535 nm (Perkin Elmer 1420 multi-label counter).

Determination of optimal metal ion concentrations, pH and temperature for DNA exonuclease activities. Quenched PicoGreen fluorescence assays were performed, quenched and analyzed as described above with minor modifications.

Determination of optimal Mg^{2+} and Mn^{2+} ion concentrations. Assays (60 µl) contained LHK-Exo protein (8 ng, 0.11 pmol of trimers), PstI-linearized pUC18 (30 ng, 0.018 pmol) in Tris–HCl (25 mM, pH 8.0), $MgCl_2$ (1–15 mM) or $MnCl_2$ (0.1–1.7 mM) at the concentrations indicated in the text; and were incubated at 25°C for 20 min.

Determination of optimal pH. Assays (60 µl) contained LHK-Exo protein (8 ng, 0.11 pmol of trimers), PstI-linearized pUC18 (30 ng, 0.018 pmol), 25 mM Tris–HCl, 7.5 mM $MgCl_2$; adjusted to the appropriate pH value (pH 7.0–9.0); and were incubated at 25°C for 20 min.

Determination of optimal temperature. Assays (60 µl) contained LHK-Exo protein (40 ng, 0.55 pmol of trimers) and PstI-linearized pUC18 (30 ng, 0.018 pmol) in nuclease assay buffer. All solutions were pre-equilibrated at the temperature indicated in the text (34–54°C) using a thermostat-regulated water bath; an incubation time of 1 min was used. Four to six independent replicates of each experimental condition were performed. Data were fitted to hyperbolae using SigmaPlot 10.0 to calculate the k_{cat} during the initial phase or reaction (digestion rate stated in nucleotides per second; per trimer of LHK-Exo protein), reported as the mean ± standard deviation.

Determination of digestion processivity. LHK-Exo (6 µg, 82 pmol of trimers) protein and 60 ng (0.015 pmol) of linear dsDNA substrate (5′-PO_4-dsDNA, EcoRV-linearized pMal-c2; or 5′-OH-dsDNA, 5′-dephosphorylated EcoRV-linearized pMal-c2) were equilibrated in 180 µl of Tris–HCl (25 mM, pH 8.0), 1 mM DTT, at 25°C for 5 min. A total of 90 µl of Tris–HCl (25 mM, pH 8.0), 1 mM DTT 15 mM $MgCl_2$ was added to initiate the reaction; then exactly 15 s later, excess heparin (Sigma; 360 µg, 20 nmol) in 90 µl of Tris–HCl (25 mM, pH 8.0), 1 mM DTT, 7.5 mM $MgCl_2$ was added to sequester unbound LHK-Exo (total reaction volume 360 µl). Assays were incubated at 25°C for 20 min, removing 60 µl aliquots at the times indicated in the text to determine dsDNA levels via fluorescent PicoGreen assays (as described above). Four to six independent replicates of each experimental condition were performed. Plotted data were fitted to hyperbolae using SigmaPlot 10.0 to obtain the digestion processivity from the extrapolated curves as previously described (26,36), which is reported as the mean ± standard deviation.

Analysis of multimericity by analytical ultracentrifugation. Sedimentation velocity analytical ultracentrifugation experiments were performed using a Beckman Proteomelab^{TM} XL-1 analytical ultracentrifuge (60 000 rpm, 262,000 g, 293 K). LHK-Exo protein was prepared at a concentration of 1 mg/ml in 150 mM NaCl. Data were analyzed using the c(S) and c(M) models in the Sedfit program (46) to determine the sedimentation coefficient and apparent mass distributions, respectively.

Crystallization and X-ray data collection. The purified recombinant LHK-Exo protein was concentrated to 20 mg/ml in crystallization buffer (20 mM Tris–HCl, pH 8.0). Crystallization conditions were screened with the Hampton Research Screens kits. Crystallization was performed by the sitting-drop vapor-diffusion method at 293 K, with 1 µl protein solution mixed with 1 µl reservoir solution and equilibrated with 100 µl reservoir solution containing 0.1 M Tris–HCl pH 8.4 and 1.0 M potassium sodium tartrate tetrahydrate. Good quality crystals were obtained within 2 days, and were flash cooled and stored in liquid nitrogen for future data collection. The X-ray diffraction data were collected to 1.9 Å resolution at 100 K at beamline BL5A of the Photon Factory (Japan) using an ADSC-Q315 detector. Reservoir solution containing 20% glycerol was used as cryoprotectant. The diffraction data were processed with the HKL2000 program suite (47). The LHK-Exo crystals obtained belong to space group $P6_3$ with unit cell parameters of: $a = b = 108.9$ Å, $c = 47.6$ Å, $\alpha = \beta = 90°$ and $\gamma = 120°$.

To obtain substrate-bound complex structures, native LHK-Exo crystals were soaked at 293 K for 24 h in a solution containing dAMP or a 5′-phosphorylated pentamer of oligothymidine [5′-PO_4-$(dT)_5$] at a final concentration of 1 mM in 0.1 M Tris–HCl pH 8.4 and 1.0 M potassium sodium tartrate tetrahydrate. Immediately prior to data collection, crystals were soaked in cryoprotectant solution consisting of 20% glycerol and the reservoir solution. X-ray diffraction data were collected in-house at 100 K on a Rigaku R-AXIS IV^{++} image plate using Cu Kα radiation ($\lambda = 1.5418$ Å) from a Rigaku MicroMax-007 rotating anode X-ray generator operating at 40 kV and 20 mA. All diffraction data were integrated, scaled and merged using the HKL2000 suite (47). A complete summary of the data collection statistics is given in Table 1.

Structure determination and refinement. The structure of LHK-Exo was solved by molecular replacement (MR) with the monomeric λ-exonuclease (PDB code: 1AVQ) employed as a search model, using the Phaser program (48) in the CCP4 suite (49). Good quality electron density maps enabled initial model building by ARP/wARP (50). Further model building and manual adjustments were facilitated using the Coot program (51) and structure refinement was performed by Refmac5 (52). The stereochemical quality of the structure model was monitored with the PROCHECK program (53).

The substrate-bound complex structures were also determined by MR with the native enzyme employed as

Table 1. X-ray data collection and structure refinement statistics

Data set	LHK-Exo	LHK-Exo:dAMP	LHK-Exo:ssDNA
Data collection statistics			
Wavelength (Å)	1.0000	1.5418	1.5418
Space group	$P6_3$	$P6_3$	$P6_3$
Cell dimensions, $a/b/c$ (Å)	108.9/108.9/47.6	108.3/108.3/47.8	108.7/108.7/48.2
Resolution (Å)	50–1.9 (1.97–1.90)	50–2.6 (2.64–2.60)	50–2.8 (2.85–2.80)
Average I/σ (I)[a]	32.6 (2.0)	43.7 (4.3)	35.5 (3.4)
Completeness (%)[a]	95.1 (70.6)	99.5 (100.0)	99.9 (100.0)
Redundancy[a]	7.1 (5.3)	4.5 (4.5)	6.8 (6.5)
R_{merge} (%)[a,b]	8.4 (58.1)	4.7 (47.3)	6.7 (49.2)
Structure refinement statistics			
Resolution (Å)	50–1.9	50–2.6	50–2.8
Average B-factor (Å[b])	37.6	65.5	69.9
R_{work}/R_{free} (%)[c]	20.4/24.1	20.8/26.2	19.3/24.8
RMSD bond lengths (Å)	0.019	0.016	0.014
RMSD bond angles (°)	1.629	1.608	1.550
Ramachandran plot			
Most favored (%)	94.7	91.8	93.5
Allowed (%)	5.3	8.2	5.9
Generously allowed (%)	0	0	0.6
Disallowed (%)	0	0	0

[a]Values in parentheses correspond to the highest resolution shell.
[b]$R_{merge} = \Sigma_i |I_i - \langle I \rangle| / \Sigma \langle I \rangle$, where I_i is an individual intensity measurement and $\langle I \rangle$ is the average intensity for all the reflections.
[c]$R_{work}/R_{free} = \Sigma ||F_o| - |F_c|| / \Sigma |F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively.

a search model. The dAMP and DNA molecules were manually built and adjusted under the guidance of $F_o - F_c$ difference maps using the Coot program (51). Structural refinement was also carried out using Refmac5 (52). A summary of the structure refinement statistics is provided in Table 1.

RESULTS AND DISCUSSION

LHK-Exo is an alkaline exonuclease that digests dsDNA with strict 5′- to 3′-polarity

The LHK_01497 gene from the HLHK9 type strain of *L. hongkongensis* was cloned into the pET28a vector as a C-terminal hexahistidine fusion, and subsequently expressed to high levels in *E. coli*. Three-step purification by immobilized Ni^{2+} ion affinity, anion exchange and gel-filtration chromatography afforded multi-milligram quantities of LHK-Exo protein of high purity suitable for biochemical and structural characterization. Analogous to the phylogentically related λ-exonuclease and SPP1-Chu (G34.1P) proteins, recombinant LHK-Exo strictly required Mg^{2+} or Mn^{2+} ions for the hydrolysis of linear dsDNA molecules (Supplementary Figure S1) (20,21,31). Analysis of quenched aliquots from a time-course incubation of LHK-Exo with linear dsDNA by agarose gel electrophoresis, indicated that LHK-Exo was functioning as an exonuclease that did not release significant amounts of partially digested products (Figure 1A). This is consistent with the activities of λ-exonuclease (30,54), SPP1-Chu (20,21), RecE (55) and SXT-Exo (26), that all function as highly processive dsDNA exonucleases.

Modified dsDNA substrates (712 bp in length), containing three consecutive 'nuclease resistant' phosphorothioate linkages near the 5′-termini of both strands (PT modified) were synthesized using a PCR-based strategy, to probe the polarity of exonucleolytic digestion (Figure 1B). Analogous dsDNA substrates with no backbone modifications (unmodified) were similarly synthesized as controls. By substituting phosphodiester linkages with phosphorothioate linkages near the 5′-termini, the activities of a 5′- to 3′-exonuclease should be severely inhibited, whereas the activities of a 3′- to 5′-exonuclease should be relatively unaffected. We have previously used this approach to probe the *in vivo* DNA hydrolytic activities of the SXT-Exo protein (26). λ-exonuclease was included as a control, as this enzyme has previously been shown to hydrolyze DNA in the 5′- to 3′-direction. LHK-Exo degraded negligible amounts of the (5′-phosphorylated) 'PT-modified' dsDNA substrate (Figure 1B, lane 5), but fully digested the analogous 'unmodified' dsDNA substrate under the conditions tested (lane 3). λ-exonuclease exhibited analogous activities (lanes 6–9). These assays clearly demonstrated that LHK-Exo specifically digests dsDNA from the 5′-termini; i.e. it hydrolyzes dsDNA with strict 5′- to 3′-polarity.

We probed the ssDNA hydrolysis activities of LHK-Exo using 5′-phosphorylated and non-phosphorylated forms of a 50-mer of oligothymidine [5′-PO_4-$(dT)_{50}$ and 5′-OH-$(dT)_{50}$, respectively] as model substrates. LHK-Exo effectively hydrolyzed both 5′-phosphorylated and non-phosphorylated ssDNA substrates (Figure 1C and D); although the 5′-phosphorylated oligonucleotide was digested ~5- to 10-fold more rapidly. As was the case for dsDNA substrates, Mg^{2+} ions were essentially required for activity (Supplementary Figure S2). Notably, the hydrolysis of 5′-OH-$(dT)_{50}$ was strongly inhibited at high salt concentrations (1 M KCl), while the degradation of the

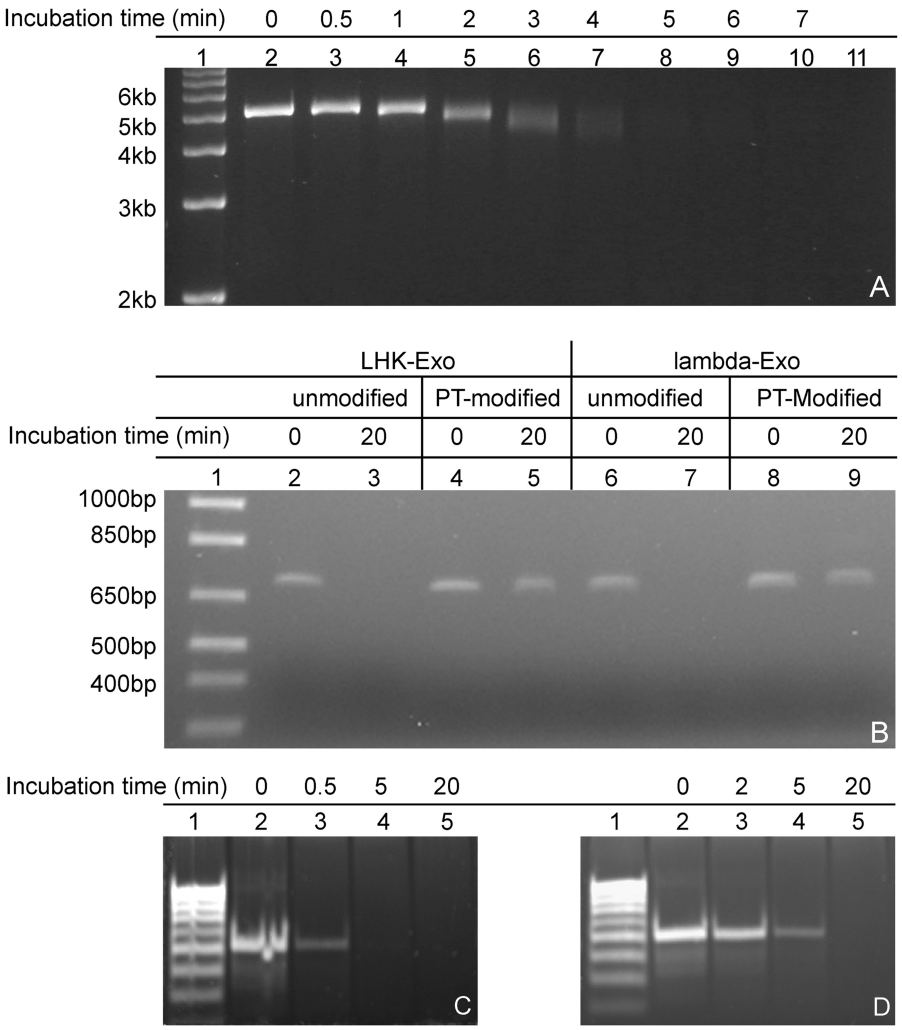

Figure 1. Qualitative analysis of ssDNA and dsDNA hydrolysis activities of LHK-Exo. (**A**) dsDNA exonuclease activities. Agarose gel showing aliquots taken (0–15 min) from an incubation of LHK-Exo (30 μg, 0.41 nmol of trimers) and BamHI-linearized pET28a (1.8 μg, 0.54 pmol) in Tris–HCl (pH 8.0, 50 mM), 50 mM NaCl, 7.5 mM $MgCl_2$ at 37°C. (**B**) Polarity of dsDNA digestion. A total of 6 μg of LHK-Exo (82 pmol of trimers, lanes 2–5) or λ-exonuclease (74 pmol of trimers, lanes 6–9) protein was incubated with 0.1 μg (0.23 pmol) of a 712-bp linear 5′-phosphorylated dsDNA substrate ('unmodified'; lanes 2, 3, 6 and 7), or an analogous 5′-phosphorylated linear dsDNA substrate containing three consecutive 'nuclease-resistant' phosphorothioate linkages at its 5′-termini ('PT-modified'; lanes 4, 5, 8, 9). Assays were quenched immediately (0 min) or incubated at 37°C for 20 min, before analysis of digestion products on 1% agarose gels. (**C**) Digestion of 5′-phosphorylated ssDNA. Reaction mixtures (80 μl) containing LHK-Exo (4.5 μg, 61.4 pmol of trimers) and 5′-PO_4-$(dT)_{50}$ (0.4 nmol) in 25 mM Tris–HCl (pH 8.0), 7.5 mM $MgCl_2$, 1 mM DTT were incubated at 37°C. Aliquots (20 μl) were withdrawn after 0, 0.5, 5 and 20 min, and immediately quenched. Reaction products were analyzed by denaturing gel electrophoresis. (**D**) Digestion of non-phosphorylated ssDNA. Analogous sets of assays were performed using non-phosphorylated 50-mers of oligothymidine [5′-OH-$(dT)_{50}$]. Fluorescent gel images were scanned after SYBR Gold staining. A ssDNA ladder [Oligo Length Standards 20/100 Ladder (IDT)] is included in lane 1.

5′-phosphorylated form was less affected (Supplementary Figure S2). This is consistent with 5′-phosphorylated DNA molecules binding to the LHK-Exo protein more effectively than the corresponding non-phosphorylated ones, and consequently being preferred as substrates (see below), as is the case for the λ-exonuclease, SXT-Exo and SPP1-Chu proteins (20,21,26,36).

The dsDNA exonuclease activities of LHK-Exo were investigated in a quantitative manner using the sensitive PicoGreen fluorescent DNA stain, as has previously been used in analogous studies (21,36,45). As may be seen in Figure 2, LHK-Exo activities were optimal in the presence of Mg^{2+} ions at a concentration of 7.5 mM (Figure 2A) or Mn^{2+} ions at a concentration of 1 mM (Figure 2B). This is roughly similar to concentrations of Mg^{2+} ions utilized by related exonucleases of viral/phage origin: λ-exonuclease [1–2.5 mM (31)]; BGLF5 from the Epstein-Barr virus [EBV; 20 mM (56); 5 mM (57)]; baculovirus alkaline exonuclease [4–8 mM (58)]; UL12 from Herpes simplex virus 1 [2–5 mM (59)]; V-TREX from *Choristoneura fumiferana* nucleopolyhedrovirus [5 mM (60)]. Equivalent levels of dsDNA digestion were observed (within experimental error) when Mg^{2+} or Mn^{2+} ions were present at their optimal concentrations; suggesting that LHK-Exo can accommodate and utilize either ion equally well. This is notably different to the metal ion selectivities of related

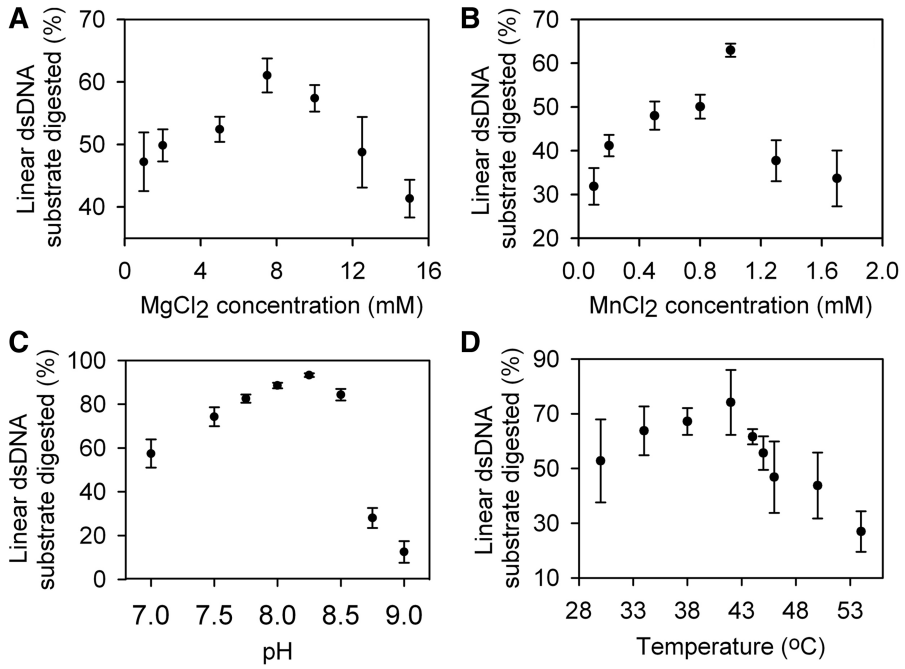

Figure 2. Determination of optimum conditions for LHK-Exo dsDNA digestion activities. PicoGreen fluorescence assays were performed to quantify the amounts of a representative 5′-phosphorylatd dsDNA substrate (PstI-linearized pUC18 that were digested by LHK-Exo under various conditions (reported as a percentage of the initial quantities of DNA). (**A** and **B**) Optimal concentrations of Mg^{2+} and Mn^{2+} ions. Amounts of PstI-linearized pUC18 (30 ng, 0.018 pmol) digested by LHK-Exo (8 ng, 0.11 pmol of trimers) in Tris–HCl (25 mM, pH 8.0) containing varying concentrations of Mg^{2+} ions (1–15 mM, A) or Mn^{2+} ions (0.1–1.7 mM, B), respectively, upon incubation at 25°C for 20 min. (**C**) Optimal pH. Amounts of PstI-linearized pUC18 (30 ng, 0.018 pmol) digested by LHK-Exo (8 ng, 0.11 pmol of trimers) in 25 mM Tris–HCl, 7.5 mM $MgCl_2$; adjusted to the appropriate pH value (pH 7.0–9.0); upon incubation at 25°C for 20 min. (**D**) Optimal temperature. Amounts of PstI-linearized pUC18 (30 ng, 0.018 pmol) digested by LHK-Exo (40 ng, 0.55 pmol of trimers) in Tris–HCl (25 mM, pH 8.0), 7.5 mM $MgCl_2$, after incubation for 1 min at the indicated temperature (34–54°C). Four to six independent replicates of each experimental condition were performed, and data are reported as the mean ± standard deviation.

exonucleases; where activities have been reported to be ~4- to 10-fold higher in the presence of optimal concentrations of Mg^{2+} ions, compared with optimal Mn^{2+} ion concentrations (31,57–59). dsDNA degradation levels were highest at pH 8.2 (Figure 2C), and at a temperature of 41°C (Figure 2D). The optimal pH value for LHK-Exo is comparable to those determined for related exonucleases: λ-exonuclease [pH 9.2–9.5 (31)]; SPP1-Chu [G34.1P; pH 9.2–9.5 (21)]; SXT-Exo [pH 8.2 (26)]; BGLF5 [pH 8.0 (56); pH 8.5–9.5 (57)]; baculovirus alkaline exonuclease [pH >9–10 (58)]; UL12 [pH 9–10 (59)]; V-TREX [pH 9.5 (60)]. When the pH and Mg^{2+} ion concentrations were optimal, LHK-Exo hydrolyzed 5′-phosphorylated linear dsDNA at a rate of 7.0 ± 1.8 nt/s (per LHK-Exo monomer, per DNA end; at 25°C; Table 2). This rate of hydrolysis is comparable to that previously determined for λ-exonuclease under a variety conditions (34,35,38,39).

Table 2. Rate and processivity of double strand DNA digestion mediated by wild-type LHK-Exo and mutants

Proteins	5′-PO_4 dsDNA		5′-OH dsDNA	
	k_{cat} (nt/s)	processivity (nt)	k_{cat} (nt/s)	processivity (nt)
WT	7.0 ± 1.8	2368 ± 72	1.6 ± 0.5	805 ± 56
R12A	1.0 ± 0.3	955 ± 38	1.1 ± 0.3	567 ± 39
E68A	2.1 ± 0.5	720 ± 32		
K114A	1.0 ± 0.3	655 ± 45		
Y134A	1.3 ± 0.4	812 ± 47		
Y134F	4.0 ± 0.8	1140 ± 59		
Q137A	1.8 ± 0.6	1122 ± 114		
Q137E	0.3 ± 0.1	440 ± 52		

Analogous sets of quenched PicoGreen fluorescence assays were performed to quantify the respective rates and processivities with which representative 5′-phosphorylated dsDNA (5′-PO_4-dsDNA) or 5′-dephosphorylated dsDNA (5′-OH-dsDNA) substrates were digested by wild-type LHK-Exo or the specified active-site point-mutant under standardized conditions. See 'Materials and Methods' section for details.

LHK-Exo digests 5′-phosphorylated dsDNA in a highly processive manner

The digestion processivity of LHK-Exo was determined by a 'heparin trap' method (61), which has similarly been used to characterize the λ-exonuclease (36), SPP1-Chu (21) and SXT-Exo (26) alkaline exonucleases. In this approach, a large excess of the heteropolysaccharide 'DNA-mimic' heparin is added to assays containing exonuclease and dsDNA substrate shortly after their initiation. This sequesters all unbound enzyme, and also prevents any disassociated enzyme from re-binding to initiate additional digestion events. DNA levels are subsequently quantified using the fluorescent PicoGreen reagent, to determine the average number of nucleotides digested during a single DNA-binding event. Under

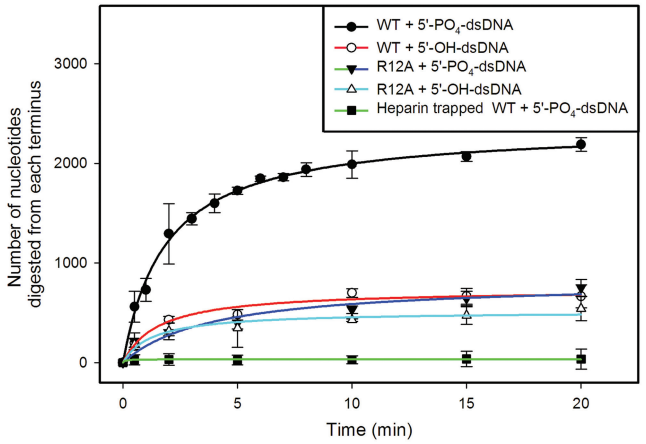

Figure 3. Processivity of double strand DNA digestion by wild-type LHK-Exo and Arg12Ala mutant. Time course analysis of the digestion of 5′-phosphorylated double strand DNA (5′-PO_4-dsDNA: EcoRV-linearized pMal-c2) and 5′-dephosphorylated double strand DNA (5′-OH-dsDNA: 5′dephosphorylated EcoRV-linearized pMal-c2) substrates by wild-type LHK-Exo and the Arg12Ala mutant form using a 'heparin trap' approach. A total of 6 µg (82 pmol of trimers) of LHK-Exo or Arg12Ala mutant protein was incubated at 25°C with 60 ng (0.015 pmol) of 5′-PO_4-dsDNA or 5′-OH-dsDNA in Tris–HCl (25 mM, pH 8.0), 1 mM DTT, 7.5 mM $MgCl_2$. After 15 s, excess heparin was added to sequester all unbound protein, and to prevent disassociated protein from re-binding. Aliquots were removed at various time points (0–20 min), and dsDNA levels were determined using fluorescent PicoGreen assays, to enable the extent of DNA digestion to be calculated. In one set of assays, heparin was added to LHK-Exo prior to the addition of dsDNA substrate, to confirm the efficacy of the heparin trap method (filled black squares, green line). Graphs show the mean number of nucleotides digested from each terminus (±SD; *y*-axis) plotted against the time of analysis (in minutes; *x*-axis).

optimal conditions (pH 8.0, 7.5 mM $MgCl_2$) at 25°C, LHK-Exo digested 2368 ± 72 nt from each terminus of 5′-phosphorylated, blunt-ended dsDNA molecules (5′-PO_4-dsDNA; EcoRV-linearized pMal-c2; 6646 bp in length; Table 2 and Figure 3). To determine whether pH had a strong influence on the digestion processivity, analogous sets of assays were performed at pH 6.8, 7.4 and 8.6 (Supplementary Table S2). Results clearly indicated that processivity was almost entirely unaffected by changes in pH over this range. This contrasts markedly with the significant effect that pH has on the rate of dsDNA digestion (Figure 2); which is 40–60% lower at these suboptimal pH values.

As the phosphorylation status of the 5′-termini has previously been shown to greatly affect both the rate and processivity of dsDNA digestion mediated by the SPP1-Chu and λ-exonuclease proteins (20,21,36), we performed analogous sets of assays using 5′-dephosphorylated EcoRV-linearized pMal-c2 as the dsDNA substrate (5′-OH-dsDNA; Table 2 and Figure 3). Results revealed that LHK-Exo trimers digested an average of 805 ± 56 nt from each 5′-dephosphorylated termini, which is two-thirds lower than the corresponding value for 5′-dephosphorylated dsDNA substrates. Similarly, LHK-Exo digested these 5′-dephosphorylated dsDNA substrates 5 times lower than the phosphorylated ones (1.6 ± 0.5 and 7.0 ± 1.8 nt/s, respectively). Taken together, these results reveal that the biochemical activities of LHK-Exo are analogous to those of phylogenetically related alkaline exonucleases; indicating there is high conservation of function within this protein family.

Overall structure of LHK-Exo

The purified recombinant LHK-Exo protein was crystallized using the sitting-drop vapor-diffusion method. Its crystal structure was solved to 1.9 Å resolution by molecular replacement using the structure of λ-exonuclease (PDB code 1AVQ) as a search model, with the crystallographic R_{work} and R_{free} converging to 20.4% and 24.1%, respectively (Table 1). The final structure includes one LHK-Exo monomer and 177 water molecules. LHK-Exo presents in the asymmetric unit as a monomer with residues 5–203 traced well in the model. Three expression vector-encoded residues located at the C-terminus of the protein were also traced in the structure. The four N-terminal residues, as well as residues 28–32 and 55–57 were unresolved, indicating that these regions may exist as flexible regions (discussed below).

The overall structure of LHK-Exo adopts an α/β 'pistol-like' fold including seven α-helices and four β-strands (Figure 4A). In addition to λ-exonuclease, a DALI search identified another four related structures, including: a phage-related exonuclease from *Haemophilus somnus* 129PT (HS_1420; PDB code 3K93; *Z*-score 17.6; r.m.s.d. 2.4 Å for 175 residues); the BGLF5 nuclease from Epstein-Barr virus (PDB code 2W4B; Z-score 11.4; r.m.s.d 2.5 Å for 168 residues) (62); the Shut-off and Exonuclease (SOX) protein from Kaposi's sarcoma-associated herpesvirus (PDB code 3FHD; *Z*-score 10.8; r.m.s.d. 2.4 Å for 160 residues) (63) and the *E. coli* RecE protein (PDB code 3H4R, *Z*-score 10.0, r.m.s.d. 3.4 Å for 152 residues) (38). The LHK-Exo, HS_1420, SOX and BGLF5 proteins belong to the λ-exonuclease family; while RecE shares more similarity with the RecB family (8). All possess DNAse and/or RNAse activities, with the exception of the HS_1420 protein, which has not yet been functionally characterized. This suggests that a conserved catalytic core domain exists in this family, despite members sharing fairly low levels of sequence identity and possessing quite different molecular weights. Superimposing the structures of these nucleases with that of LHK-Exo enabled the conserved nuclease 'core domain' to be clearly identified. This comprised a central arrangement of β-strands surrounded by α-helices (Supplementary Figure S3, panels A–F). Insertions within this central domain, as well as N-terminal and C-terminal extensions, exhibited significant diversity in shape and size.

LHK-Exo forms a toroidal trimer

The five related nuclease structures mentioned above have diverse oligomeric states: being monomeric (BGLF5 and SOX), trimeric (λ-exonuclease and HS_1420) and tetrameric (RecE). Consequently, analytical ultracentrifugation was performed to determine the oligomeric state of LHK-Exo. The sedimentation velocity showed a single

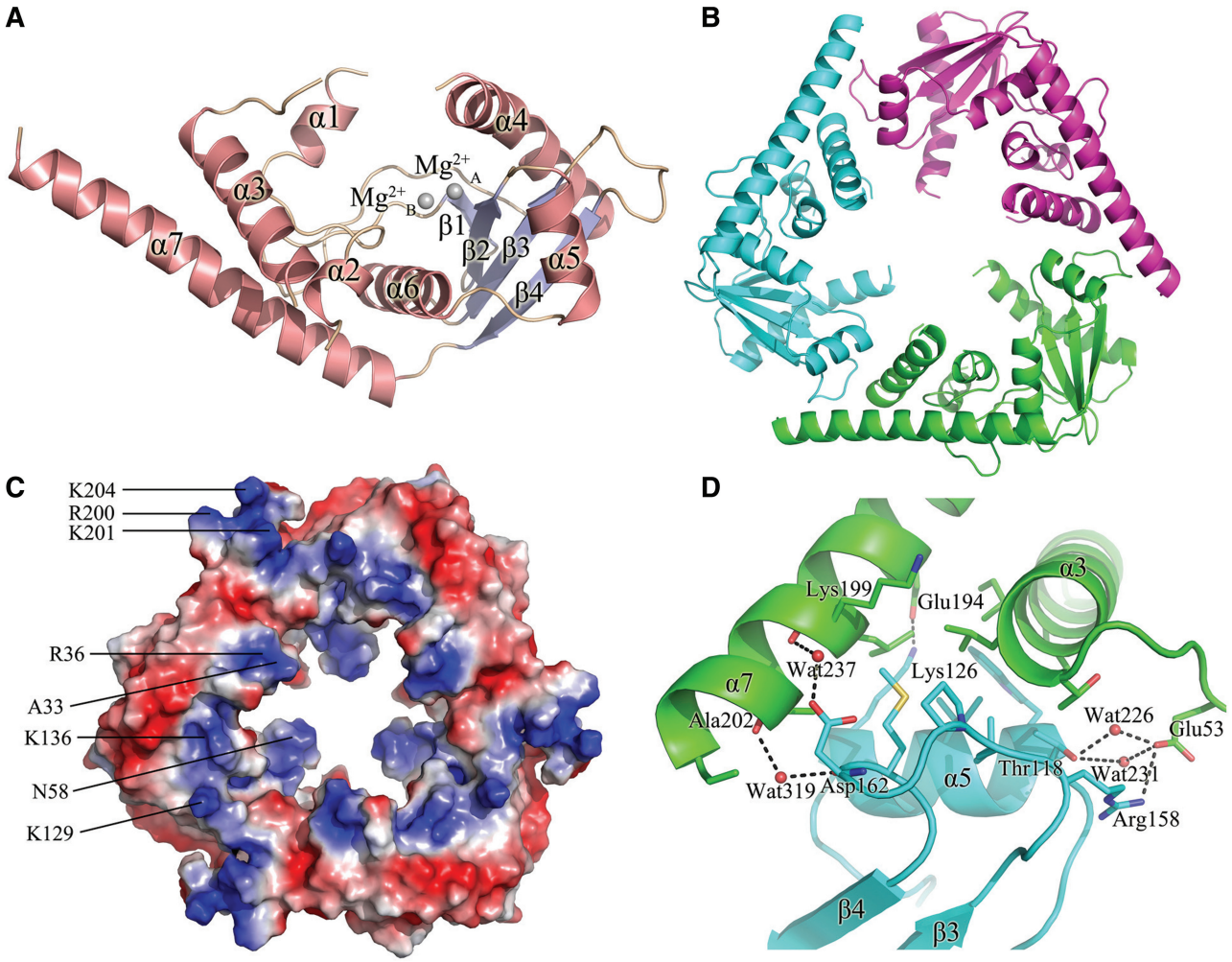

Figure 4. Crystal structure of LHK-Exo and the toroidal structure of the LHK-Exo homotrimer. (**A**) Ribbon diagram of the structure of the LHK-Exo monomer. The α-helix, β-sheet and loop regions are colored salmon-pink, light blue and wheat-yellow, respectively. The two bound Mg^{2+} ions are shown as gray spheres. (**B**) Toroidal structure of the LHK-Exo homotrimer, which has a 3-fold axis of symmetry. The three constituent monomers are colored cyan, magenta and green. (**C**) Electrostatic surface potential of the 'doughnut-like' LHK-Exo trimer, with positively charged regions colored blue and negatively charged regions in red. Residues contributing to the positively charged regions are labeled (on one protein monomer). It should be noted that residues A33 and N58 located at the C-terminal ends of the two disordered loops, respectively, over-represent the amount of the positive charge in the central channel due to their free amino groups. The toroidal LHK-Exo structure is shown with the wide-end of the tapered central channel facing the front (same orientation as in B). (**D**) Interactions at the interface between two neighboring LHK-Exo protein monomers, which are colored green and cyan. The residues involved in the interactions are shown with stick representation, and the water molecules located at the subunit interface are shown as red spheres. Salt bridges and hydrogen bonds are indicated with dashed lines.

4.4 S species corresponding to a mass of 79 ± 4 kDa (Supplementary Figure S4). The calculated mass for a trimer of LHK-Exo proteins is ~74 kDa, indicating that LHK-Exo forms a stable trimer in solution, analogous to λ-exonuclease. Although LHK-Exo presents in the asymmetric unit as a monomer, the trimeric conformation can be identified via a symmetric operation involving two neighboring molecules. Forming a 'doughnut-like' toroidal structure (Figure 4B), the constructed conformation of the LHK-Exo trimer is highly similar to that previously obtained for λ-exonuclease (41). A total solvent-accessible surface area of 992 $Å^2$ (9.1%) is buried at each monomer interface; formed mainly by α3, the α3–α4 loop and α7 in one molecule, and α5 and the β3–β4 loop in the neighboring molecule (Figure 4D). Hydrophobic interactions between the neighboring subunits appear to be the main contributors towards the overall stabilization of the trimeric conformation. Two salt bridges: $Glu53_A$-$Arg158_B$ and $Lys126_B$-$Glu194_A$ also appear to play important roles (where the A and B subscripts refer to residues present on adjacent monomers). Buried water molecules located at the subunit interfaces may also play a role in trimer assembly, via the formation of hydrogen bonds. Wat237 links the Oδ1 atom of $Asp162_B$ and the carbonyl oxygen atom of $Lys199_A$; Wat319 links the two carbonyl oxygen atoms of $Asp162_B$ and $Ala202_A$; and $Glu53_A$ is connected to $Thr118_B$ through Wat226 and Wat231. The hydrophobic interactions are mostly concentrated toward the centers of the interfaces; whilst the two salt bridges and the hydrogen

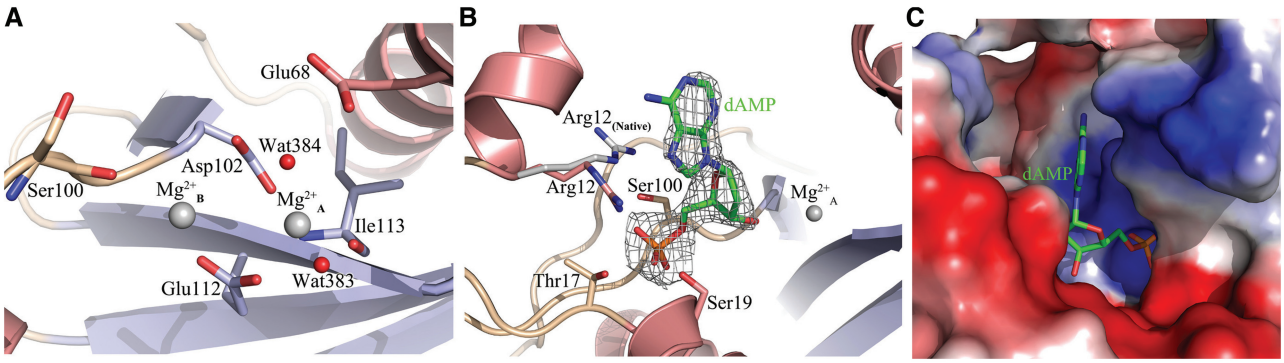

Figure 5. The LHK-Exo active site and binding site for the 5′-phosphate group of the DNA substrate. A ribbon representation of the LHK-Exo structure is colored according to the scheme used in Figure 4A. The Mg^{2+} ions and the water molecules are shown as gray and red spheres, respectively. (**A**) The residues involved in Mg^{2+} coordination, and those putatively involved in DNA phosphoesterase activity are labeled, with their chains shown in stick form. (**B**) The binding of deoxyadenosine 5′-monophosphate (dAMP) with the phosphate-recognition pocket in LHK-Exo. The dAMP molecule is shown in green stick representation and the electron density ($2F_o$-F_c contoured at 1σ, calculated at 2.60 Å) is shown in gray. The altered conformation (i.e. dAMP bound form) of the Arg12 residue side chain is shown in salmon-colored stick form, and its native conformation is shown in gray. (**C**) Surface representation of the dAMP-bound structure of LHK-Exo reveals a binding pocket for the 5′-phosphate moiety of the deoxynucleotide molecule.

bond network mediated by the ordered water molecules are positioned toward the peripheries of the interfaces.

A tapered central channel enables the passage of a single strand of DNA

A tapered channel ~20 Å in length passes through the centre of the protein trimer (the 'doughnut hole'), whose diameter is ~35 Å at the wide end and ~20 Å at the narrow end (Figure 4C). This central channel appears to be more 'open' (i.e. slightly wider) at both ends than the corresponding channel present in the λ-exonuclease trimer, due to the absence of the two (untraced) flexible loops. The RecE tetramer forms a similar central channel with a length of ~40 Å, which is tapered from ~30 Å at the wide end to ~10 Å at the narrow end (38).

The striking similarities in structural organization shared between the λ-exonuclease trimer, the RecE tetramer and the LHK-Exo trimer strongly suggests that they share a conserved mode of action. It has previously been proposed for λ-exonuclease and RecE that the dsDNA substrate enters through the wide end of the channel, and the non-digested 3′-ended ssDNA exits through the narrow end (38,41). For the LHK-Exo trimer, the positive electrostatic potential is predominantly concentrated at both ends of the central channel. However, the central channel does not have a strong overall positive electrostatic potential, and the active site possesses a strong negative charged surface (Figure 4C). Highly similar patterns of electrostatic potential distribution are also found within the central channels of the λ-exonuclease and RecE multimers. This suggests that for this group of exonucleases, the binding of the 5′-ended DNA strand (to be digested) may be primarily mediated by specific electrostatic and hydrogen bonding interactions. It also indicates that the ingress of the dsDNA substrate and the passage of the undigested strand out of the channel may be promoted primarily by electrostatic interactions.

In the LHK-Exo trimer structure, the two untraced flexible loop regions are positioned at both ends of the central channel. Residues 28–32 (TKSGY) that form the untraced loop between α2 and α3, extend out from the rim of the wide end (entrance) of the channel. Untraced residues 55–57 (RFS) lie between α3 and α4 on the opposite face of the toroid, lining the narrow channel 'exit'. In the λ-exonuclease and RecE proteins, the two loops that correspond to these two stretches of untraced residues have been proposed to form important contacts with the incoming dsDNA substrate (residues 28–32), as well as the exiting strand of undigested DNA (residues 55–57) (38,41). From a functional perspective, it seems likely that these two loops are flexible in order to accommodate the translocation of the DNA molecule through the centre of the torroid, as the termini of the 5′-strand are processively hydrolyzed. However, sequence alignments reveal that both of these stretches of untraced residues show very low levels of homology with related exonucleases (Supplementary Figure S5); with the loop between α2 and α3 (residues 28–32) being absent in some homologs. The Arg55 residue, which is located in the small untraced loop proximal to the active site, may project toward the centre of the channel exit; analogous to Lys76 in λ-exonuclease and Arg858 in RecE (38,41). Correspondingly, this residue may form important binding contacts with the undigested 3′-strand as it translocates through the central channel of the enzyme.

Coordination of active-site metal ions

Active-site metal ions were not detected in the native crystal structures of either the λ-exonuclease or RecE proteins, and their positions were only identified after Mn^{2+} ions were soaked into the crystals (38,41). Notably, in the native structure of LHK-Exo, two bound magnesium ions were identified in a negatively charged pocket within the active site (Figure 5A). One magnesium (II) ion (Mg^{2+}_A) is coordinated to the Oε2 atom of Glu112 (2.3 Å); the Oδ2 atom of Asp102

(2.4 Å); the carbonyl oxygen atom of Ile113 (2.5 Å); and to two water molecules, Wat383 (2.4 Å) and Wat384 (2.3 Å). The Mg_A^{2+}-binding site is essentially identical to the corresponding metal ion-binding site in λ-exonuclease, with the exception that two water molecules also contribute to the coordination geometry in our structure. It may also be noted that Mn^{2+} ions are chelated by histidine residues located in analogous positions in the active sites of both the RecE (His652) (38) and RecB (His956) exonucleases (64). The second, putatively more weakly bound magnesium ion (Mg_B^{2+}), is located 3.7 Å away from Mg_A^{2+}. A corresponding metal ion is absent in the λ-exonuclease structure. Mg_B^{2+} is coordinated by the Oε1 atom of Glu112 (2.7 Å); the carbonyl oxygen atom of Ser100 (2.7 Å); the Oδ1 atom of Asp102 (3.3 Å); and two water molecules, Wat387 (2.7 Å) and Wat388 (3.2 Å). In contrast to the octahedral coordination geometry typically observed in Mg^{2+}-bound enzymes, both Mg^{2+} ions in the LHK-Exo structure have a partial octahedral coordination, leaving positions available for potential interactions with DNA substrates or water molecules during the catalytic process (see below).

Based on the sequence analysis of exonuclease homologs, it has previously been proposed that the highly conserved Glu68 residue (Glu85 in λ-exonuclease) is a potential ligand for Mg^{2+} ions (8). In our structure, Glu68 seems slightly too far from the Mg_A^{2+} ion (4.1 Å) to act as a direct ligand; however, a water molecule (Wat384) forms a bridge between Glu68 and the Mg_A^{2+} ion. To probe the putative involvement of this residue, the Glu68Ala mutant was constructed, expressed and biochemically characterized. This mutant protein hydrolyzed linear 5′-phosphorylated dsDNA at a rate of 2.1 ± 0.5 nt/s (pH 8.0, 7.5 mM $MgCl_2$, 25°C), which was 70% lower than that of the wild-type LHK-Exo enzyme (7.0 ± 1.8 nt/s). However, the hydrolytic activities of the Glu68Ala mutant increased 1.6-fold (3.4 nt/s) when the $MgCl_2$ concentrations were increased to 20 mM, while those of the wild-type enzyme decreased ~35% to 4.7 nt/s (data not shown). This indicated that the Glu68Ala mutant had a reduced affinity for Mg^{2+} ions, which could be partially compensated for by increasing their concentration in solution. The coordination and functional roles of these two Mg^{2+} ions are discussed further below.

Recognition of 5′-phosphorylated DNA termini

It has previously been shown that λ-exonuclease digests dsDNA substrates that have a phosphate group at their 5′-termini (5′-PO_4) significantly more efficiently than analogous (dephosphorylated) substrates lacking this group (5′-OH) (30,35). Subramanian *et al.* (36) subsequently proposed that a phosphate ion observed in a pocket located 7 Å away from the Mg^{2+} ion in the λ-exonuclease crystal structure most likely occupied the position of the 5′-phosphorylated terminus of the DNA strand to be digested (41). They identified the Arg28, Ser35, Ser117 and Gln157 residues as forming key electrostatic interactions with this phosphate ion; which are very highly conserved amongst λ-exonuclease family members (Supplementary Figure S5). Mutation of the Arg28 residue of λ-exonuclease significantly reduced its digestion rate and processivity, and essentially abolished its ability to discriminate between 5′-phosphorylated and non-phosphorylated dsDNA substrates; demonstrating that it played a key role in binding the 5′-phosphate group. Contrastingly, an analogous residue is absent in the RecE structure (38), and correspondingly the phosphorlyation status of the 5′-strand does not appear to significantly affect the rate of dsDNA hydrolysis (55).

In the LHK-Exo structure, the putative phosphate-binding pocket incorporates the Arg12, Thr17, Ser19, Ser100 and Gln137 residues. To confirm the proposed involvement of these residues in binding the 5′-phosphorylated DNA terminus, we determined the structure of LHK-Exo in complex with dAMP to a resolution of 2.6 Å. The 5′-phosphate group of this deoxy-mononucleotide fits comfortably into the binding pocket, coordinated tightly by Arg12, Thr17, Ser19 and Ser100 (Figure 5B and C; Supplementary Figure S6). Compared with λ-exonuclease, the phosphate-binding site in LHK-Exo is located 2 Å further away from the active site. Consequently, Thr17 rather than Gln137 appears to be involved in stabilizing the phosphate group. An equivalent residue is present in λ-exonuclease (Thr33), SXT-Exo (Thr22) and PLU2936 (Thr34), but is noticeably absent in other homologous proteins. Furthermore, the adenine base moiety of the nucleotide does not appear to form any specific interactions with LHK-Exo. It interacts loosely with Trp8 and Phe9, which may supply a flexible binding environment capable of accommodating the four different nucleobase moieties present in DNA substrates.

Notably, Arg12 undergoes a significant conformational change upon dAMP binding (Figure 5B); adopting an orientation consistent with that of Arg28, which binds the phosphate ion in the λ-exonuclease structure. To determine its putative involvement in binding the terminal nucleotide residue of the DNA chain to be hydrolyzed, the exonuclease activities of the Arg12Ala mutant were investigated. The mutant protein digested 5′-phosphorylated linear (5′-PO_4) dsDNA substrates at a rate of 1.0 ± 0.3 nt/s, which is 7-fold lower than that of wild-type LHK-Exo. However, it digested the analogous 5′-dephosphorylated (5′-OH) dsDNA substrates at an essentially identical rate (1.1 ± 0.3 nt/s) (Table 2). Consistent with the proposed role of Arg28 in λ-exonuclease (36), these results indicate that the corresponding arginine residue in LHK-Exo plays a pivotal role in discriminating between 5′-phosphorylated and non-phosphorylated forms of DNA.

Structure of the LHK-Exo protein complexed with a ssDNA oligomer reveals a binding channel for the hydrolyzed DNA strand

The λ-exonuclease and RecE exonuclease proteins have the ability to distinguish and separate the two strands of the DNA duplex, and to specifically digest the 5′-ended strand in a highly processive manner (30,36,54,55). In order to better understand the mechanistic basis for 5′-stand recognition and binding, we soaked crystals

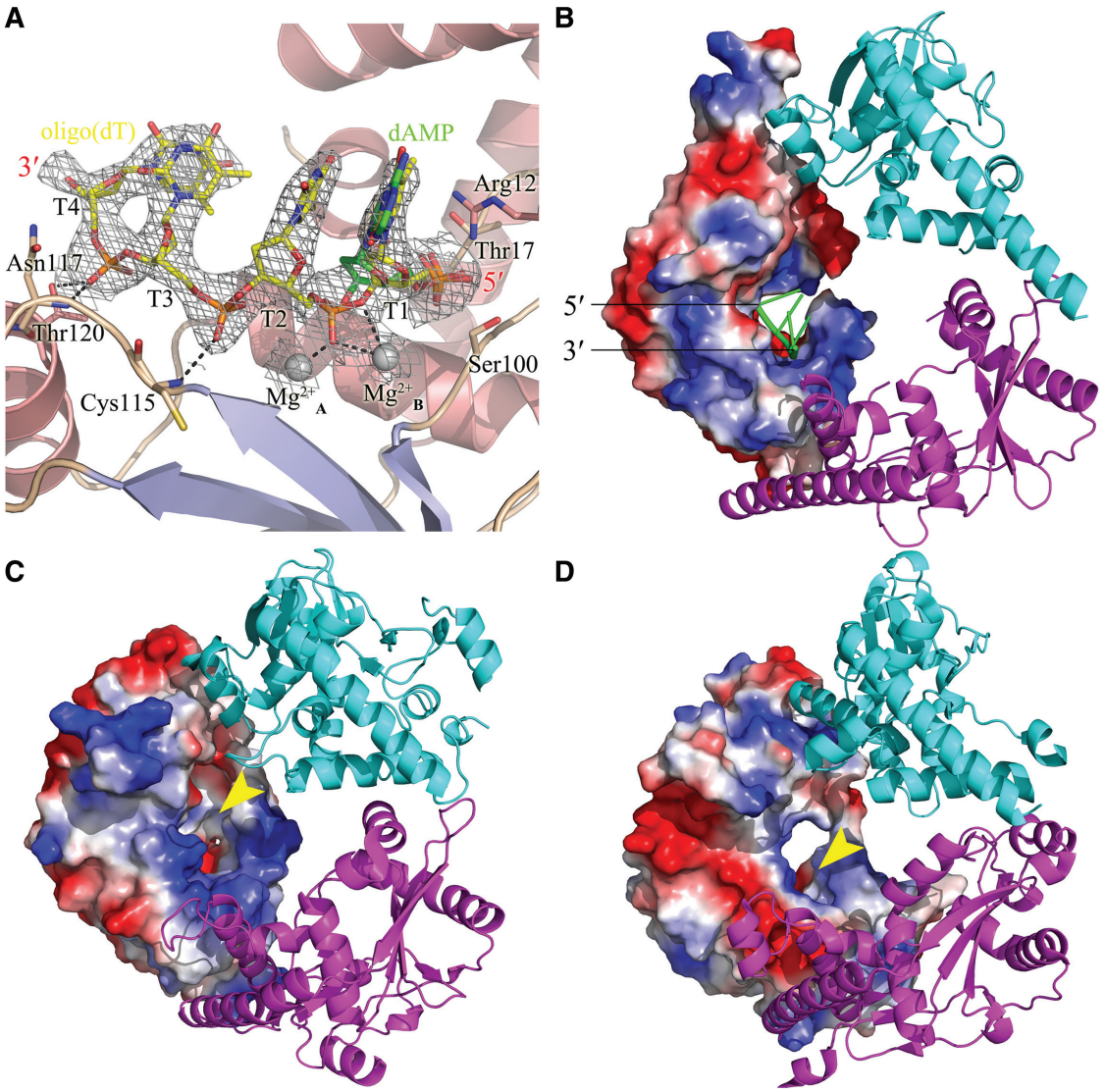

Figure 6. Structure of a ssDNA molecule bound within the active site of LHK-Exo. (**A**) The structure of LHK-Exo complexed with 5′-phosphorylated pentamer of oligothymidine [5′-PO_4-$(dT)_5$]. The oligonucleotide is shown in yellow-colored stick form and is covered by a $2F_o - F_c$ electron density map contoured at 1σ (calculated at 2.80 Å). Bound Mg^{2+} ions are shown as gray spheres, and the residues that interact with the ssDNA molecule are labeled and shown in stick form. dAMP ligand (green) is superposed on the ssDNA molecule and the 5′-phosphate groups of dAMP and ssDNA bind in identical positions. (**B**) After mapping the electrostatic potential onto the protein surface, four residues of the ssDNA ligand (green) can be seen to fit comfortably into a binding channel within one of the LHK-Exo monomers. The 5′- and 3′-termini of the complexed ssDNA molecule are indicated. The trimer is shown with the wide-end of the tapered central channel facing the front, i.e. DNA enters from the front face. (**C** and **D**) Surface representations of the native structures of λ-exonuclease (PDB code: 1AVQ) and the HS_1420 exonuclease from *Haemophilus somnus 129PT* (PDB code: 3K93), respectively. Both exonucleases contain similar channels (one is indicated with a yellow arrow) that may accommodate the 5′-end of the DNA strand to be hydrolyzed; with a more 'closed' channel observable for λ-exonuclease (C). Both protein trimers are shown orientated in a manner analogous to that of LHK-Exo in B, revealing an equivalent positioning of the ssDNA-binding channels adjacent to the central tapered channel.

of LHK-Exo with buffer containing a synthetic 5-mer of oligothymidine that had been enzymatically phosphorylated at the 5′-termini [5′-PO_4-$(dT)_5$]. We were subsequently able to solve the structure of LHK-Exo-oligonucleotide complex to 2.8 Å. We could clearly identify four complete nucleotide residues (T1–T4) in each monomer (Figure 6A and Supplementary Figure S6). The ssDNA chain fits comfortably into a channel formed by α4, α6 and the α5–β2 loop (Figure 6B). The T1 residue was located in the same site as the dAMP 5′-mononucleotide was (see above); coordinated tightly by Arg12, Thr17, Ser19 and Ser100. The 5′-phosphate group (of T1) was positioned identically to the 5′-phosphate group of dAMP in the protein complex (Figure 6A). However, the 3′-OH of the T1 deoxyribose ring was located 1.5 Å closer to the active site, and was hydrogen bonded to Mg^{2+}_B (3.0 Å). The fact that the dAMP and T1 oligonucleotide residues occupied the same binding site in the LHK-Exo crystal structure, suggests that this is the most stable nucleotide binding

site. This seems reasonable, as this is the residue that is hydrolytically cleaved from the DNA chain; being subsequently released as a 5′-mononucleotide.

The scissile phosphate between the T1 and T2 nucleotides was stabilized by the two Mg^{2+} ions via the O1 atom. In addition, the O1 atom of the phosphate moiety between T2 and T3 formed a 2.9 Å hydrogen bond with the main chain amide nitrogen of Cys115. The O1 atom of the phosphate moiety between T3 and T4 was stabilized by hydrogen bonds to the hydroxyl group of Thr120 (2.6 Å) and the main chain amide nitrogen of Asn117 (3.1 Å). Notably, no specific interactions appeared to be formed between the base moieties of the oligonucleotide and the LHK-Exo protein. This would be consistent with the exonuclease protein exhibiting little or no preference for dsDNA substrates with differing nucleotide compositions. We could not identify the T5 residue at the 3′-termini of the oligonucleotide in the complex structure, suggesting that this nucleotide does not form stable binding interactions with the enzyme.

Our findings indicate that LHK-Exo recognizes and tightly binds at least four consecutive nucleotides at the 5′-termini of the DNA strand to be digested (T1–T4), via the sugar–phosphate backbone. This stretch of residues at the 5′-termini is most probably separated from the complementary residues of the (undigested) DNA strand. However, it is difficult to extrapolate our results to estimate the degree to which the B-form DNA duplex is unwound by enzyme, or precisely how many base pairs are separated. Notably, our X-ray data is consistent with the findings of Dahlroth *et al.* (63), who modeled an oligonucleotide into the active site of the SOX exonuclease crystal structure. They predicted that SOX was also likely to accommodate 4 nt.

It should be noted that the soaking experiments did not lead to ssDNA molecule cleavage in our complex structure, which was similarly observed for the EcoRV endonuclease (65). While LHK-Exo possessed notable ssDNA exonuclease activities in the presence of 7.5 mM Mg^{2+} ions (in solution phase at pH 8.0, 37°C), there was negligible activity when Mg^{2+} ions were omitted (Supplementary Figure S2). The crystal soaking experiments were conducted at high salt concentrations, at 20°C, in the absence of free Mg^{2+} ions, which may have facilitated the retention of an intact tetranucleotide molecule within the active site. In addition, some quaternary structure changes may occur during the catalysis process that are in conflict with the crystal packing forces, which may be responsible for the prevention of ssDNA cleavage. It may also be the case that the position of the tetranucleotide molecule within the active site corresponds to a protected region of the substrate DNA as it is processed by LHK-Exo, analogous to a proposed ~11-bp protected region of DNA in λ-exonuclease (41), although further experiments are needed to confirm this.

A similar ssDNA-binding channel exists in λ-exonuclease and the HS_1420 phage-related exonuclease from *Haemophilus somnus 129PT*, indicating that they may share a common binding mechanism for the 5′-DNA strand (Figure 6C and D). However, the corresponding ssDNA-binding channel of the λ-exonuclease trimer appears to be slightly more 'closed' at the 5′-phosphate recognition site, due to the presence of an extra α-helix and 3_{10}-helix located at its N-terminus (Figure 6B–D).

We further noted an additional binding pattern in the structure of the oligonucleotide–protein complex. The three central phosphodiester groups of the DNA chain (T1–T2, T2–T3 and T3–T4) are only coordinated to the protein via their O1 atoms, leaving their O2 atoms orientated toward the same inner-surface of the channel (Supplementary Figure S7). This surface is composed of Lys114, Tyr134 and Gln137; three residues that are very highly conserved among λ-exonuclease family members (Supplementary Figure S5). The distances between the three O2 atoms and the respective polar side-chain groups of these residues are maintained within 4.5 Å. We speculated these three residues may play important roles in binding and orienting the 5′-DNA strand during its step-wise hydrolysis. To probe the role of these universally conserved residues, we prepared the Lys114Ala, Tyr134Ala, Tyr134Phe, Gln137Ala and Gln137Glu mutants and characterized their activities. The Lys114Ala and Tyr134Ala mutants digested 5′-phosphorylated dsDNA with rates of 1.0 and 1.7 nt/s, respectively; which was roughly 15% and 25% that of the wild-type enzyme (Table 2). However, the digestion rate of the Tyr134Phe mutant was 4.0 nt/s, which was ~60% that of wild-type LHK-Exo protein. The processivity of digestion mediated by these three mutants were all roughly similar (~1100 nt hydrolyzed per binding event), which is ~50% that of the wild-type enzyme. This data indicate that neither of these individual residues plays vital roles in DNA hydrolysis or translocation. However, the fact that the Tyr134Phe mutant was significantly more active than the Tyr134Ala mutant suggests that the presence of the benzyl side-chain moiety (compared to a methyl group) is beneficial for the overall protein architecture, or at least plays an ancillary role in substrate binding or catalysis. The dsDNA digestion rate of the Gln137Ala mutant was 4-fold lower than the wild-type (1.8 ± 0.6 nt/s), while the Gln137Glu mutant was 23-fold less active. This ~6-fold difference in catalytic activities between these two mutant forms most likely results from considerable electrostatic repulsion between the negatively charged pendant carboxylate group and the phosphate of the DNA substrate. Notably, it has been proposed that the corresponding glutamine residue in RecE (Gln781), is hydrogen bonded to Lys761; thereby modulating the orientation or catalytic activity of this highly conserved lysine residue (see below) (38).

Proposed two metal ion catalytic mechanism for LHK-Exo

Although the crystal structure of λ-exonuclease was solved more than a decade ago, its precise catalytic mechanism remains to be established. Our crystal structure of the LHK-Exo protein in complex with a ssDNA molecule and Mg^{2+} ions gives a highly informative snapshot of the active site of this exonuclease enzyme. In particular, the relative positioning of the two Mg^{2+} ions suggests that

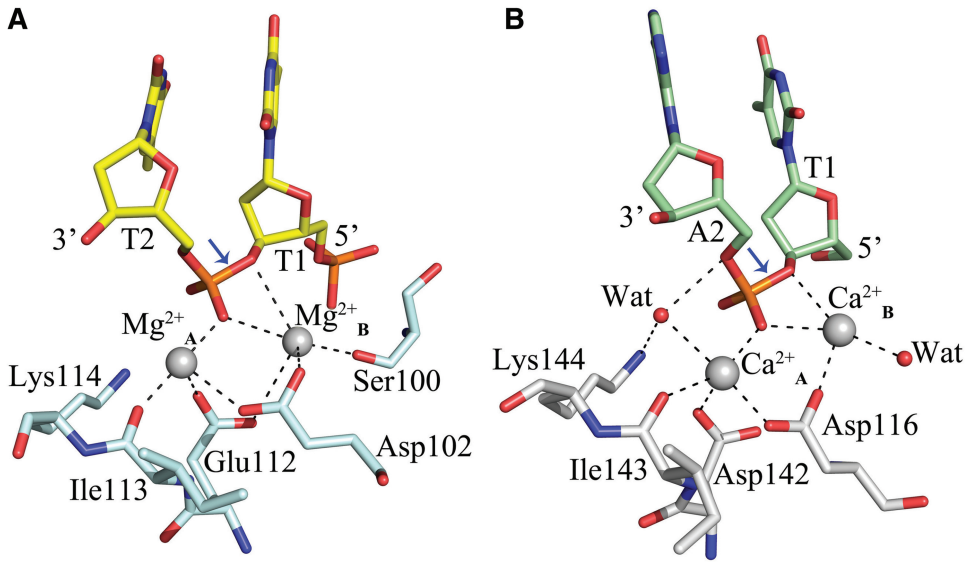

Figure 7. Conservation of catalytic site architecture between the LHK-Exo DNA exonuclease and the BglI DNA endonuclease. (**A**) Binding interactions formed between the scissile phosphodiester group (blue arrow) of the oligonucleotide and the two Mg^{2+} ions and active site residues in LHK-Exo. Two thymidine residues of the bound DNA molecule (T1, T2) are shown in yellow-colored stick form. The two complexed Mg^{2+} ions (gray spheres) and LHK-Exo active site residues (pale cyan-colored stick form) are labeled. Mg^{2+}_A may contribute to the deprotonation and subsequent stabilization of the nucleophilic water molecule. Mg^{2+}_B could stabilize the scissile P-O bond and facilitate the departure of the ribose 3′ hydroxyl moiety. Lys114 is postulated to act as a conserved general base catalyst. (**B**) Architecture of the active site of the BglI restriction endonuclease with two complexed Ca^{2+} ions (PDB code: 1DMU). Adenosine and thymidine nucleotides of the bound DNA molecule are shown in pale green-colored stick form. Active site residues are shown in gray-colored stick form and the bound Ca^{2+} ions are shown as gray spheres. Lys144 has been proposed to act as a conserved general base catalyst, in a 'two-metal' catalytic mechanism. The two water molecules that putatively act as the nucleophile and general acid are colored red. The blue arrow indicates the scissile P-O bond.

this protein may utilize a two-metal ion catalytic mechanism for DNA hydrolysis (Figure 7A). This has previously been proposed to occur in a variety of polymerases (66,67) and nucleases (68–72), but within diverse catalytic scenarios. The two bound Mg^{2+} ions are located approximately 4 Å away from each other, parallel to the scissile bond. Mg^{2+}_A is tightly coordinated by Glu112, Asp102, Ile113 and the scissile phosphate; but Mg^{2+}_B is more weakly bound by Glu112, Ser100, Asp102 and the scissile P-O. In the presence of DNA, the two Mg^{2+} ions are jointly coordinated by the side-chain carboxylate groups of two residues that are highly conserved within this protein family: Asp102 and Glu112 (Supplementary Figure S5), as well as the O1 atom of the scissile phosphate. It must be noted however, that neither of these magnesium ions have a fully occupied octahedral coordination in our structure. It is possible that water molecules crucially involved in the catalytic process may be missing due to the modest resolution of our data set (discussed below). As described previously, Mg^{2+}_A could contribute to the deprotonation and subsequent stabilization of the nucleophilic water molecule. Consistent with the classical two-metal ion mechanism for a range of polymerases and nucleases, including λ-exonuclease, Mg^{2+}_B may stabilize the scissile P-O bond and neutralize the negative charge transferred to the 3′-hydroxyl group of the terminal nucleotide-5′-phosphate moiety, facilitating its departure (73). In addition, both metal cations may help to neutralize the negative charge associated with the pentavalent phosphate moiety transiently generated during the attack of the nucleophilic water molecule.

In the proposed two metal ion catalytic mechanism, the highly conserved Lys114 (equivalent to Lys131 in λ-exonuclease, Lys761 in RecE, Lys144 in BglI) may function as a conserved general base for the activation of the nucleophilic water molecule, which has previously been suggested (38,41). Mg^{2+}_A and the (protonated) ε-amino group of Lys114 may stabilize the negatively charged pentacoordinate phosphate intermediate. Substituting Lys114 for an alanine resulted in a ∼85% loss in activity compared with wild-type LHK-Exo (1.0 ± 0.3 nt/s compared to 7.0 ± 1.8 nt/s; Table 2). Its digestion processivity was similarly reduced by ∼75% (655 ± 45 nt compared to 2368 ± 72 nt digested per binding event). This indicates that while this highly conserved lysine residue may play an important role in facilitating phosphodiester bond hydrolysis, its individual functions are not essential.

We further compared the DNA-bound structure with that of the BglI restriction endonuclease in complex with a specific substrate DNA model and Ca^{2+} ions (PDB code: 1DMU) (74), in order to gage the similarity between LHK-Exo and a representative type II restriction enzyme that employs a two metal ion catalytic mechanism (Figure 7A and B). Two crucial water molecules bound in the active site of BglI appear be involved in the catalytic process: the water molecule hydrogen bonded to Lys144 and metal I (Ca^{2+}_A) performs an in-line attack of the scissile phosphate; the other water molecule bound to metal II (Ca^{2+}_B) acts as a general acid, protonating the leaving group. Unfortunately, it is not possible to identify any functionally equivalent water molecules in our ssDNA

bound structure of LHK-Exo due to the modest resolution and slightly smeared electron density. Furthermore, we are not able to identify the precise mechanism by which LHK-Exo promotes the formation and stabilization of the pentacoordinate phosphate intermediate. However, the analogous composition and orientation of the active site residues and metal ions within the BglI and LHK-Exo crystal structures supports the operation of an equivalent two-metal ion catalytic mechanism. A more precise identification and positioning of the crucial residues and water molecules within the LHK-Exo active site will require higher resolution structures in complex with ssDNA or even (partially) dsDNA model substrates.

The biophysical, biochemical and structural data presented here suggest that a considerable number of residues within the central channel of the LHK-Exo trimer are actively involved in binding, orienting and translocating the DNA substrate; as well as playing a direct catalytic role in the iterative hydrolysis of nucleotides from the terminus of the 5′-strand. Our results are consistent with the mutational analysis of the RecE exonuclease, where the precise roles of the individual active site residues proved similarly difficult to unambiguously establish, and there appears to be a significant amount of catalytic redundancy (38). While many key mechanistic points remain to be established, the structures of the mononucleotide and oligonucleotide-bound complexes of LHK-Exo described here are the first such structures reported for a member of this notable class of exonucleases. They provide structural evidence consistent with the operation of a two metal ion catalytic mechanism for this protein, and offer significant additional insight into the catalytic mechanisms of related nucleases of phage/viral origin.

ADDENDUM

While revising our manuscript, we became aware of an 'early edition' article reporting two DNA-bound complexes of λ-exonuclease (75). The authors used 12-mer dsDNA (Form 1) or a 14-mer:12-mer DNA duplex (containing a 2 nt 5′-overhang; Form 2) for co-crystallization with the wild-type and catalytically inactive K131A mutant forms of λ-exonuclease, respectively. Interestingly, the position of the ssDNA in our LHK-Exo complex structure overlaps precisely with the 5′-end of the 14-mer DNA strand in the Form 2 structure, including the 2-nt overhang. The λ-exonuclease: dsDNA structures reported by Zhang and colleagues are highly consistent with our LHK-Exo:ssDNA structure reported here. Taken together, our studies reveal that these two exonuclease proteins share an equivalent arrangement of metal cofactor and substrate bonding interactions, and employ a functionally identical two metal ion catalytic mechanism.

ACCESSION NUMBERS

The coordinates of the crystal structures of LHK-Exo, and the LHK-Exo:dAMP and LHK-Exo:ssDNA complexes have been deposited in the RCSB Protein Data Bank with the accession codes 3SYY, 3SZ4 and 3SZ5, respectively.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors are grateful to Zhiyong Lou for assistance with data collection and processing; all of the beamline scientists at the Photon Factory in Japan for technical support; and Xiaoxia Yu for analysis of the analytical ultracentrifugation data. The authors also thank the technical staff of the Centralized Research Laboratory of the Faculty of Dentistry (HKU) for their assistance.

FUNDING

Ministry of Science and Technology of China Project 973 (grant number 2007CB914301 to M.B.); General Research Fund (GRF) award from the Research Grants Council of Hong Kong (grant number 779109 to R.M.W.); Infection and Immunology Strategic Research Theme of the University of Hong Kong (to R.M.W.). Funding for open access charge: Ministry of Science and Technology of China Project 973 (grant number 2007CB914301 to M.B.).

Conflict of interest statement. None declared.

REFERENCES

1. Kreuzer,K.N. (2005) Interplay between DNA replication and recombination in prokaryotes. *Annu. Rev. Microbiol.*, **59**, 43–67.
2. San Filippo,J., Sung,P. and Klein,H. (2008) Mechanism of eukaryotic homologous recombination. *Annu. Rev. Biochem.*, **77**, 229–257.
3. Cox,M.M. (2001) Recombinational DNA repair of damaged replication forks in Escherichia coli: questions. *Annu. Rev. Genet.*, **35**, 53–82.
4. Cox,M.M. (1998) A broadening view of recombinational DNA repair in bacteria. *Genes Cells*, **3**, 65–78.
5. Thomas,C.M. and Nielsen,K.M. (2005) Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.*, **3**, 711–721.
6. Szczepanska,A.K. (2009) Bacteriophage-encoded functions engaged in initiation of homologous recombination events. *Crit. Rev. Microbiol.*, **35**, 197–220.
7. Kuzminov,A. (1999) Recombinational repair of DNA damage in Escherichia coli and bacteriophage lambda. *Microbiol. Mol. Biol. Rev.*, **63**, 751–813, (table of contents).
8. Aravind,L., Makarova,K.S. and Koonin,E.V. (2000) SURVEY AND SUMMARY: holliday junction resolvases and related nucleases: identification of new families, phyletic distribution and evolutionary trajectories. *Nucleic Acids Res.*, **28**, 3417–3432.
9. Lopes,A., Amarir-Bouhram,J., Faure,G., Petit,M.A. and Guerois,R. (2010) Detection of novel recombinases in bacteriophage genomes unveils Rad52, Rad51 and Gp2.5 remote homologs. *Nucleic Acids Res.*, **38**, 3952–3962.
10. Iyer,L.M., Koonin,E.V. and Aravind,L. (2002) Classification and evolutionary history of the single-strand annealing proteins, RecT, Redbeta, ERF and RAD52. *BMC Genomics*, **3**, 8.

11. Li,Z., Karakousis,G., Chiu,S.K., Reddy,G. and Radding,C.M. (1998) The beta protein of phage lambda promotes strand exchange. *J. Mol. Biol.*, **276**, 733–744.
12. Kolodner,R., Hall,S.D. and Luisi-DeLuca,C. (1994) Homologous pairing proteins encoded by the Escherichia coli recE and recT genes. *Mol. Microbiol.*, **11**, 23–30.
13. Ayora,S., Missich,R., Mesa,P., Lurz,R., Yang,S., Egelman,E.H. and Alonso,J.C. (2002) Homologous-pairing activity of the Bacillus subtilis bacteriophage SPP1 replication protein G35P. *J. Biol. Chem.*, **277**, 35969–35979.
14. Court,D.L., Sawitzke,J.A. and Thomason,L.C. (2002) Genetic engineering using homologous recombination. *Annu. Rev. Genet.*, **36**, 361–388.
15. Mosberg,J.A., Lajoie,M.J. and Church,G.M. (2010) Lambda red recombineering in Escherichia coli occurs through a fully single-stranded intermediate. *Genetics*, **186**, 791–799.
16. Maresca,M., Erler,A., Fu,J., Friedrich,A., Zhang,Y. and Stewart,A.F. (2010) Single-stranded heteroduplex intermediates in lambda Red homologous recombination. *BMC Mol. Biol.*, **11**, 54.
17. Poteete,A.R. (2001) What makes the bacteriophage lambda Red system useful for genetic engineering: molecular mechanism and biological function. *FEMS Microbiol. Lett.*, **201**, 9–14.
18. Zhang,Y., Buchholz,F., Muyrers,J.P. and Stewart,A.F. (1998) A new logic for DNA engineering using recombination in Escherichia coli. *Nat. Genet.*, **20**, 123–128.
19. Muyrers,J.P., Zhang,Y., Buchholz,F. and Stewart,A.F. (2000) RecE/RecT and Redalpha/Redbeta initiate double-stranded break repair by specifically interacting with their respective partners. *Genes Dev.*, **14**, 1971–1982.
20. Martinez-Jimenez,M.I., Alonso,J.C. and Ayora,S. (2005) Bacillus subtilis bacteriophage SPP1-encoded gene 34.1 product is a recombination-dependent DNA replication protein. *J. Mol. Biol.*, **351**, 1007–1019.
21. Vellani,T.S. and Myers,R.S. (2003) Bacteriophage SPP1 Chu is an alkaline exonuclease in the SynExo family of viral two-component recombinases. *J. Bacteriol.*, **185**, 2465–2474.
22. van Kessel,J.C. and Hatfull,G.F. (2007) Recombineering in Mycobacterium tuberculosis. *Nat. Methods*, **4**, 147–152.
23. van Kessel,J.C., Marinelli,L.J. and Hatfull,G.F. (2008) Recombineering mycobacteria and their phages. *Nat. Rev. Microbiol.*, **6**, 851–857.
24. Wozniak,R.A., Fouts,D.E., Spagnoletti,M., Colombo,M.M., Ceccarelli,D., Garriss,G., Dery,C., Burrus,V. and Waldor,M.K. (2009) Comparative ICE genomics: insights into the evolution of the SXT/R391 family of ICEs. *PLoS Genet.*, **5**, e1000786.
25. Garriss,G., Waldor,M.K. and Burrus,V. (2009) Mobile antibiotic resistance encoding elements promote their own diversity. *PLoS Genet.*, **5**, e1000775.
26. Chen,W.Y., Ho,J.W.S., Huang,J.D. and Watt,R.M. (2011) Functional characterization of an alkaline exonuclease and single strand annealing protein from the SXT genetic element of Vibrio cholerae. *BMC Mol. Biol.*, **12**, 16.
27. Datta,S., Costantino,N., Zhou,X. and Court,D.L. (2008) Identification and analysis of recombineering functions from Gram-negative and Gram-positive bacteria and their phages. *Proc. Natl Acad. Sci. USA*, **105**, 1626–1631.
28. Swingle,B., Bao,Z., Markel,E., Chambers,A. and Cartinhour,S. (2010) Recombineering using RecTE from Pseudomonas syringae. *Appl. Environ. Microbiol.*, **76**, 4960–4968.
29. Sawitzke,J.A., Thomason,L.C., Costantino,N., Bubunenko,M., Datta,S. and Court,D.L. (2007) Recombineering: in vivo genetic engineering in E. coli, S. enterica, and beyond. *Methods Enzymol.*, **421**, 171–199.
30. Little,J.W. (1967) An exonuclease induced by bacteriophage lambda. II. Nature of the enzymatic reaction. *J. Biol. Chem.*, **242**, 679–686.
31. Little,J.W., Lehman,I.R. and Kaiser,A.D. (1967) An exonuclease induced by bacteriophage lambda. I. Preparation of the crystalline enzyme. *J. Biol. Chem.*, **242**, 672–678.
32. Bujnicki,J.M. and Rychlewski,L. (2001) The herpesvirus alkaline exonuclease belongs to the restriction endonuclease PD-(D/E)XK superfamily: insight from molecular modeling and phylogenetic analysis. *Virus Genes*, **22**, 219–230.
33. Carter,D.M. and Radding,C.M. (1971) The role of exonuclease and beta protein of phage lambda in genetic recombination. II. Substrate specificity and the mode of action of lambda exonuclease. *J. Biol. Chem.*, **246**, 2502–2512.
34. van Oijen,A.M., Blainey,P.C., Crampton,D.J., Richardson,C.C., Ellenberger,T. and Xie,X.S. (2003) Single-molecule kinetics of lambda exonuclease reveal base dependence and dynamic disorder. *Science*, **301**, 1235–1238.
35. Mitsis,P.G. and Kwagh,J.G. (1999) Characterization of the interaction of lambda exonuclease with the ends of DNA. *Nucleic Acids Res.*, **27**, 3057–3063.
36. Subramanian,K., Rutvisuttinunt,W., Scott,W. and Myers,R.S. (2003) The enzymatic basis of processivity in lambda exonuclease. *Nucleic Acids Res.*, **31**, 1585–1596.
37. Sriprakash,K.S., Lundh,N., Huh,M.-O. and Radding,C.M. (1975) The specificity of lambda exonuclease. Interactions with single-stranded DNA. *J. Biol. Chem.*, **250**, 5438–5445.
38. Zhang,J., Xing,X., Herr,A.B. and Bell,C.E. (2009) Crystal structure of E. coli RecE protein reveals a toroidal tetramer for processing double-stranded DNA breaks. *Structure*, **17**, 690–702.
39. Dapprich,J. (1999) Single-molecule DNA digestion by lambda-exonuclease. *Cytometry*, **36**, 163–168.
40. Matsuura,S., Komatsu,J., Hirano,K., Yasuda,H., Takashima,K., Katsura,S. and Mizuno,A. (2001) Real-time observation of a single DNA digestion by lambda exonuclease under a fluorescence microscope field. *Nucleic Acids Res.*, **29**, E79.
41. Kovall,R. and Matthews,B.W. (1997) Toroidal structure of lambda-exonuclease. *Science*, **277**, 1824–1827.
42. Woo,P.C., Lau,S.K., Tse,H., Teng,J.L., Curreem,S.O., Tsang,A.K., Fan,R.Y., Wong,G.K., Huang,Y., Loman,N.J. *et al.* (2009) The complete genome and proteome of Laribacter hongkongensis reveal potential mechanisms for adaptations to different temperatures and habitats. *PLoS Genet.*, **5**, e1000416.
43. Woo,P.C.Y., Lau,S.K.P., Teng,J.L.L., Que,T.L., Yung,R.W.H., Luk,W.K., Lai,R.W.M., Hui,W.T., Wong,S.S.Y., Yau,H.H. *et al.* (2004) Association of Laribacter hongkongensis in community-acquired gastroenteritis with travel and eating fish: a multicentre case-control study. *Lancet*, **363**, 1941–1947.
44. Woo,P.C., Lau,S.K., Teng,J.L. and Yuen,K.Y. (2005) Current status and future directions for Laribacter hongkongensis, a novel bacterium associated with gastroenteritis and traveller's diarrhoea. *Curr. Opin. Infect. Dis.*, **18**, 413–419.
45. Tolun,G. and Myers,R.S. (2003) A real-time DNase assay (ReDA) based on PicoGreen fluorescence. *Nucleic Acids Res.*, **31**, e111.
46. Schuck,P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.*, **78**, 1606–1619.
47. Otwinowski,Z.a.W.M. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, 307–326.
48. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
49. Collaborative Computational Project, Number 4. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **50**, 760–763.
50. Perrakis,A., Morris,R. and Lamzin,V.S. (1999) Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.*, **6**, 458–463.
51. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
52. Murshudov,G.N., Vagin,A.A. and Dodson,E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.*, **53**, 240–255.
53. Laskowski,R.A., MacArthur,M.W., Moss,D.S. and Thornton,J.M. (1993) Procheck-a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr*, **26**, 283–291.
54. Thomas,K.R. and Olivera,B.M. (1978) Processivity of DNA exonucleases. *J. Biol. Chem.*, **253**, 424–429.

55. Joseph,J.W. and Kolodner,R. (1983) Exonuclease VIII of Escherichia coli. II. Mechanism of action. *J. Biol. Chem.*, **258**, 10418–10424.
56. Stolzenberg,M.C. and Ooka,T. (1990) Purification and properties of Epstein-Barr virus DNase expressed in Escherichia coli. *J. Virol.*, **64**, 96–104.
57. Baylis,S.A., Purifoy,D.J. and Littler,E. (1989) The characterization of the EBV alkaline deoxyribonuclease cloned and expressed in E. coli. *Nucleic Acids Res.*, **17**, 7609–7622.
58. Li,L. and Rohrmann,G.F. (2000) Characterization of a baculovirus alkaline nuclease. *J. Virol.*, **74**, 6401–6407.
59. Bronstein,J.C. and Weber,P.C. (1996) Purification and characterization of herpes simplex virus type 1 alkaline exonuclease expressed in Escherichia coli. *J. Virol.*, **70**, 2008–2013.
60. Yang,D.H., de Jong,J.G., Makhmoudova,A., Arif,B.M. and Krell,P.J. (2004) Choristoneura fumiferana nucleopolyhedrovirus encodes a functional 3'-5' exonuclease. *J. Gen. Virol.*, **85**, 3569–3573.
61. Reddy,M.K., Weitzel,S.E. and von Hippel,P.H. (1992) Processive proofreading is intrinsic to T4 DNA polymerase. *J. Biol. Chem.*, **267**, 14157–14166.
62. Buisson,M., Geoui,T., Flot,D., Tarbouriech,N., Ressing,M.E., Wiertz,E.J. and Burmeister,W.P. (2009) A bridge crosses the active-site canyon of the Epstein-Barr virus nuclease with DNase and RNase activities. *J. Mol. Biol.*, **391**, 717–728.
63. Dahlroth,S.L., Gurmu,D., Haas,J., Erlandsen,H. and Nordlund,P. (2009) Crystal structure of the shutoff and exonuclease protein from the oncogenic Kaposi's sarcoma-associated herpesvirus. *FEBS J.*, **276**, 6636–6645.
64. Singleton,M.R., Dillingham,M.S., Gaudier,M., Kowalczykowski,S.C. and Wigley,D.B. (2004) Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature*, **432**, 187–193.
65. Kostrewa,D. and Winkler,F.K. (1995) Mg2+ binding to the active site of EcoRV endonuclease: a crystallographic study of complexes with substrate and product DNA at 2 A resolution. *Biochemistry*, **34**, 683–696.
66. Doublie,S., Tabor,S., Long,A.M., Richardson,C.C. and Ellenberger,T. (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 A resolution. *Nature*, **391**, 251–258.
67. Steitz,T.A. (1998) A mechanism for all polymerases. *Nature*, **391**, 231–232.
68. Beese,L.S. and Steitz,T.A. (1991) Structural basis for the 3′-5′ exonuclease activity of Escherichia coli DNA polymerase I: a two metal ion mechanism. *EMBO J*, **10**, 25–33.
69. Lee,J.Y., Chang,J., Joseph,N., Ghirlando,R., Rao,D.N. and Yang,W. (2005) MutH complexed with hemi- and unmethylated DNAs: coupling base recognition and DNA cleavage. *Mol. Cell.*, **20**, 155–166.
70. Nowotny,M. and Yang,W. (2006) Stepwise analyses of metal ions in RNase H catalysis from substrate destabilization to product release. *EMBO J.*, **25**, 1924–1933.
71. Stahley,M.R. and Strobel,S.A. (2006) RNA splicing: group I intron crystal structures reveal the basis of splice site selection and metal ion catalysis. *Curr. Opin. Struct. Biol.*, **16**, 319–326.
72. Yang,W. (2011) Nucleases: diversity of structure, function and mechanism. *Q Rev. Biophys.*, **44**, 1–93.
73. Yang,W., Lee,J.Y. and Nowotny,M. (2006) Making and breaking nucleic acids: two-Mg2+-ion catalysis and substrate specificity. *Mol. Cell.*, **22**, 5–13.
74. Newman,M., Lunnen,K., Wilson,G., Greci,J., Schildkraut,I. and Phillips,S.E. (1998) Crystal structure of restriction endonuclease BglI bound to its interrupted DNA recognition sequence. *EMBO J.*, **17**, 5466–5476.
75. Zhang,J., McCabe,K.A. and Bell,C.E. (2011) Crystal structures of {lambda} exonuclease in complex with DNA suggest an electrostatic ratchet mechanism for processivity. *Proc. Natl Acad. Sci. USA*, **108**, 11872–11877.