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

Christopher O. Barnes\textsuperscript{a,b,1}, Ying Wu\textsuperscript{a,1}, Jinhu Song\textsuperscript{a}, Guowu Lin\textsuperscript{a}, Elizabeth L. Baxter\textsuperscript{c}, Aaron S. Brewster\textsuperscript{d}, V. Nagarajan\textsuperscript{a}, Andrew Holmes\textsuperscript{e}, S. Michael Soltis\textsuperscript{f}, Nicholas K. Sauter\textsuperscript{g}, Jinwoo Ahn\textsuperscript{h}, Aina E. Cohen\textsuperscript{i,j}, and Guillermo Calero\textsuperscript{k,2}

\textsuperscript{a}Department of Structural Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15260; \textsuperscript{b}Department of Pharmacology and Chemical Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15260; \textsuperscript{c}Structural Molecular Biology, Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Stanford University, Menlo Park, CA 94025; \textsuperscript{d}Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720; \textsuperscript{e}JAN Scientific Inc., Seattle, WA 98105; and \textsuperscript{f}Swanson School of Engineering, University of Pittsburgh, Pittsburgh, PA 15260

Edited by Alexander M. Kilbanov, Massachusetts Institute of Technology, Cambridge, MA, and approved March 22, 2019 (received for review November 26, 2018)

Deoxynucleotide triphosphohydrolases (dNTPases) play a critical role in cellular survival and DNA replication through the proper maintenance of cellular dNTP pools. While the vast majority of these enzymes display broad activity toward canonical dNTPs, such as the dNTPase SAMHD1 that blocks reverse transcription of retroviruses in macrophages by maintaining dNTP pools at low levels, \textit{Escherichia coli} (Ec)-dGTPase is the only known enzyme that specifically hydrolyzes dGTP. However, the mechanism behind dGTP selectivity is unclear. Here we present the free-, ligand (dGTP)-, and inhibitor (GTP)-bound structures of hexameric Ec-dGTPase, including an X-ray free-electron laser structure of the free Ec-dGTPase enzyme to 3.2 Å. To obtain this structure, we developed a method that applied UV-fluorescence microscopy, video analysis, and highly automated goniometer-based instrumentation to map and rapidly position individual crystals randomly located on fixed target holders, resulting in the highest indexing rates observed for a serial femtosecond crystallography experiment. Our structures show a highly dynamic active site where conformational changes are coupled to substrate (dGTP), but not inhibitor binding, since GTP locks dGTPase in its apo-form. Moreover, despite no sequence homology, Ec-dGTPase and SAMHD1 share similar active-site and HD motif architectures; however, Ec-dGTPase residues at the end of the substrate-binding pocket mimic Watson–Crick interactions providing guanine base specificity, while a 7-Å cleft separates SAMHD1 residues from dNTP bases, abolishing nucleotide-type discrimination. Furthermore, the structures shed light on the mechanism by which long distance binding (25 Å) of single-stranded DNA in an allosteric site primes the active site by conformationally opening a tyrosine gate allowing enhanced substrate binding.

Significance

While cellular dNTPases display broad activity toward dNTPs (e.g., SAMHD1), \textit{Escherichia coli} (Ec)-dGTPase is the only known enzyme that specifically hydrolyzes dGTP. Here, we present methods for highly efficient, fixed-target X-ray free-electron laser data collection, which is broadly applicable to multiple crystal systems including RNA polymerase II complexes, and the free Ec-dGTPase enzyme. Structures of free and bound Ec-dGTPase shed light on the mechanisms of dGTP selectivity, highlighted by a dynamic active site where conformational changes are coupled to dGTP binding. Moreover, despite no sequence homology between Ec-dGTPase and SAMHD1, both enzymes share similar active-site architectures; however, dGTPase residues at the end of the substrate-binding pocket provide dGTP specificity, while a 7-Å cleft separates SAMHD1 residues from dNTP.

Author contributions: C.O.B., Y.W., J.A., A.E.C., and G.C. designed research; C.O.B., Y.W., J.S., G.L., E.L.B., A.S.B., V.N., and A.H. performed research; C.O.B., A.H., and G.C. were responsible for MCH design; C.O.B., J.S., G.L., V.N., S.M.S., and A.E.C. were responsible for imaging and incorporation of data collection software at LCLS-XPP beamline; C.O.B., Y.W., E.L.B., A.S.B., N.K.S., and G.C. analyzed XFEL data; C.O.B., Y.W., and G.C. refined structures of the broadly acting tetrameric dNTPases, exemplified by SAMHD1, have enabled a better understanding of phosphodihydrolase activity by revealing the binding modes of activator or allosteric regulator nucleotides, as well as dNTP substrates (10–12, 16). However, unlike the vast majority of dNTPases displaying broad activity toward canonical and noncanonical dNTPs (11, 12, 14, 17, 18), Ec-dGTPase displays a strong preference for hydrolyzing only dGTP substrates (2, 10, 11, 14, 16, 18, 19). The structural basis of Ec-dGTPase specificity toward dGTP is less well understood despite recent crystal structures of a free and ssDNA-bound Ec-dGTPase (20). Indeed, the potential of ssDNA cofactors affecting the dGTP-binding site (20, 21) suggests a unique mechanism significantly different in the allosteric activation of phosphohydrolase activity compared with other dNTPase enzymes.

C

Cellular regulation of deoxynucleotide triphosphate (dNTP) pools is a vital process for DNA replication and survival in both prokaryotic and eukaryotic cells. It is tightly controlled by deoxynucleotide triphosphohydrolases (dNTPases) through hydrolyzing dNTPs into deoxyribonucleosides and inorganic phosphate, in contrast to ribonucleotide reductase, which supplies cellular dNTP pools by converting ribonucleotides to the corresponding deoxy forms (1–5). As a class of metalloenzymes, dNTPases contain a histidine–aspartate (HD) motif, which helps coordinate a divalent cation near the active site to promote phosphohydrolase activity (2, 7–12). Given the importance of intracellular dNTP concentrations for cellular survival, studies have shown that tight control of dNTP pools by dNTPases can function as a host mechanism in the cellular defense against pathogens. For example, in primates the sterile alpha motif (SAM) and HD domain-containing protein 1 (SAMHD1) (13) block reverse transcription of retroviruses (e.g., HIV-1, SIV) by maintaining dNTP pools at low levels in infected cells (7–10, 12, 14). In addition, observations in prokaryotes have revealed a dNTPase–pathogen interplay, illustrated by \textit{Escherichia coli} (Ec) dGTPase, where the gene 1.2 product encoded by bacteriophage T7 inhibits Ec-dGTPase to promote productive infection (15).
In the present work, we aimed to understand the structural basis of Ec-dGTPase substrate specificity and the mechanism of phosphohydrolase activity in the presence of nucleotide substrates. Given that Mn\(^{2+}\) is required for optimal dGTPase activity, which can be functionally replaced by Mg\(^{2+}\) (2, 18, 22), we determined the radiation-damage free crystal structure in the presence of Mn\(^{2+}\)cations to 3.2 Å using data collected at the Linac Coherent Light Source (LCLS) X-ray Free Electron Laser (XFEL). These experiments employed a methodology for highly efficient serial femtosecond crystallography (SFX) experiments using micrometer-sized crystals that builds upon conventional crystal visualization and fixed target diffraction techniques (23, 24). Our methodology combines (i) the use of specialized multicrystal holders (MCH) compatible with cryo-protectant or native crystallization conditions, (ii) the identification of individual microcrystals on MCHs through the analysis of UV-microscopy images, and (iii) the development of automated routines for serial positioning of the identified crystals during data collection. Implementation of these methods improved on similar reference and alignment schemes (25–31) enabling rapid mapping and positioning of multiple crystals in random locations on the holder, which consistently resulted in >80% crystal hit-rates and the highest indexing rates reported to date for any SFX experiment.

In addition, we investigated substrate specificity by applying chemical cross-linking methods to introduce nucleotide substrates into the catalytic site of Ec-dGTPase crystals and obtained dGTP-, dGTP-1-thiol-, and GTP-bound Ec-dGTPase structures. We found that structural elements from an adjacent monomer, which protrude into the enzymatic active site, are responsible for nucleotide discrimination and dGTP specificity. Analysis of catalytic residues and Mn\(^{2+}\) cations of free and bound Ec-dGTPase structures revealed a regulatory mechanism for Tyr\(^{272}\) that may explain activation or inhibition by ssDNA and GTP, respectively. Overall, these structures provide detailed mechanistic insights into Ec-dGTPase function and demonstrate a conserved binding mode for nucleotide substrates across dNTPase enzymes.

**Results**

**XFEL Data Collection of Microcrystals on MCHs.** To address the scarcity of XFEL beam time, we sought to improve the efficiency and reliability of data collection methods by simplifying related sample preparation and operational requirements for fixed-target setups. To this end, we developed an MCH (SI Appendix, Fig. S1 A and B) capable of holding microcrystals under native- or cryo-crystallization conditions that is compatible with UV fluorescence microscopy for crystal imaging and identification (Fig. 1; SI Appendix, Fig. S1; Methods). Given the high contrast of the UV-fluorescence signal of crystals versus the solvent background, bright areas corresponding to the location of individual crystals on MCHs were identified and mapped relative to four fiducial marks (Fig. 1B and SI Appendix, Fig. S1). Algorithms previously incorporated into the Blu-Ice/DCSS experimental control system to position grid-based sample holders and microfluidic traps at LCLS-XPP (23, 26, 27) were adapted to automatically position crystals based on their relationship to the four fiducial coordinates of the MCH for efficient serial diffraction experiments (Fig. 1B and SI Appendix, Fig. S1 D and E and Table S1). To demonstrate this technology, our protocol was applied to various crystal morphologies generated from either Ec-dGTPase or RNA polymerase II (Pol II) complexes mounted on MCHs (SI Appendix, Fig. S1) collected during the same 24-h XFEL shift. Our methods enabled diffraction data to be collected in an automated fashion with improved intensities for high-angle Bragg reflections, consistent with previous results using MCHs (32, 33).

In addition, we implemented multishot data collection strategies for crystals measuring >100 μm along a single axis (SI Appendix, Fig. S1 F–I). Similar to helical data collection protocols at LCLS, translation of ~50 μm into uncompromised crystal volumes with angular offsets from the origin was introduced to increase dataset completeness and limit the effects of radiation or mechanical damage introduced from previous exposures (23, 24, 34). To test if microcrystals displayed problems of preferential orientation on the MCH surface, we analyzed the reciprocal basis vectors for indexed Ec-dGTPase and Pol II crystals that were singly exposed for one MCH (SI Appendix, Fig. S2) and for the entire Ec-dGTPase dataset (Fig. 1C), using described methods (35). Our results revealed a spherical-like projection indicative of randomly oriented crystals on MCHs (23, 24, 26, 34). As a result, all datasets collected showed >90% completeness to high resolution (SI Appendix, Table S2), with only 221 still images being required to solve the structure of the free Ec-dGTPase enzyme and <1,000 images for Pol II complexes (further analysis of Pol II structures is beyond the scope of this work). Overall, application of MCH methodologies resulted in efficient data collection with minimal background scattering to preserve weak Bragg reflections, highly accurate microcrystal hit rates (>80%), increased indexing rates, and dataset completeness from a minimal number of exposed crystals (SI Appendix, Table S2).

**Comparison of Synchrotron and XFEL Ec-dGTPase Structures.** To gain structural insight into the effects of radiation-induced damage at the metal-binding site, we compared Ec-dGTPase structures determined from an individual crystal collected at a synchrotron source with our XFEL dataset. The synchrotron-based structure was solved from selenium–methionine derivatized protein to 2.9 Å resolution (Fig. 2A and SI Appendix, Fig. S24 and Table S3) and is consistent with previously published dGTPase apo-structures from *Pseudomonas syringae* [Protein Data Bank (PDB) ID code 2PGS], *E. coli* (PDB ID code 4XDS), and the *E. coli* structure bound to ssDNA, PDB ID code 4X9E). Interestingly, inspection of the anomalous map revealed the presence of density within HD-motif residues His\(^{99}\), His\(^{112}\), and Asp\(^{98}\) (SI Appendix, Fig. S2B). Given that the crystallization condition did not contain any anomalous scatterers, it was inferred that the observed signal was due to a metal bound to the protein during cell growth. A fluorescence scan of the crystals showed a peak at the Mn\(^{2+}\) edge (6.53 keV) but not at the Fe\(^{2+}\) edge (7.1 keV), suggesting the presence of Mn\(^{2+}\) ions in our crystals. This is consistent with kinetic experiments that showed Mn\(^{2+}\) (at 20 μM) as the preferred cofactor, despite similar kinetic activity achieved with 12 mM MgCl\(_2\) (8).
The radiation damage-free XFEL structure of the Ec-dGTPase apo-enzyme to 3.2 Å was solved from <150 crystals exposed to the XFEL beam (Fig. 2), representing the fewest number of randomly oriented microcrystals used to obtain a complete SFX dataset. Initial phases were generated by molecular replacement using the selenium–methionine phased structure. The initial unbiased sigma-A-weighted difference map for Mn$^{2+}$ coordination for XFEL (blue) and low-dose synchrotron structure (olive green) showed that Asp$^{233}$ residues (based structures with bound substrates and metals in all active sites) and Asp$^{B}$ were indicated as dashes between residues. (C) Electron density of the Sigma-A weighted $2F_{\text{obs}} - F_{\text{calc}}$ map contoured at 1.5σ for residues comprising the HD motif (synchrotron data); Mn$^{2+}$ ion is illustrated as a yellow sphere. (D) Electron density of the Sigma-A weighted $2F_{\text{obs}} - F_{\text{calc}}$ map contoured at 1.5σ for residues comprising the HD motif (XFEL data). A water molecule (indicated in red) forms part of the Mn$^{2+}$ coordination (SI Appendix, Fig. S2C).

Structural Basis for dGTP Binding and Phosphohydrolase Activity. Next, we wanted to characterize the substrate-bound form of Ec-dGTPase by determining the structure in the presence of dGTP and the nonhydrolyzable dGTP analog, dGTP-1-thiol. Initial cocystalization experiments of Ec-dGTPase with Mn$^{2+}$ ions and dGTP-1-thiol revealed no ligand-bound structures, likely due to the high-salt condition (1.8 M ammonium sulfate) from which the crystals were harvested. Thus, we employed chemical cross-linking using glutaraldehyde to stabilize the integrity of the crystal lattice, while decreasing the salt concentrations to physiological levels for overnight soaking experiments with dGTP or dGTP-1-thiol (Methods). Crystals remained stable during this procedure resulting in conventional synchrotron-based structures with bound substrates and metals in all active sites (SI Appendix, Fig. S3 A–E). Cross-linked Ec-dGTPase crystals also showed catalytic activity when incubated with dGTP, suggesting that the cross-linked enzyme was capable of binding and hydrolyzing the dGTP substrate, albeit in a slightly diminished capacity compared with unmodified enzyme (SI Appendix, Fig. S3B).

Comparisons between the substrate-bound and apo-structures show a root mean square deviation (rmsd) of 0.9 Å for the catalytic versus 0.3 Å for the noncatalytic regions of the protein. These differences are a result of conformational changes induced by dGTP binding, as active-site residues rearrange to provide a tight-fitting pocket for the substrate (Fig. 3 A and B). Critical residues involved in dGTP binding, as well as dGTP-1-thiol binding, include (i) π-π stacking interactions between the electron-rich aromatic ring of Phe$^{369}$ and the guanosine ring; (ii) hydrogen bonding between Arg$^{353}$, Arg$^{342}$, Glu$^{40}$, and Val$^{54}$ and the guanosine ring; (iii) stacking interactions between Tyr$^{277}$ [a highly conserved tyrosine found within dNTP triphosphohydrolases (14, 18, 37)] and the deoxyribose ring; (iv) 3′-OH discrimination via Gln$^{186}$ and Asp$^{276}$; (v) hydrogen bonds (in green) structures showing the conformational changes observed in the substrate-free structure and the second one positioned in-line with the hydroxyl group of the dGTP α-phosphate (Fig. 3D). This latter conformation is consistent with the proposed nucleophilic substitution reaction mechanism for dNTP hydrolysis via a nearby water molecule (35).

Owing to the contrast in side-chain confirmation for free- and substrate-bound structures, as well as the observed “sandwiching” between Tyr$^{277}$, His$^{126}$, and the dGTP ligand (Fig. 3C), we mutated

![Image](http://example.com/image.png)

Fig. 2. The hexameric Ec-dGTPase XFEL crystal structure. (A) Cartoon representation of the hexameric Ec-dGTPase XFEL crystal structure solved using 221 still images from randomly oriented crystals. (B) Differences in Mn$^{2+}$ coordination for XFEL (blue) and low-dose synchrotron structure (olive green). Potential H-bond interactions with distances shorter than 3.5 Å are indicated as dashes between residues. (C) Electron density of the Sigma-A weighted $2F_{\text{obs}} - F_{\text{calc}}$ map contoured at 1.5σ for residues comprising the HD motif (synchrotron data); Mn$^{2+}$ ion is illustrated as a yellow sphere. (D) Electron density of the Sigma-A weighted $2F_{\text{obs}} - F_{\text{calc}}$ map contoured at 1.5σ for residues comprising the HD motif (XFEL data). A water molecule (indicated in red) forms part of the Mn$^{2+}$ coordination (SI Appendix, Fig. S2C).

![Image](http://example.com/image.png)

Fig. 3. Interactions of dGTP substrate with Ec-dGTPase. Wire (A) and surface (B) representations of the overlay between the apo- (blue) and dGTP-bound (green) structures showing the conformational changes observed in the active-site pocket upon dGTP binding (red spheres). Red arrows indicate contraction, and black arrows indicate expansion of the pocket. (C) Ball-and-stick representation of key residues (green) involved in dGTP (red) binding. Hydrogen bonds are illustrated as dashes and water molecules as red spheres. (D) Ball-and-stick representation of Mn$^{2+}$ (yellow sphere) coordination by dGTP residues (green). Distances are indicated next to dashes. The position of the apo-XFEL Mn$^{2+}$ ion and coordinating water after overlay with the dGTP bound-structure are indicated as blue spheres.
these residues to alanine to identify their roles in phosphohydrolase activity. The single alanine mutation in the active site had no effect on expression, purification, or hexamer formation for the H126A, E129A, or Y272A variants compared with wild type. Thus, we assayed enzymatic activity by monitoring deoxyguanosine product formation using reverse-phase chromatography. All three variants had abolished hydrolytic activity compared with wild-type enzyme, even after a 2-h incubation in the presence of 100 µM dGTP and 5 mM MgCl₂ (SI Appendix, Fig. S1G). While the roles of His₁²⁶ and Glu₁²⁹ in catalysis have been well-established (10, 37), the function of Tyr₂⁷² is less clear. Similar to structures of the tetrameric dNTPases [PDB ID code 3IRH, PDB ID code 4TNQ, PDB ID code 2DQB (9, 11, 37)], His⁻²⁶ in line with the α-phosphate during catalysis. Interestingly, while mutations in the catalytic dyad showed slight product formation at the 2-h time point, the Y272A variant resulted in a consistently inactive enzyme (SI Appendix, Fig. S3G). This suggests that Tyr⁻²⁷² plays a critical role in Ec-dGTPase phosphohydrolase activity, possibly by stabilizing the α-phosphate for nucleophilic attack.

Characterization of Ec-dGTPase Specificity and Nucleotide Discrimination. Since Ec-dGTPase shows a very high affinity and preference for dGTP substrates (2, 18), we determined the mode by which the enzyme can specifically recognize this substrate compared with other purine or pyrimidine rings, namely dTTP, dATP, and dCTP. Recognition of dGTP within the active site is achieved by the enzyme can specifically recognize this substrate compared with other purine or pyrimidine rings, namely dTTP, dATP, and dCTP. Recognition of dGTP within the active site is achieved by Glu⁻¹²⁹ and the backbone carbonyl of Val⁻¹³², which interacts with the purine amine groups. In addition, residues Arg⁻¹³³ and Arg⁻¹⁴² project into the binding pocket from an adjacent monomer to interact with the ketone group of the purine ring (Fig. 3C and SI Appendix, Fig. S3E). These residues create a tight fit around the substrate by forming an extensive hydrogen bond network of interactions that stabilize dGTP for subsequent hydrolysis (Fig. 3B and C). In contrast, modeling of dTTP, dATP, and dCTP binding to the active site shows that these nucleotides bind loosely in the pocket (due to their smaller size) and establish fewer hydrogen bonds with the purine or pyrimidine rings (SI Appendix, Fig. S3H).

GTP Binding and Inhibition of Ec-dGTPase Activity. Given that Ec-dGTPase showed high specificity toward the guanosine ring, and cellular nucleotide concentrations of GTP can achieve levels 100- to 1,000-fold higher than dGTP (36, 37), we suspected that GTP may play a role in Ec-dGTPase cellular regulation. To test our hypothesis, we analyzed the effects of GTP on Ec-dGTPase activity in the presence of 100 µM dGTP. We observed a negative correlation between increasing GTP concentration and decreased Ec-dGTPase activity (Fig. 4A), suggesting that at physiologically relevant nucleotide ratios (>10:1 GTP:dGTP concentrations) GTP can act as a competitive inhibitor of Ec-dGTPase, resulting in the loss of dGTP hydrolysis (Fig. 4A). To further understand this result, we obtained the structure of the GTP-bound enzyme to 3.25 Å by similar cross-linking and soaking methods under physiological conditions.

Analogous to the dGTP and dGTP-1-thiol structures, GTP was observed in all six active sites. Superposition of the dGTP- and GTP-bound structures revealed conformational changes to the overall architecture of the active-site pocket (rmsd of 1.2 Å for 155 Cα atoms that comprise the active site), as well as differences relative to substrate binding (Fig. 4B and C and SI Appendix, Fig. S4A). In contrast to the dGTP-bound structure, the ribose 2′-OH of GTP occupies the position of the 3′-OH (in the dGTP-bound form), and the ribose 3′-OH draws closer to the Mn²⁺ ion, becoming part of its coordination sphere (Fig. 4B). As a result, GTP does not bind deeply in the pocket, losing interactions with the carbonyl of Val⁻¹³², which allows Tyr⁻²⁷² to stack against the purine ring and reach within hydrogen bond distance of the α-phosphate.
α-phosphate, pulling it away from the coordination sphere of the Mn²⁺ ion (Fig. 4 B and C). Intriguingly, overlay of the active sites for the GTP-bound and apo- Ec-dGTPase structures shows minimal conformational changes between the two states (rmsd of 0.39 Å on 155 Ca atoms comprising the active site; Fig. 4D and SI Appendix, Fig. S4 B and C). Thus, in contrast to the extensive conformational changes observed during dGTP binding (SI Appendix, Fig. S4B), GTP does not induce a transitional “catalytic state” of the active site.

Discussion
Despite recent Ec-dGTPase structures detailing the potential of ssDNA acting as an effector molecule, fundamental questions about dGTP recognition and dGTPase activity regulation have not been elucidated. The work presented herein establishes (i) unique approaches in both XFEL and cross-linking methodologies useful to the general crystallographic community for structure determination and (ii) a comprehensive structural framework for understanding dGTPase substrate recognition and activity.

We detailed protocols for fixed-target data collection at XFEL sources in an automated fashion (Fig. 1 and SI Appendix, Figs. S1 and S2). Notwithstanding the proven success of the injector setups for structure determination, goniometer-based fixed-target approaches at the XFEL are advantageous for (i) data collection using delicate crystals, (ii) crystals in limited supply, (iii) large, radiation-sensitive crystals, or (iv) crystal quality screening to prepare for injector-based experiments. Our highly efficient fixed-target data collection methodology was demonstrated to provide greater than 80% crystal hit rates for SFX experiments using randomly oriented micrometer-sized crystals with varying morphologies. While crystals in MCHs were mapped before the experiment using UV-fluorescence microscopy at the home laboratory, improvements to the standard goniometer setup in the LCLS MFX instrument (38) will incorporate UV-imaging capabilities for “on-the-fly” crystal identification and mapping. This will improve the precision of crystal positioning and provide a straightforward means to fully automate fixed-target SFX experiments using a variety of MCH form factors. As UV imaging is broadly applicable to identify a wide range of macromolecular crystals, the general crystallographic community may easily adopt this fully automated approach for multicrystal experiments at both synchrotron and XFEL sources. This method is also compatible with new methods for in situ crystal growth and data collection (Martiel, Müller-Werkmeister and Cohen, see ref. 39).

The success of our approach provided the opportunity to solve a radiation, damage-free structure of the apo-Ec-dGTPase enzyme from a limited number of crystals (Fig. 2). The XFEL structure revealed a distinctly closer contact of Mn²⁺ coordination toward the HD motif compared with an X-ray apo-structure and two previously published structures (20). Comparisons between the apo- and dGTP-bound structures show that conformational changes in the active site of Ec-dGTPase allow substrate binding. Such conformational changes involve (i) rigid body displacements that contract and expand the substrate cage to accommodate dGTP and (ii) individual residue motions to establish hydrogen bonds with the substrates. Moreover, remodeling of the binding pocket upon dGTP, but not GTP, binding reveals a “moldable” active site where conformational changes are coupled to selectivity.

The structure of Ec-dGTPase sheds light on the mechanism of nucleotide selectivity in dNTPases. The dNTPase activity of SAMHD1 is regulated by dGTP or GTP/dNTP binding at its primary/secondary allosteric sites, respectively, and is mediated

![Fig. 5. Structural and enzymatic insight into the mechanism of Ec-dGTPase activity. (A) Ribbon representation of the overlay between SAMHD1 (PDB ID code 4BZC) (purple) and Ec-dGTPase (green) illustrating fold conservation of the enzymatic cores. (B) Stereo and ball-and-stick representation of the overlay between SAMHD1 and Ec-dGTPase active sites illustrating residue type and geometry conservation. (C and D) Surface and ball-and-stick representation of SAMHD1 active-site residues illustrating that most contacts with the dNTP involve interaction with the ribose, the phosphates, and the purine or pyrimidine ring (circle). No interactions with SAMHD1 residues that could confer specificity are possible since a 7-Å gap separates them from the dNTP. Thus, dNTPases bind shared motifs D and F. A similar set of interactions takes place in dGTPase; however, dGTP selectivity occurs through formation of four hydrogen bonds.](image-url)
by tetramerization (40). The active form of SAMHD1 can bind and hydrolyze all four dNTPs with similar affinities and $K_{cat}$ (41). Structural studies have revealed the mechanism of nucleotide binding and have shown that the shape of the catalytic pocket remains nearly identical upon binding of the four dNTP substrates (12). Overlay of the Ec-dGTPase and SAMHD1 structures shows that, in addition to sharing similar active-site architectures (Fig. 5A), HD motif residues and those involved in stacking, 3′OH groups at the active site) triggered conformational changes affecting the catalytically essential Tyr275 (SI Appendix, Fig. S2C). The structure of Ec-dGTPase bound to dGTP sheds light on this matter. Comparisons between the active-site residues of the apo-, dGTP-, and ssDNA-bound forms illustrate that the latter combines features of the apo- and dGTP-bound structures (SI Appendix, Fig. S5). The net effect of ssDNA binding is an increase in the volume of the binding pocket by opening the tyrosine “gate” and hence improving binding efficiency, reconciling the observed threefold decrease in $K_m$. Altogether, we utilized a combination of XFEL and chemical cross-linking methods to successfully reveal the molecular basis of Ec-dGTPase substrate recognition and the regulation mechanism. Together with the limited known structures of dNTPases in both prokaryotic and eukaryotic organisms, our results provide structural insight into the metabolic regulation of tightly controlled dNTP pools in DNA replication and cellular survival.

Methods

Cloning, Protein Purification, and Crystallization. The cDNA encoding wild-type Ec-dGTPase was a gift from Charles C. Richardson (Harvard Medical School, Boston, MA). It was cloned into the pET21 vector with a His-tag at the N terminus. Site-specific mutants of Ec-dGTPase were prepared using QuikChange mutagenesis kits (Agilent). Wild-type and mutant Ec-dGTPase proteins were expressed and purified as previously described (40). Before crystallization, Ec-dGTPase was buffer exchanged into ssDNA, reflecting substrate-binding enhancement (20). Moreover, the crystal structure revealed that ssDNA binding at the interface of the monomer (at the active site) triggered conformational changes affecting the catalytically essential Tyr275 (SI Appendix, Fig. S2C). The structure of Ec-dGTPase bound to dGTP sheds light on this matter. Comparisons between the active-site residues of the apo-, dGTP-, and ssDNA-bound forms illustrate that the latter combines features of the apo- and dGTP-bound structures (SI Appendix, Fig. S5). The net effect of ssDNA binding is an increase in the volume of the binding pocket by opening the tyrosine “gate” and hence improving binding efficiency, reconciling the observed threefold decrease in $K_m$.

To mount each MCH onto the beamline goniometer followed by a manual, semiautomated alignment procedure where the MCH is rotated face on to the on-axis microscope and the four reference markers are clicked in a clockwise order from within a video display of the software interface. Following this procedure, the location of the crystal coordinates is displayed over the video image of the mount (SI Appendix, Fig. S1E) and are visually inspected. If necessary, a graphical interface enables the experimenter to remove or shift egregious crystal positions or shift the location of the reference points to improve accuracy. Updated crystal positions are stored, and the user is prompted to begin automated data collection. During automated data collection, a crystal is translated into the beam position between each X-ray pulse. This process was repeated for each MCH in the cassette.

Diffraction experiments on Ec-dGTPase and Pol II complexes were done using 9.5-keV X-ray pulses with a 40-s duration and an 8-μm beam focus at the X-ray interaction point. Diffraction images were recorded on a Rayonix MX325 detector and processed using the cctbx.xfel software package (45, 46). Synchrotron-based X-ray diffraction experiments of single dGTPase crystals were performed on SSRRL beamline 12.2 and the APS beamlines 22ID and 23IDD. Data were processed using XDS and SCALA software packages (47, 48). Single anomalous diffraction experiments of selenium–methionine-labeled dGTPase crystals were collected at 12.656 keV with inverse beam every 15° of oscillation data.

Structure Determination and Refinement. Selenium substrates were determined with SHELXSCD (49), using a resolution cutoff of 4.4Å corresponding to a CC anom of 0.301. Substructure solutions were utilized in the CRANK pipeline (50), resulting in an initial, experimentally phased structure of Ec-dGTPase (SI Appendix, Fig. S5A), which was then manually built in Coot and refined in BUSTER. Subsequent Ec-dGTPase apo- (XFEL), GTP-bound, dGTP-1-thiol, and dGTP-bound structures were solved by PHASER (51) using the SeMet structure as a search model. All structures were refined using Phenix (52) and BUSTER (53), followed by several cycles of manual refinement in Coot (54, 55). All superpositions and figures were rendered in PyMOL. Potential hydrogen bonds were assigned using a distance of <3.5 Å and an A-D-H angle of >90°, while the maximum distance allowed for a van der Waals interaction was 4.0 Å.

Enzymatic Assay of Ec-dGTPase Activity. Purified wild-type and active-site variants were dialyzed overnight into reaction buffer (20 mM Tris, pH 7.8, 50 mM NaCl, 3% glycerol, 5 mM MgCl2) and concentrated to 2 mg/mL. For phospohydrolase experiments, 2 μM enzyme was incubated with 100 μM dGTP (TriLink Biotech) at room temperature. Enzymatic activity was monitored at 5-, 10-, 30-, 60-, and 120-min time points by quenching the reaction as wicking away excess solvent may also result in the loss of crystalline sample.
with 50 mM EDTA. Analysis of the deoxyguanosine product at the various time points was achieved by reverse-phase HPLC. Briefly, quenched reactions were injected into a C18 M column (Phenomenex) against 10 mM ammonium phosphate (pH 7.8) and 5% methanol, and the deoxyguanosine product was eluted with a gradient to 30% methanol. Individual peak heights were integrated and compared between Ec-dGTPase constructs. To test the effect of GTP on enzymatic activity, enzyme was assayed in a similar manner in the presence of 100 μM GTP and increasing concentrations of GTP (0–2 mM).

To test the enzymatic activity of cross-linked crystals, we washed glutaraldehyde-fixed crystals extensively with low-salt reservoir buffer (20 mM Tris-HCl pH 8.0, 0.2 M sodium chloride, 2 mM manganese chloride, 0.02% azide, 3% glycerol) as described above. Enzymatic activity was monitored by incubating washed crystals in the presence of 100 μM dGTP for 60 min at 37 °C. Reverse-phase HPLC was used to analyze the deoxyguanosine product for cross-linked enzyme in crystals, unmodified enzyme in solution, and unmodified enzyme in solution in the presence of 500 μM GTP.


Acknowledgments. We thank Michael Becker and Craig Ogata (GM/CA at Argonne National Laboratory); Irmin Mathews, Clyde Smith, and Ana Gonzales at the Stanford Synchrotron Radiation LightSource (SSRL) for their support during data collection; D. Lee for home-source X-ray technical support; and Charles C. Richardson (Harvard University) for providing cDNA of dGTPase. This work was supported by NIH Grants R01GM116642 (to J.A.), R01GM121866 and P50GM082251 (to G.C.), and R01GM171726 (to N.K.S.). G.C. also acknowledges support from BioXFEL-STC1231306. Use of the UCLS and SSRl SLAC National Accelerator Laboratory is supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract DE-AC02-76SF00515. The SSRl Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research and by the NIH, National Institute of General Medical Sciences (NIGMS) (including Grant P41GM103393). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the NIGMS or NIH.