# SUPPORTING FIGURES AND TABLE

A	XNI_ECO24 : DPOln_THEA : EXO5_BPT5 : RNH_BPT4 : FEN1_HUMAN :	20 MRGMLPLFEPKGR MRSKSWGKFIEEEEAEMASRRM MGLQGLAKLIADVAPSAIRENDIKSYFGRK A	40 ALN IRRIHAVQGGLTTS GHH AYRTFHALKGLTTS GTN GFRFKHNNSPDKE ASMSIYQFLIAVRQGGDVLQNE	60 8 SPCVETCOHA DOL MHSO RGEPVQAVYGFAKS LKA KE-D KRPFASSYVST OSLAKSYS KINLSNVRHLILNS KFN KKAK EGETTSHLMGMFYRTIRM ENGI	0 - : 41 - : 58 - : 59 T : 60 - : 79
	XNI_ECO24 : DPOln_THEA : EXO5_BPT5 : RNH_BPT4 : FEN1_HUMAN :	PTHA A FDDENRSG HORLPD AGRPPM GDAV V DDAKAPS HEAYGG AGRAP -ART-T V GOKGKSV LEHLPE ONODE LGYTKI LCID-AKSGY RDFAYY KNRGKA -KP Y TD	120 	140 16 P ELHDEPAP P DFPRQLAL D OFFEYLKD STWDW GYFESSHK GAEQEVEKFTKRLVKVTKQHNDE	0 L : 84 I : 99 A : 111 V : 111 C : 141
	XNI_ECO24 : DPOln_THEA : EXO5_BPT5 : RNH_BPT4 : FEN1_HUMAN :	180 RAAFE-QRGVPCWSTSGNENG LAAT AVKVTQA KELVO-LLGLARLEVPEYRG VLAS A'KAEKE FELCK-TT-PFTFTRGVENG MANY V LIGHT IDELKAYNPYIVHDIDKYENG HINV V KFSLF KHLLS-LMGIPYLDAP-SEA SSCALV AKG-	200 V GRIQATIVSTD KGYCOIL-SPT SGYEVRITAD & LYOIL-SDR JYDHWWISTD & NDTIL-TDK SGHKILISD C FTOHKYPM VYAAAHED C CTTGSF	220 24 RIRDYFOKRWIDAPF HULHPEGYLITPAW SRFSTRREYHLRD KOWSPMRKKWVKI VLMRHLTASEAKKLPIQEFHLSR	0 1 : 154 : 168 M : 180 K : 181 I : 212
	XNI_ECO24 : DPO1n_THEA : EXO5_BPT5 : RNH_BPT4 : FEN1_HUMAN :	2604 W DKEFG -OP OLPDYNG ACISSSK PG A WERYG -RPD WADYRA TOESS N PG K YEHHN DDV OFISIKA MOLG N RG E SGSA IDCMTKI KCDKK N AS KVRSI LQELG -NQ FVDLCI LSS VCES R	280 CICP SATO IVEF CICE TARK JEEM CICA ROYN IREF DFWFTRVEGERTPSMKTS VEAI CICP RAVD IQH GhG	300 -OS G YEN DAVAEK-WRKK -CS A LKN DRLKPA-IREK -CN LD IDO PLPGKQKYION ANDR QAKVLLT -KS E VRRDPNK-YP	0 E : 219 L : 233 N : 247 E : 245 P : 271
	XNI_ECO24 : DPOln_THEA : EXO5_BPT5 : RNH_BPT4 : FEN1_HUMAN :	340 THKEMAFLCRD AR-L-QTDLHIDGNLQQ AHMDDLKLS-WDLAKVRTDLPLEVDFAN ASEELIFRNLI VD-LPTYCVDAIAAVQQJ SEYNRYKENLV ID-FDYIPDNIASNIVNYYN E-NWLHKEAHQ FLEPEVLDPESVELKW	360 DLRLVR	380 40 HEFGLLESPKALEEAPWPPPEGA TNSINEF IRSGVKRLSKSRQGSTQGRLDDF	0 - : 251 F : 306 - : 291 - : 305 F : 344
		20	40	60 8	0
В	XNI_ECO24 XNI_KLEP7 XNI_YERPA XNI_VIBCH		THAVQGSP THAVQGSP THAVQGSP THAVQGSP VHSAQPDPNDI	-CVETCQHALDQLIMHSQPTHAV -CVDTCLHALEQLIVHSQPTHAV -CVKACQHALQQLIQHSQPSHAV RTAETTTTLQRIINEAQPSHMI	A: 47 A: 47 A: 47 A: 52
	B.subtilis S.aureus DPO1n_THEA EXO5_BPT5	MNNNKULLVDGUAD FR MPN-KILLVDGUAD FR MRGMLPLFEPKGRVLLVDGHHDAYR MSKSWGKFIEEEEAEMASRRNLMIVDGTNDGFR	AFFATAVHRNFUINDSCVP HFYATSLHKQFMYNSQCVP TFHALKGUTTSRGEP FKHNNSKKP	NGVNGFLKHLITAVETFOPTHV NGIQGFVRHIFSATHEIRPTHVA QAVYGFAKSLLKALKEDG-DAV -FASSYVSTIQSLAKSYSARTTI	C: 61 V: 60 V: 64 V: 65
	XNI_ECO24 XNI_KLEP7 XNI_YERPA XNI_VIBCH	100 Vederresse References Vederresse References Vederress References Vederress References Vederress References Vederresse References	120 	140 166 PCWSTSCNEADD AACLAVKVTQ RCWASPCSEADD AACLAVKVTQ ACHHSPCNEADD AACLAVKVAQ DSLLSEODSADD AACLVVKVAG	0 A : 117 A : 117 A : 117 H : 122
	B.subtilis S.aureus DPO1n_THEA EXO5_BPT5	C DMGSKTYRNDLFQDYKINRSAPPYELIPO C DMGOSTRNDMFDCYKONRSAPPELIPO VEDAKAPSERHBAYGYKICRAFTEEDFPRO LCDKGKSVERLEHLPEYKGNRDEKYAQRTEE	FDLA BAABLC FDVKEISEOFG LALIKELVDLLG EKALDEQFFEYLKDAFELCKTF	MNIGFACYBADDCIGTLADLFAN VNIGVRNYBADDVIGTLAQYST ARLEVPCYBADDVIASLAKKAEK PTFTIRCVBADDVASYIVKLIGH	E : 129 D : 128 E : 132 L : 143
	XNI_ECO24 XNI_KLEP7 XNI_YERPA XNI_VIBCH	180 GHQATIVED GKCDISPTRIRDYFQK GHQATIVED GKCDISPTRIRDYFQK GHQTIVED GKCDISPTRIRDYFQK QEKUTIETD GKCDISPTROTRDYFQC	200 WLDAPFIDKEFGVQ-PQQIPD WLDAPFIASEFGVT-PEQIAD WLDMPFVKOEFGVL-PRQIPD WLDEPFIAQEFGVT-PAQLID	220 24 WG JACISS SKV FOVACIOPKSAT WG JACISS SKV FOVACIOPKSAA WG JACISS SKV FOVACIOPKSAA WG JACISS SKI FOVACIOPKA K	0 Q : 192 Q : 192 L : 192 E : 197
	B.subtilis S.aureus DPO1n_THEA EXO5_BPT5	A-DITVYEDEDLDLLDKYSVALLQKGIGNY N-DYYIIGDYDLLCINDNYEVWLIKKGFNIY GYEVRIITADXDLYCLSDRHHVHPEG YDHVWLISTDCDMDTLLTDKYSRFSFTTR	VYTKETFYBETGVM-PKALID NRYTLHRFNBEVALE-PQOLID YLTTPAWLWEKYGLR-PDOWAD EYHLRDMYBHHNVDDVEQFISI	KA MCD SSDNYD GVKGI GEKIAY KAFMCD ADGYAGYKGI GEKIAI RA TCDESDNI POVKGI GEKIAR KAIMCDLGDNI RGVEGI GAKRGY	K : 207 K : 206 K : 206 N : 219
	XNI_ECO24 XNI_KLEP7 XNI_YERPA XNI_VIBCH	260 IIV FO TAG YEN DA AEKW KITETHKEMAI IN BODI GUYAR AE PEKW KITAARGEMAI IO CADTU YON ES PECO IIN IS TEAVAS - PE PAKY KITOPHIEMAI	280 FLCRD AR O D H DGNLQUE FTCRE AR O D O DGNLQUE FTCRE KR GDD G DGNLQUE RICKO SATK DIE GFNLQUI	300 32 LVR LTR	0 - : 251 - : 251 - : 230 - : 259
	B.subtilis S.aureus DPO1n_THEA EXO5_BPT5	LIREVE-TOR LEN SL PRGOOG TOOGLSDLI LIQOYOSVEN VEN DA SAGQENKINDNLDEL LLER GSTALLKN DR KPAIRE ILAHMDDLI IIRE GNULD IDO PLPGKOKYIQNLNASEEL	EMSRKUAEIHCSVPUACTLKI YLSKRUAEIHOVPIDSDALFE KLSWDUAKVRUDIPUEVDFAKR LFRNLUVDLPTYCVDAIAAVG(	ALFTLQMEQAADMLRRHQIKGIE MSFATTLNHILSICNEHELHVSG EPDRERLRAFLERLEFGSLLHEF DVLDKFTKDILEIAEQ	R : 285 K : 286 G : 286 - : 291

## Figure S1. Sequence alignments of E. coli ExolX against other FENs

(A) Alignment of ExolX (XNI\_ECO24) against the N-terminal domain of *Thermus aquaticus* DNA polymerase (DPO1n\_THEA, N terminal domain of Uniprot DPO1\_THEAQ), Bacteriophage T5 FEN (EXO5\_BPT5), Bacteriophage T4 RNase H (RNH\_BPT4) and human FEN-1 (FEN1\_HUMAN). Black background indicates residues conserved across all these sequences, gray are partially conserved. Red boxes indicate Cat1 site residues, and blue boxes indicate the Cat2 site residues.

**(B)** Alignment of ExoIX (XNI\_ECO24) against the Xni proteins from *Klebsiella pneumoniae* (XNI\_KLEP7), *Yersinia pestis* (XNI\_YERPA), *Vibrio cholerae* (XNI\_VIBC), *B. subtilis, Staphylococcus aureus*, the N-terminal domain of *Thermus aquaticus* DNA polymerase (DPO1n\_THEA) and bacteriophage T5 FEN (EXO5\_BPT5). Black background indicates residues conserved across all these sequences, gray are partially conserved. The first four sequences are ExoIX-like Xni proteins as they lack the three Cat2 site aspartates (red letters in blue boxes).



#### Figure S2. Oligonucleotide sequences and activity

(A,B) Sequences of Dup1 and Dup2 used in fluorescence anisotropy studies, orange circle indicates position of fluorescein label.

(C) Sequence of Flap1 used in fluorescence anisotropy studies and in initial co-crystallization study. Orange circle indicates position of fluorescein label.

(D,E) Sequences of palindromic oligonucelotides 5ov4 and 5ov6, respectively, used in the later co-crystallization studies. The position of the twofold symmetry axis is indicated by a filled black ellipse. (F) ExoIX displays nuclease activity on a flap substrate: fluorescently labelled Flap1 substrate (Figure S2C) was incubated with either T5 exonuclease (T5), wild type ExoIX (ExoIX WT) or ExoIX Lys67Ala mutant (ExoIX K67A) in the presence or absence of magnesium cofactor as indicated for 3hrs (left panel) or overnight (right panel). Substrate and product were analyzed by PAGE and visualized using a Fuji phosphorimager.



### Figure S3. Stereodiagrams of electron density maps.

(A) View of the initial  $F_{obs}$ - $F_{calc}$  difference map for the ExolX:Flap1 complex showing density (blue) for the DNA duplex bound to the ExolX molecule. Only the protein had been built into the model at this stage. The map is contoured at 2.5  $\sigma$ . The ExolX molecule is represented as a green cartoon with the H3tH motif highlighted in yellow and the beta sheet in orange.

**(B)** Stereo version of Figure 2E.  $2F_{obs}$ - $F_{calc}$  map (grey, contoured 1  $\sigma$ ) at showing the electron density in the region of the K<sup>+</sup> ion (purple sphere) and its coordinating main chain carbonyls and DNA phosphate group in the native ExolX:50v6 complex structure. A  $F_{obs}$ - $F_{calc}$  simulated annealing omit map (with K<sup>+</sup> and DNA excluded from the refinement) contoured at 5  $\sigma$  is shown in green. The protein carbon atoms are shown in green, the DNA carbons in yellow. Nitrogen, oxygen and phosphorus atoms are colored blue, red and orange. Metal - ligand interactions are shown as dotted lines.

(C) Stereo version of Figure 4A.  $F_{obs}$ - $F_{calc}$  simulated annealing omit map (green positive, red negative density, contoured at 3  $\sigma$ ; Mg<sup>2+</sup> and nearby oxygens omitted from refinement) of the Cat1 site in the ExoIX:5ov4:Mg<sup>2+</sup> complex showing the coordination of the two Mg<sup>2+</sup> ions (blue spheres). An  $2F_{obs}$ - $F_{calc}$  map (grey, contoured at 1  $\sigma$ ) is also shown. Other colors as Figure S3B.

(D)  $2F_{obs}$ - $F_{calc}$  map for the K<sup>+</sup> site in the ExoIX:5ov4:Mg<sup>2+</sup> complex (grey density, contoured at 1  $\sigma$ ) and  $F_{obs}$ - $F_{calc}$  simulated annealing omit map in green (metal ions and nearby water molecules omitted, contoured at 5  $\sigma$ ). The K<sup>+</sup> ion (0.4 occupancy, purple sphere) is partially displaced by the Mg<sup>2+</sup> ion (0.6 occupancy, blue sphere).

(E)  $2F_{obs}$ - $F_{calc}$  map (grey. contoured at 1  $\sigma$ ) and simulated annealing omit map (green, contoured at 5  $\sigma$ ) for the K<sup>+</sup> site for the ExoIX:5ov4:Ca<sup>2+</sup> complex. The K<sup>+</sup> ion (0.4 occupancy, purple sphere) is partially displaced by the Ca<sup>2+</sup> ion (0.6 occupancy, green sphere).



## Figure S4. Schematic diagram of the interactions between ExoIX and the 5ov6 substrate.

Deoxyribose rings are represented as pentagons containing the residue number and the phosphodiester linkages as circles enclosing the letter P. Hydrogen bonds are shown as black (intra-DNA) or blue (protein-DNA) dotted lines. The crystallographic two-fold axis is represented by a filled black oval. The interactions between the reference ExoIX molecule and the DNA duplex are shown on the left, the identical ones from the other symmetry-related ExoIX molecule in the crystal are not shown.

		MID Phasing		Othor Nativ	Ctructures		DNA Comple	Ctructure of	
Dataset (pdb code)	Native	HgSO4	Au(CN) <sub>2</sub>	K <sup>+</sup> bound	K <sup>+</sup> Free	Flap1	$50v4 + Mg^{2+}$	$50v4 + Ca^{2+}$	50v6
Space Group	رonze) 1ل	Pl	P1	(эхиу) Р2 <sub>1</sub>	(Jozue) C2	( <b>32ua)</b> P2 <sub>1</sub> 2 <sub>1</sub> 2	( <b>JZUD)</b> P2 <sub>1</sub> 2 <sub>1</sub> 2	( <b>32uc</b> ) P2 <sub>1</sub> 2 <sub>1</sub> 2	( <b>ɔzuu)</b> P2 <sub>1</sub> 2 <sub>1</sub> 2
$\alpha$ (°) $\alpha$ (°)	43.4 110.4	43.4 110.6	43.8 110.9	53.5 90.0	128.5 90.0	66.2 90.0	66.2 90.0	66.2 90.0	66.4 90.0
c (Å) β (°) β (°)	56.8 95.3 60.4 94.9	56.5 95.4 60.2 95.0	57.3 95.4 60.7 95.1	58.5 108.0 59.7 90.0	37.4 117.7 66.7 90.0	154.8 90.0 34.5 90.0	152.3 90.0 34.5 90.0	152.7 90.0 34.5 90.0	150.0 90.0 34.2 90.0
Resolution (Å) <sup>b</sup>	20.0 - 2.0	20.0 - 2.50	37.0 - 2.80	40.0 - 2.00	40.0 - 2.45	51.64 - 1.50	33.67 - 1.47	40.00 - 1.53	31.15 - 1.50
Danmlina a	(2.05-2.00) חו פע 14 1	(2.64-2.50) Shee	(2.95-2.80) SHEE	(2.05-2.00) SHEE	(2.58-2.45) Shee	(1.54-1.50)	(1.51-1.47)	(1.57-1.53)	(1.54-1.50)
Detector	CCD	MAR 345	MAR 345	MAR345	MAR345	ADSC 0315 CCD	ADSC 0315 CCD	ADSC 0315 CCD	ADSC 0315 CCD
Wavelength (Å)	0.870	1.542	1.542	1.542	1.542	0.976	0.980	0.980	0.976
Observed/unique refls	82,213/34,557	23,621/15,496	24,719/12,138	41,565/14,512	18,686/9,692	571,962/57,905	246,258/60,541	223,277/53,386	307,026/54,849
$K_{merge}(\mathcal{Y}_0)^{-}$	4./(30.8)	14.2 (36.4) 84 6 (84 6)	00 5 (85 7)	9.1 (36.9)	02 / (95 0)	9.7 (47.7)	00 0 (23.8)	00 3 (07 0)	(0.1C) C./
$I/\sigma^{b}$	8.8 (2.1)	4.9 (1.2)	6.2 (1.5)	11.4 (4.5)	3.8 (1.4)	22.5 (3.4)	15.2 (4.7)	13.7 (3.2)	12.6 (3.5)
[Heavy atom] (mM)	1	2	2	۲.	1	۲.	1	۲.	•
No. of sites	ı	2	4		ı				ı
$\frac{R_{iso}}{R_{iso}}$		13.6	13.7	,	·				
Phasing power 3.0A Routis 3.0Å		0.77	0.60						
Refinement Statistics									
Number of reflections (working/free sets)	31,471/1,708	ı	ı	13,356/714	8,558/474	54,940/2,950	54,391/3,074	47,991/2,697	51,998/2,783
Number of atoms (protein/other)	3,726/249			1,947/100	1,895/17	1,915/206/257	1,918/245/211	1,901/245/170	1,957/206/229
Rwork/Rfree	0.160/0.227			0.159/0.246	0.170/0.275	0.211/0.235	0.201/0.231	0.205/0.229	0.215/0.254
KMS deviation from ideality:	1 1 1 L 1			0 01 c Å	r vuo r	0 01 5 Å	ړ د د ب	8 610 0	1 010 Å
Bond angles	0.0107			0 CL U.	1 07 °	1 550	1 1/0	1 /2º	0.010 A
Dihadral angles	0 0 O C	1	1	11 6 0	20 7 0	10.20	10.70	1.4.0	12.02
Pfactore of honded atoms - overall	$133 Å^2$			$160 Å^2$	$11 \leq \lambda^2$	10.J ΑΑ <sup>2</sup>	10.2 C 0 Å 2	۲.1 د م گر 2	10.82
main chain	$67 Å^2$	ı	ı	$8 4 Å^2$	$10.7 ^{32}$	$4 \ 1 \ \text{\AA}^2$	0.0 A	$3 \wedge k^2$	$1.7 \Lambda$ $1.4 \lambda^2$
side chain	$21.2 \text{ Å}^2$	ı	ı	$25.3 \text{ Å}^2$	$12.3 \text{ Å}^2$	$9.0 \text{ Å}^2$	$8.3 \text{ Å}^2$	$7.1 Å^2$	$2.2 \text{ Å}^2$
DNA	ı	ı	ı		ı	36.7 Å <sup>2</sup>	$32.7 \text{ Å}^2$	$18.6 Å^2$	$8.0 \text{ Å}^2$
Ramachandran plot. Proportion of residues in:									
allowed regions	98.3 %			98.8 %	97.5 %	99.2%	99.2%	99.6%	98.0%
additional allowed regions	1.7 %	ı	ı	1.2 %	2.1 %	0.8%	0.8%	0.4%	1.6%
disallowed regions	0.0%	ı	ı	0.0 %	0.4 %	0.0%	0.0%	0.0%	0.4%
Average B-factors (main chain/side chain)	32.2/51.1 Å <sup>2</sup>			22.8/43.6 A <sup>2</sup>	46.5/52.0 Å <sup>2</sup>	25.4/31.5 Å <sup>2</sup>	22.9/28.0 Å <sup>2</sup>	25.0/30.8 A <sup>2</sup>	23.8/26.7 Å <sup>2</sup>
Average B-factors of DNA				· · ·		$82.4 \text{ A}^2$	$65.7 \text{ A}^2$	58.5 A <sup>2</sup>	$65.7 \text{ A}^2$
Number of waters/Average B-factor	247/47.5 Å <sup>2</sup>	ı	ı	100/35.0 Å <sup>2</sup>	17/42.0 Å <sup>2</sup>	243/34.8 Å <sup>2</sup>	203/31.1 Å <sup>2</sup>	$159/33.1 \text{ Å}^2$	220/35.6 Å <sup>2</sup>
Number of metals/Average B-factor	2/26.4 Å <sup>2</sup>	ı	ı	$1/19.9 \text{ Å}^2$	ı	2/32.0 Å <sup>2</sup>	$4/19.2 \text{ Å}^2$	3/25.1 Å <sup>2</sup>	1/22.6 Å <sup>2</sup>
<sup>a</sup> SHEF, on a MAR345 detector / Rigaku Micromax 007 rotati	ing anode generator at	University of Sheffield	; DL, at CCLRC	R	cullis = < rms lack of clo	sure>/ <rms di<="" isomorphous="" td=""><td>ference&gt; for acentric reflect</td><td>ions. (values for centric relec</td><td>ctions given in</td></rms>	ference> for acentric reflect	ions. (values for centric relec	ctions given in
$R_{merge} = \Sigma F_{nu} - F_{\nu}/\Sigma F_{\nu}$ , where $F_{vu}$ and $F_{\nu}$ are the derivative and $R_{merge} = \Sigma F_{nu} - F_{\nu}/\Sigma F_{\nu}$ , where $F_{vu}$ and $F_{\nu}$ are the derivative and $R_{\nu} = 2F_{nu} - F_{\nu}/\Sigma F_{\nu}$ .	d native structure facto	r amplitudes.		H, od	a endeses.) alues for the outermos $t_{work} = \Sigma   F(obs)  -  F(ca)$	t resolution shell are given i lc)∥/∑ F(obs)  for the 95% c	n parentheses. f the reflection data used in r	efinement	
		allon maca.						CITICITY CITY	

ı

Table S1. Data collection, MIR phasing and refinement statistics.

Phasing power = <rms heavy atom structure factor>/<rms lack of closure> for acentric reflections (values for centric relections given in parentheses.)

 ${}^{d}R_{free} = \Sigma ||F(obs)| - |F(cale)||/\Sigma|F(obs)|$  for the remaining 5 % of the reflection data excluded from the refineme