PNAS
www.pnas.org

Supplementary Information for

The crystal structure of dGTPase reveals the molecular basis of dGTP selectivity

Christopher O. Barnes[1,2]†, Ying Wu[1]†, Jinhu Song[3], Guowu Lin[1], Elizabeth L. Baxter[3], Aaron S. Brewster[4], Veeranagu Nagarajan[5], Andrew Holmes[6], Michael Soltis[3], Nicholas K. Sauter[4], Jinwoo Ahn[1], Aina E. Cohen[3*] and Guillermo Calero[1*]

Co-correspondence should be addressed to Aina E. Cohen and Guillermo Calero
Emails: acohen@slac.stanford.edu, guc9@pitt.edu

This PDF file includes:

Supplementary text
Figs. S1 to S5
Tables S1 to S4

Supplementary Methods

[1] Department of Structural Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15260, USA.
[2] Department of Pharmacology and Chemical Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15260, USA
[3] Macromolecular Crystallographic Group, Stanford Synchrotron Radiation Lightsource, National Accelerator Laboratory, Stanford University, Menlo Park, CA 94025, USA.
[4] Molecular Biophysics & Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.
[5] JAN Scientific, Inc, 4726 11th Ave NE, Ste 101, Seattle, WA, 98105, USA.
[6] Swanson School of Engineering, University of Pittsburgh, Pittsburgh, PA 15260, USA.

Additional Footnotes: †These authors contributed equally to this work.

Design and Fabrication of Multi-Crystal Holders (MCHs)

MCHs were developed to meet the following requirements: 1) minimal UV-background, 2) a larger size than commercially available mounts to hold substantial amounts of crystals, and 3) compatibility with the SAM robot system for robotic exchange during the diffraction experiment. Initial MCH prototypes were fabricated at the University of Pittsburgh, Swanson School of Engineering using a commercially available Stereo Lithography Machine, a 3D Systems Viper High Resolution SLA and Somos 11122XC Resin. The pointed end of the diamond shaped MCH was necessary to penetrate a crystal containing drop and mount crystals without drop displacement. Despite the success of the first-generation 3D-printed MCHs, subsequent generations of MCHs were manufactured by micron laser technology (Oregon, USA) from Mylar sheets (McMaster Carr, IL, USA) of 50 μm thickness (Fig. S1A, B). The laser-cut MCHs had better defined reference features (fiducial marks for crystal positioning), and were resilient to the physical manipulation during the crystal loading process. MCHs were affixed using epoxy to the end of a standard Hampton Research base-pin assembly for robotic exchange (Fig. S1B).

Macro parameters for detecting crystals on MCHs

UV microscopy images were loaded into *ImageJ* and brightness/contrast features were adjusted to highlight UV positive regions (Fig. 1B and Figs. S1, S2). Images were subsequently processed using a macro developed in our laboratory. Sample input parameters for our macro are listed below:

```
run("Smooth");
```

```
run("Threshold", "method=Huang ignore_black white setthreshold");
setThreshold(40, 900);
setOption("BlackBackground", false);
run("Erode");
run("Analyze Particles...", "size=150-10000 circularity=0.25-1.00 show=Outlines
display summarize in situ");
```

Modifications to brightness threshold, particle size, and circularity are critical to the selection of bright areas that indicate UV-visualized crystals, which are converted into crystal profiles (Fig. S2C) and a list specifying crystal size and centroid position (in x-y jpeg pixel coordinates, see Table S1). The pixel coordinates of the four MCH reference points were identified in relation to the crystal coordinates (Fig. 1B and Fig. S1E, red dots) through examination of a corresponding brightfield image (Figs. S1C and S2A, D). For larger crystals (Fig. S2D-F), a library of masks was created to define spacing along a crystal to translate into unexposed crystal volumes. To utilize this feature, the macro was altered to allow for the overlay of the mask as follows:

```
run("Subtract...", "value=50");
run("Smooth");
run("Threshold", "method=Huang ignore_black white setthreshold");
setThreshold(30, 255);
setOption("BlackBackground", false);
run("Make Binary", "thresholded remaining");
```

The resulting image was then overlaid with the mask (Fig. S2E, inset) followed by:

```
run("Make Binary", "thresholded remaining");
run("Analyze Particles...", "size=40-2000 circularity=0-1.00 show=Outlines
display summarize record in situ");
```

The minimum size was set to the same size (in pixels) as the mask size and circularity began at 0 since the mask creates square shapes. Coordinates for each distinct beam position were saved with an angular offset for helical data collection.

Supplementary Figures and Legends

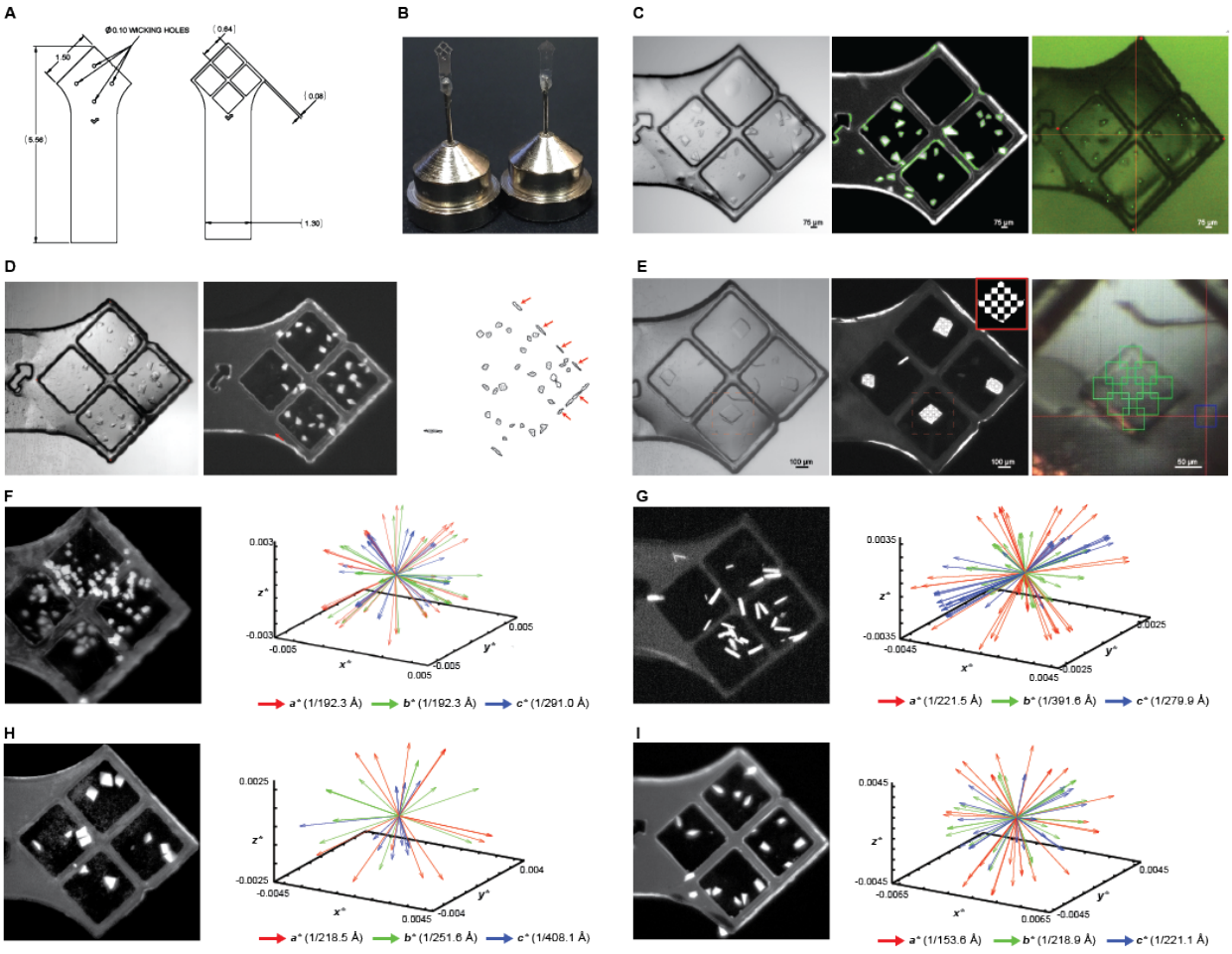


Figure S1. Design and identification of crystals mounted on Multi Crystal Holders. A) AutoCAD DXF schematic illustrating the dimensions of MCHs for laser printing onto Mylar sheets. **B)** Assembly of MCHs achieved by epoxying laser-printed MCHs onto the pin-base aperture of a Hampton loop. **C)** Left and Middle Panels: Brightfield and UV fluorescence microscopy images of Pol-Spt4/5-DNA mounted crystals. Green outlines indicate the crystal positions determined after particle analysis in ImageJ. Right panel: *BluIce* Image of Pol-Spt4/5 crystals mounted at XPP endstation. Coordinates generated in ImageJ (see Table S1) were uploaded, and after identification of reference fiducial marks (red asterisks) and beam positions were assigned to each crystal (green boxes). **D)** Brightfield image, UV fluorescence image, and

crystal positions as revealed by ImageJ. Coordinates for incorrectly identified UV fluorescence signal on the edge of the MCH (red arrows, right panel) are manually removed from the final coordinate list before data collection. **E)** Large crystals of the Pol II – TFIIB – DNA complex illustrate the multi-shot strategy with user specified spacing along the crystals surface. **F-I)** Reciprocal space representation of the basis vectors of indexed crystals that were singularly exposed for **F)** *Ec*-dGTPase, **G)** RNA Polymerase II T834P variant, **H)** RNA Polymerase II – TFIIB – DNA, and **I)** RNA Polymerase II – Spt4/5 – DNA crystals.

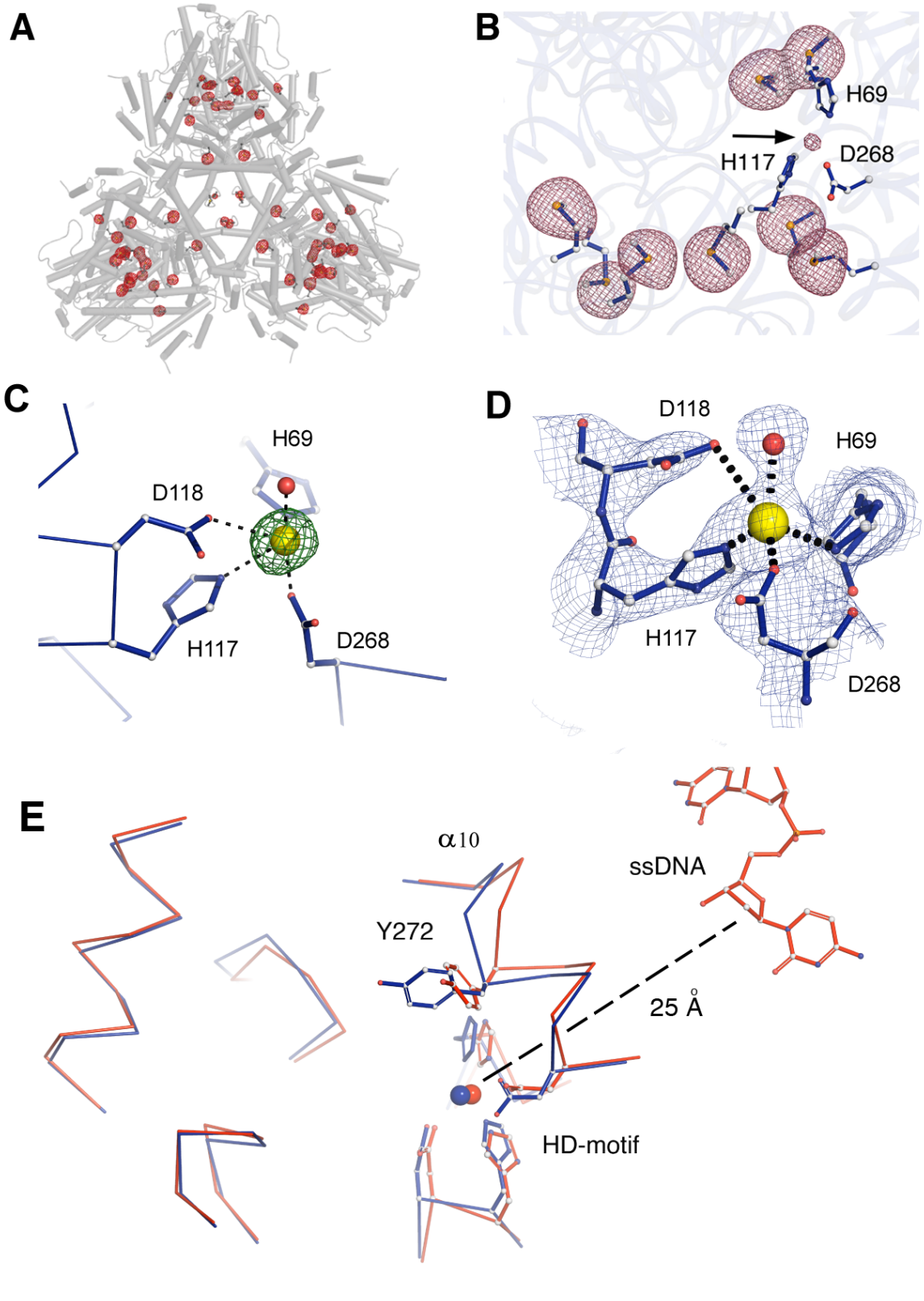
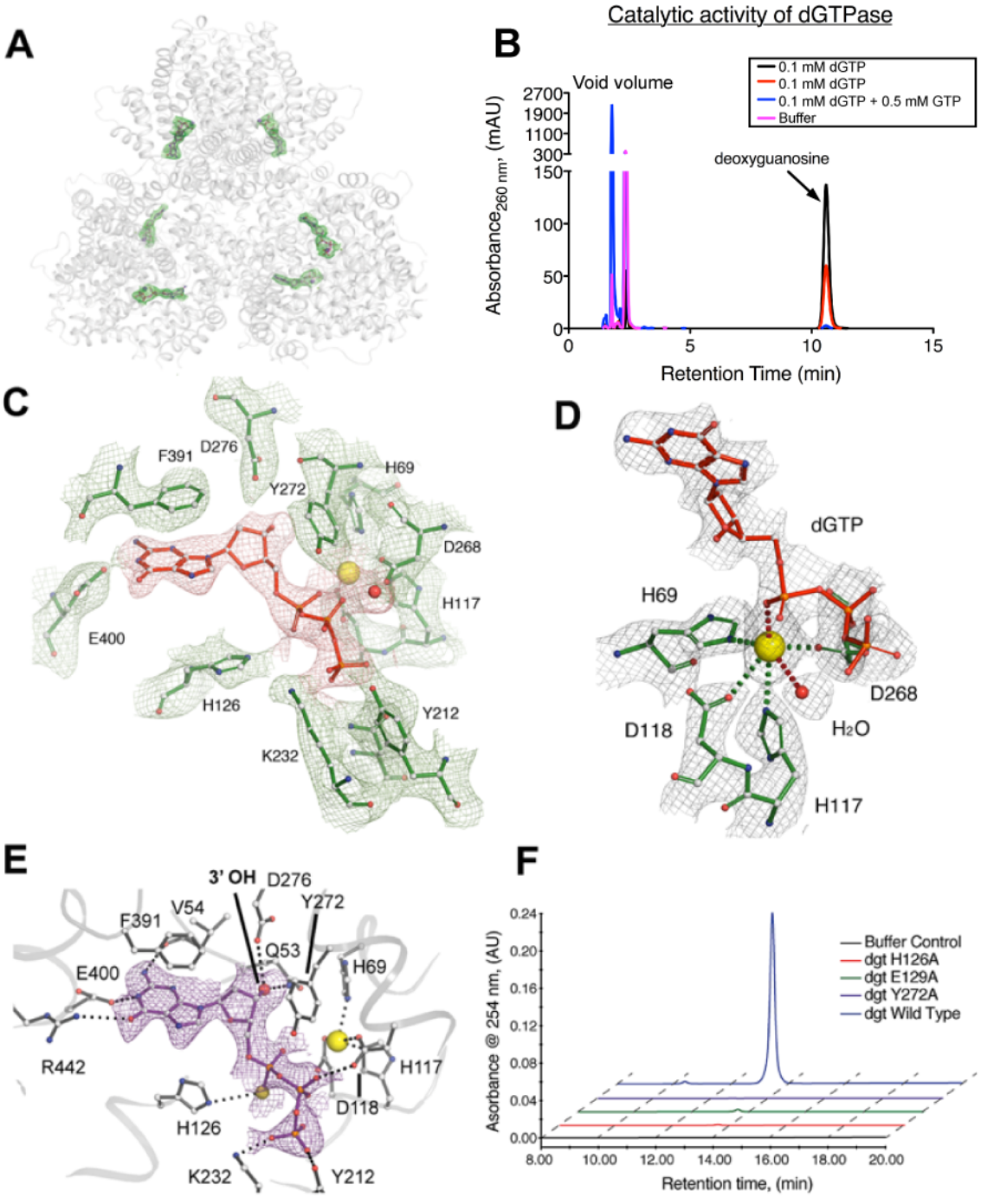

A
B
H69
H117
D268
C
H69
D118
H117
D268
D
D118
H69
H117
D268
E
α10
ssDNA
Y272
25 Å
HD-motif

Figure S2. Se-Met substructure and comparison of active site residues in apo- *Ec*-dGTPase structures. A) Tube representation illustrating the hexameric *Ec*-dGTPase structure determined by experimental phasing from selenium methionine labeled protein. The Se anomalous difference map is shown in red, contoured at 8σ. **B)** Anomalous difference map contoured at 5σ illustrating the presence of density for Mn^{2+} (arrow) at the active site. Anomalous density for Se atoms is shown for comparison. **C)** The F_o-F_c map countered at 6σ of the apo-XFEL structure calculated after removal of Mn^{2+} ions from the refined structure (modeled Mn^{2+} ion shown in yellow). **D)** $2F_o$-F_c composite omit map to illustrate electron density for a water molecule which forms part of the coordination sphere of Mn^{2+}. **E)** Overlay of the *Ec*-dGTPase apo-XFEL (blue) and ssDNA (red, PDB 4X9E) structures illustrating conformational differences mainly in the α10 helix involving Tyr^{272} due to the presence of the ssDNA located 25 Å away from the binding pocket (broken line).


A
B
Catalytic activity of dGTPase
0.1 mM dGTP
0.1 mM dGTP
0.1 mM dGTP + 0.5 mM GTP
Buffer
Void volume
deoxyguanosine
Absorbance260 nm, (mAU)
2700
1900
1100
300
150
100
50
0
0
5
10
15
Retention Time (min)
C
F391
D276
Y272
H69
D268
H117
E400
H126
Y212
K232
D
dGTP
H69
D268
D118
H2O
H117
E
3' OH
D276
V54
Y272
F391
Q53
E400
H69
R442
H117
D118
H126
K232
Y212
F
Buffer Control
dgt H126A
dgt E129A
dgt Y272A
dgt Wild Type
Asorbance @ 254 nm, (AU)
0.24
0.20
0.16
0.12
0.08
0.04
0.00
8.00
10.00
12.00
14.00
16.00
18.00
20.00
Retention time, (min)

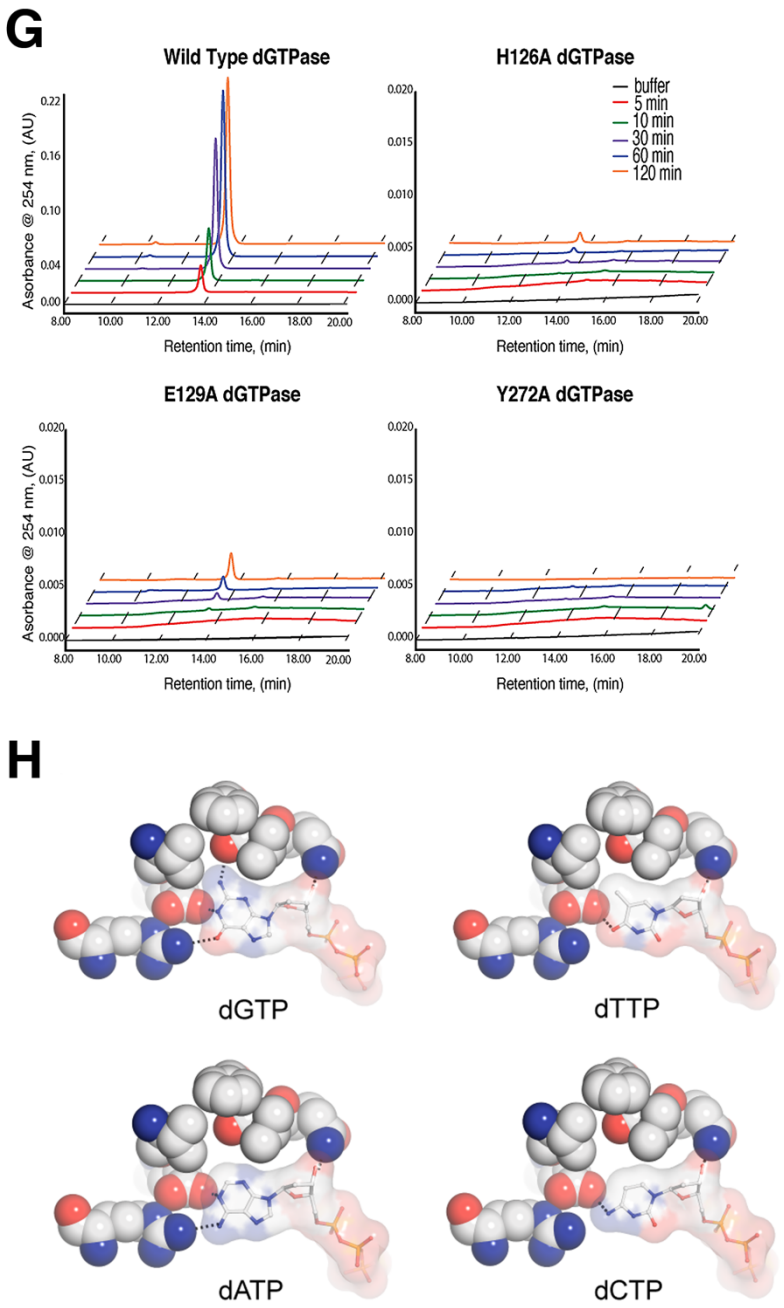


Figure S3. Nucleotide substrate binding, specificity, and *Ec*-dGTPase activity. A) Fo-Fc map countered at 3σ showing electron density corresponding to dGTP. The presence of the substrate was observed in every pocket of the hexamer (modeled dGTP shown). **B)** Catalytic activity of *Ec*-dGTPase crystals in the presence of 100 μM dGTP, illustrating that cross-linked crystals (red line) are still catalytically active when compared to wild-type (black line) or *Ec*-

dGTPase in the presence of 0.5 mM GTP (blue line). **C)** Final refined 2Fo-Fc map countered at 1.5 σ after applying negative B-factor sharpening =150. The quality of the map allowed full tracing of the structure. **D)** Final refined $2F_{obs} - F_{calc}$ map contoured at 2.0 σ and ball and stick representation of Mn^{2+} octahedral coordination by HD residues (forest green), dGTP and a water molecule (W1) are shown as ball and stick model and red sphere respectively.

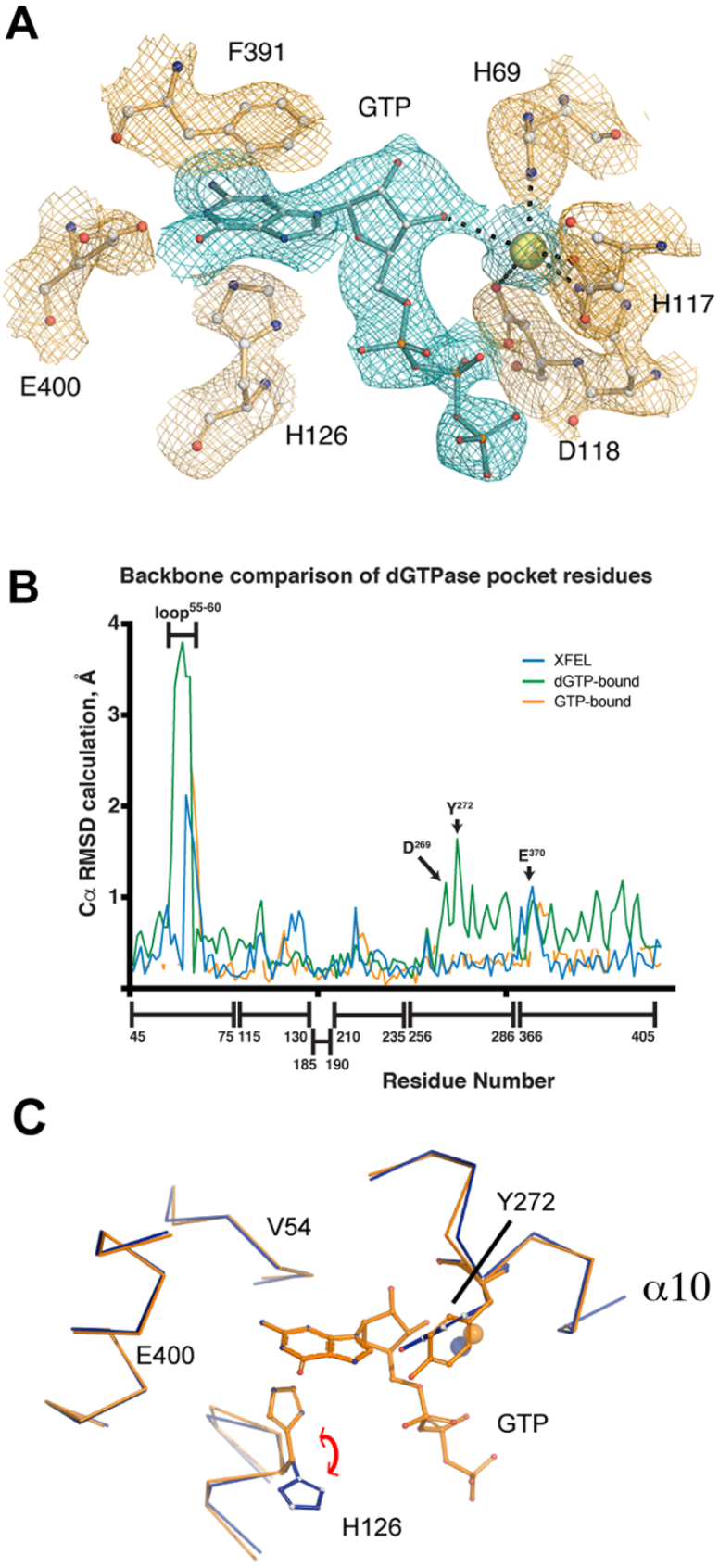
E) Ball and stick representation of residues involved in dGTP-1-thiol binding and final refined $2F_{obs} - F_{calc}$ map around dGTP-1-thiol contoured at 1.2σ. The overall and active site R.M.S.D. between the two substrates bound structures dGTP-1-thiol and dGTP is 0.4 Å and 0.3 Å respectively.

F) Enzymatic activity analysis after 2-hour incubation at room temperature in the presence of 100 μM dGTP substrate was performed by monitoring total deoxyguanosine product using reverse-phase chromatography for wild-type (blue), Y272A (purple), E129A (green) and H126A (red) dGTPase enzymes. The representative chromatogram indicates that the activity of dGTPase mutants is significantly decreased when compared to wild-type enzyme.

G) Enzymatic activity of individual dGTPase constructs. Representative reverse-phase HPLC chromatograms illustrating the amount of deoxyguanosine (dG) product attained after 5 (red), 10 (green), 30 (purple), 60 (blue), and 120 (orange) minutes of incubation at room temperature in the presence of 100 μm dGTP for wild-type, H126A mutant, E129A mutant, and Y272A mutant constructs. Wild-type enzyme is capable of producing an order of magnitude more dG product after 5 minutes (red line) than the mutant constructs after 120 minutes incubation (orange lines).

H) Comparison of dGTP with models of dTTP, dATP and dCTP binding to illustrate that active site residues can form four hydrogen bonds with the amide and ketone groups of dGTP but not with the other NTPs where one interaction is observed at best.

Figure S4. Structural differences related to inhibition of *Ec*-dGTPase. **A**) Final refined $2F_{obs} - F_{calc}$ of GTP bound to the active site pocket contoured at 1.2σ after applying negative B-factor



sharpening=160. **B.** Graphical representation of Cα RMSD illustrating: 1) the similarities between the apo and GTP structures (orange and blue traces) and 2) the conformational changes of the binding pocket residues triggered by dGTP binding (green trace). Positional differences for Cαs of Val^{54}, Asp^{268}, Tyr^{272} and Glu^{400}. The position of the catalytic His^{126} varies among the three structures illustrating its conformational flexibility. **C)** Overlay between the apo and GTP-bound structures showing the similar conformations between the two of them.

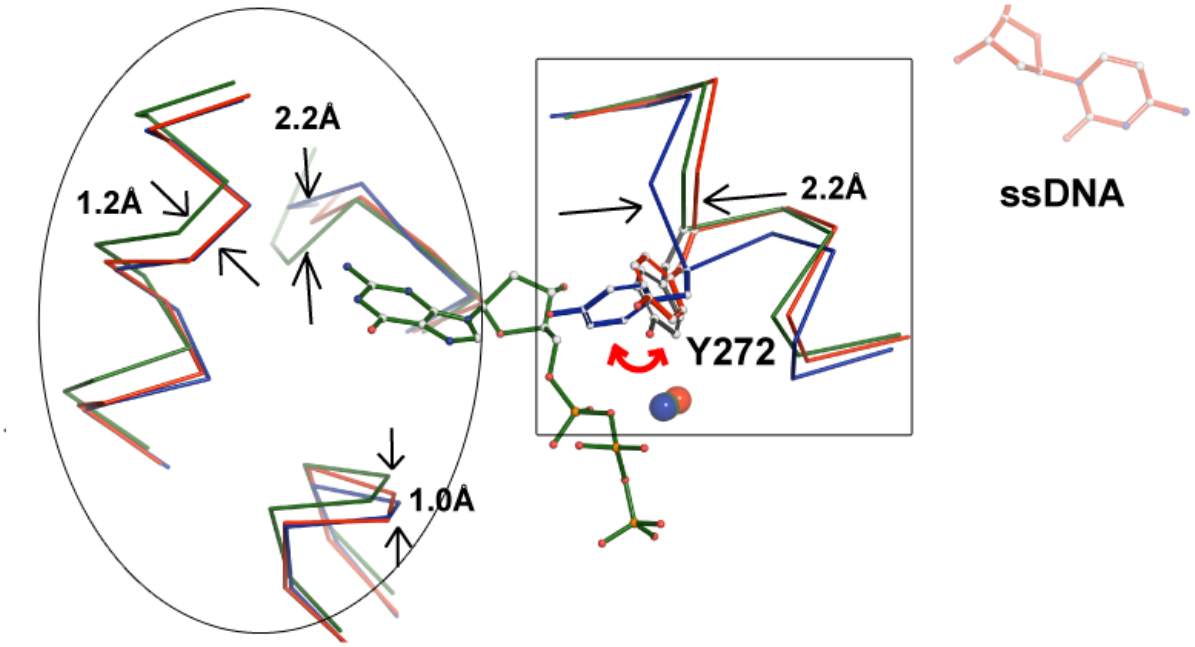


Figure S5. Overlay between *Ec*-dGTPase apo-structure bound to ssDNA (PDB 4X9E, red wire and red ball and stick model) with the apo-XFEL (blue wire) and the dGTP-bound structures (forest green wire). The ssDNA structure combines apo-like (overlapped red and blue traces, left ellipse) and substrate-bound like (overlapped green and red traces, right square) features. These conformational changes lead to a 20% increase in the size of the substrate binding pocket. The RMSD for Cαs between the structures is indicated between arrows.

Supplemental Tables

Table S1. Example of crystal coordinates and fiducial marks derived from ImageJ for MCH represented in Figure 1.

Fiducial Coordinates	*X-Coordinate*	*Y-Coordinate*	*Angular Offset*[a]
Position 1	844	36	0
Position 2	1430	690	0
Position 3	794	1258	0
Position 4	240	644	0

Crystal Coordinates	*X-Coordinate*	*Y-Coordinate*	*Angular Offset*
Crystal 1	672.8	217.9	0
Crystal 2	1045.2	258.0	0
Crystal 3	910.4	259.7	0
Crystal 4	397.9	310.3	0
Crystal 5	965.8	322.4	0
Crystal 6	616.2	342.1	0
Crystal 7	638.1	379.0	0
Crystal 8	1078.2	383.8	0
Crystal 9	747.7	413.5	0
Crystal 10	939.3	426.6	0
Crystal 11	1207.8	438.5	0
Crystal 12	298.2	438.0	0
Crystal 13	534.2	453.5	0
Crystal 14	328.4	480.2	0
Crystal 15	804.5	516.6	0
Crystal 16	761.4	531.3	0
Crystal 17	215.7	529.6	0
Crystal 18	652.6	560.7	0
Crystal 19	384.0	559.1	0
Crystal 20	977.7	607.3	0
Crystal 21	506.4	624.8	0
Crystal 22	360.8	621.6	0
Crystal 23	817.9	642.7	0
Crystal 24	324.7	710.5	0
Crystal 25	452.9	731.0	0
Crystal 26	1046.7	759.6	0
Crystal 27	841.8	746.8	0
Crystal 28	250.9	784.2	0
Crystal 29	696.7	816.8	0
Crystal 30	1226.5	812.9	0
Crystal 31	922.8	839.8	0
Crystal 32	411.3	862.5	0
Crystal 33	589.0	893.2	0
Crystal 34	1113.9	911.4	0

[a]For crystals capable of multiple exposures, and angular offset can be input to improve data completeness.

Table S2. XFEL data collection statistics and efficiency using MCHs

	dGTPase (LCLS - XPP)	*Pol II – TFIIB - DNA (LCLS - XPP)*	*Pol II – Spt4/5 – DNA (LCLS - XPP)*
Experimental Parameters			
Average Crystal Size (*a x b*, μm)	80 x 80	100 x 150	40 x 70
Total Number of Images	335	613	807
Total Hits (% of total images)	302 (90.1)	523 (85.3)	643 (79.7)
Indexed/Merged (% of total images)	221 (65.9)	421 (68.7)	594 (73.6)
Data Collection Statistics			
Space Group	$P4_32_12$	$P2_12_12_1$	$P2_12_12_1$
Resolution (Å)	25-3.2	25-3.5	25-4.2
Completeness (%)	97.9 (91.3)	91.3 (67.5)	95.8 (78.2)
Redundancy	7.3 (3.3)	7.8 (5.4)	9.3 (6.1)
$\langle I/\sigma I \rangle$	6.1 (0.8)	5.7 (0.4)	4.8 (0.3)
R_{split} (%)	37.3 (85.2)	36.9 (102)	34.9 (88)
Data Processing Program	cctbx.xfel	cctbx.xfel	cctbx.xfel

Table S3. Crystallographic data collection and refinement statistics

	dGTPase - SeMet (APS-GM/CA) Inflection	*dGTPase - XFEL (LCLS - XPP)*	*dGTPase-dGTP-1-thiol (APS-GM/CA)*	*dGTPase - GTP (SSRL 12-2)*	*dGTPase - dGTP (SSRL 12-2)*
PDB ID code	**6OI7**	**6OIV**	**6OIW**	**6OIX**	**6OIY**
Data Collection[a]					
Space Group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$
Unit cell (Å)	191.9, 191.9, 286.9	192.3, 192.3, 291.0	191.2, 191.2, 298.6	191.6, 191.6, 292.9	192.2, 192.2, 299.6
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Wavelength (Å)	0.968	1.304	1.0332	0.9798	0.9798
Resolution (Å)[b]	49-2.85	25-3.2	49-3.35	40-3.25	40.3.28
Unique Reflections	124,944	88,170	80,111	81,683	85,095
Completeness (%)	100 (100)	97.9 (91.3)	100 (100)	94.8 (86.3)	99.0 (98)
Redundancy	15.7 (15.5)	7.3 (3.3)	9.3 (9.1)	6.4 (2.2)	8.7 (2.5)
$CC_{1/2}$ (%)	99.4 (33.0)	82.8 (25.5)	99.0 (28.9)	99.8 (60.1)	94.9 (45)
$\langle I/\sigma I \rangle$	15.7 (1.0)	6.1 (0.8)	4.7 (0.9)	5.5 (0.7)	7.1(1.2)
Mosaicity (º)	0.12	0.5	0.2	0.35	0.3
R_{merge} (%)	27.7 (308)	64.7 (84.5)	29.9 (248)	33.7 (302)	25.2 (79)
R_{pim} (%)	7.5 (83.2)	37.3 (85.2)[c]	10.2 (86.5)	12.2 (94.1)	6.1 (75.6)
Data Processing Program	HKL2000	cctbx.xfel[d]	XDS/Scala	XDS/Scala	HKL2000
Refinement					
No. Atoms	27,701	27,352	25,059	25,047	26,560
R_{cryst}/R_{free} (%)	17.6/20.6	23.2/24.5	17.9/21.2	18.3/21.5	20.5 /24.1
Ramachadran plot					
Outliers	1.4	1.4	2.1	1.5	2
Allowed	4.6	6.6	6.9	4.5	5
Favored	94	92	91	94	93
r.m.s deviations					
Bond Length (Å)	0.007	0.01	0.02	0.02	0.013
Overall B ($Å^2$)	95	112	135	123	117
Refinement Program	Buster	Buster	Buster	Buster	Buster

[a]Numbers in parentheses correspond to the highest resolution shell
[b]Resolution limits were extended to include weak intensity data.

Table S4. Structural overall and active site comparison R.M.S.D (Å) statistics

Overall	dGTPase - XFEL	dGTPase-dGTP	dGTPase-dGTP-1-thiol	dGTPase - GTP
dGTPase - XFEL	X	0.80	0.83	0.50
dGTPase-dGTP		X	0.30	0.83
dGTPase-dGTP-1-thiol			X	0.85
dGTPase - GTP				X

Active site	dGTPase - XFEL	dGTPase-dGTP	dGTPase-dGTP-1-thiol	dGTPase - GTP
dGTPase - XFEL	X	0.68	0.72	0.45
dGTPase-dGTP		X	0.40	0.70
dGTPase-dGTP-1-thiol			X	0.70
dGTPase - GTP				X