PNAS

Supplementary Information for
Mechanism by which T7 bacteriophage protein Gp1.2 inhibits *Escherichia coli* dGTPase

Bradley P. Klemm,[a,1] Deepa Singh,[a,1] Cassandra E. Smith,[a] Allen L. Hsu,[a] Lucas B. Dillard,[a] Juno M. Krahn,[a] Robert E. London,[a] Geoffrey A. Mueller,[a] Mario J. Borgnia,[a] and Roel M. Schaaper[a,2]

[a]Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709, USA;

[1]B.P.K. and D.S. contributed equally to this work.

Roel M. Schaaper
Email: schaaper@niehs.nih.gov

This PDF file includes:

Figures S1 to S10
Tables S1 to S3
SI References

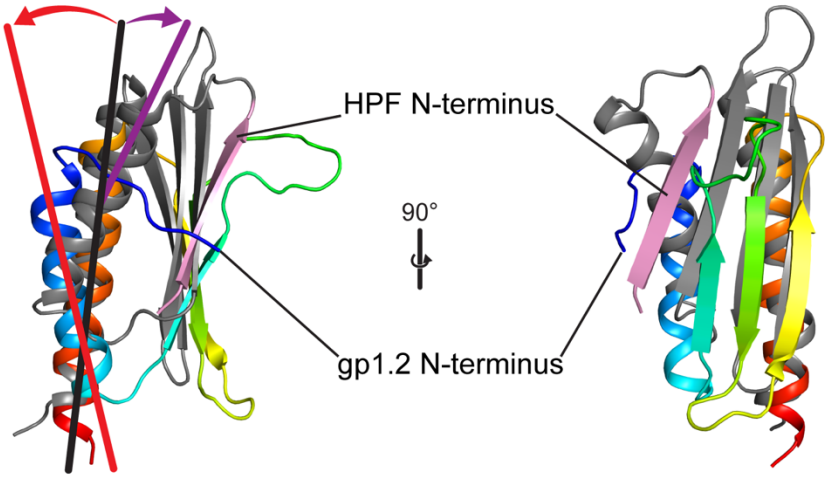

Fig. S1. Alignment of Gp1.2 (PDB ID: 2MDP) with HPF (PDB ID: 2RQL) (1). Gp1.2 is colored as a rainbow as in Fig. 1, with the N-terminus blue and C-terminus red. HPF is grey, with the N-terminus colored pink. Relative to the α2 helical axis (black), HPF's α1 helical axis (purple) is rotated towards the β-sheet by ~20 degrees (purple arrow). In contrast, Gp1.2's α1 helical axis (red) is instead rotated away from the β-sheet by ~20 degrees (red arrow). Generated in PyMOL.

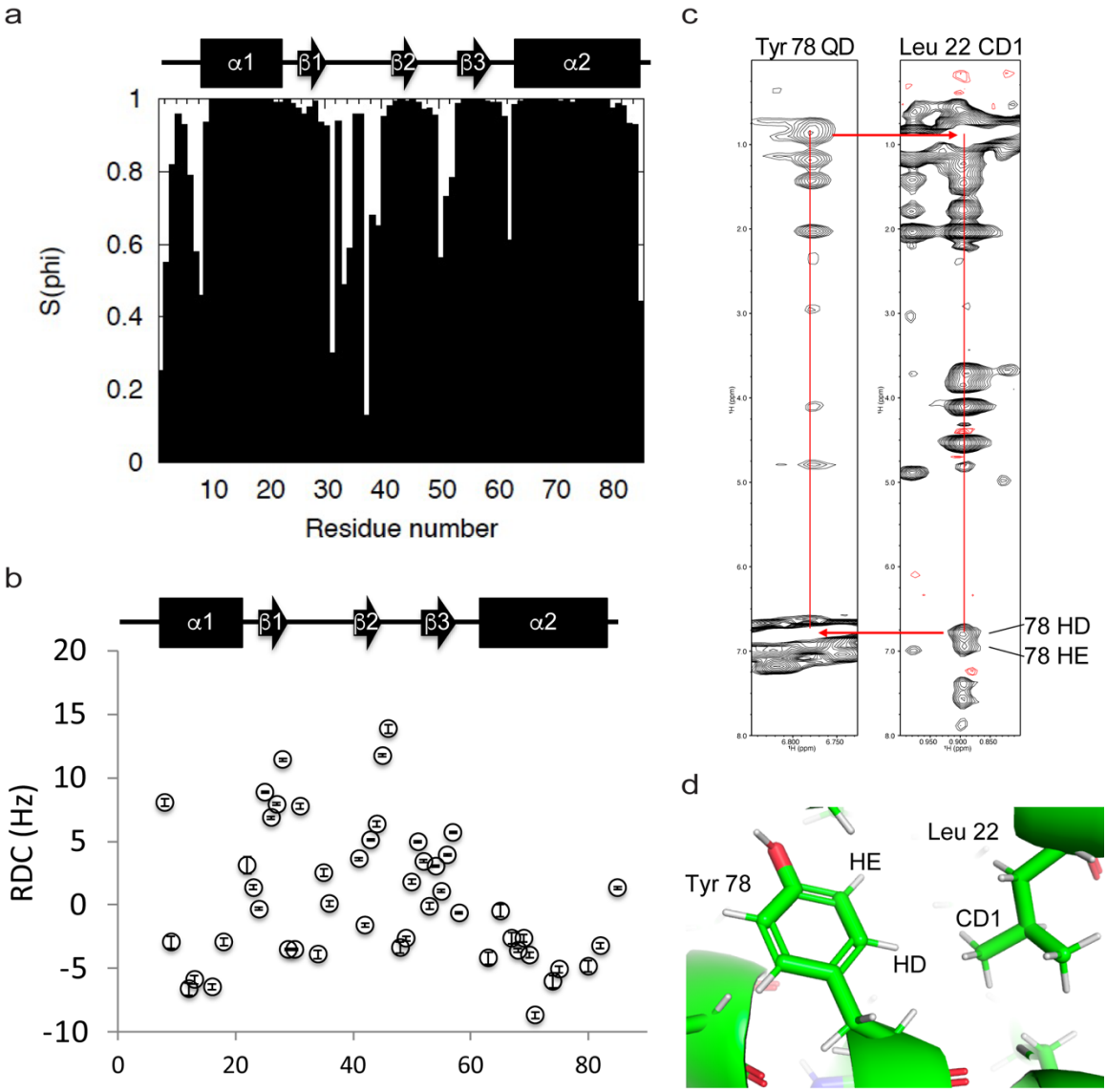

Fig. S2. NMR structure factors for Gp1.2. (a) Angle phi of the NMR ensemble is a pseudo-order parameter (2). The regions with more apparent disorder include the N- and C-termini and the β1–β2 and β2–β3 loops (compare to Fig. S6a). (b) The measured residual dipolar couplings (RDC) by residue, which provide additional restraints for the structure calculations. The secondary structure in the NMR model is shown between the panels, for reference. The measured RDCs are in good agreement with the structure. Importantly, this provides strong evidence of precise positioning of the angle between helices 1 and 2, which was used to position Gp1.2 into the Cryo-EM density. (c) Example NOE data. Left: Strip plot from ^{1}H-^{13}C-^{1}H NOESY at the plane corresponding the folded frequency of Tyr78 CD1 and CD2. An NOE to Leu22 HD1 is highlighted. Right: Strip plot at folded frequency of Leu 22 CD1. The return NOE to Tyr 78 HD is highlighted as well as an NOE to TYR 78 HE. (d) Tyr 78 and Leu 22 are adjacent in the structure of Gp1.2.

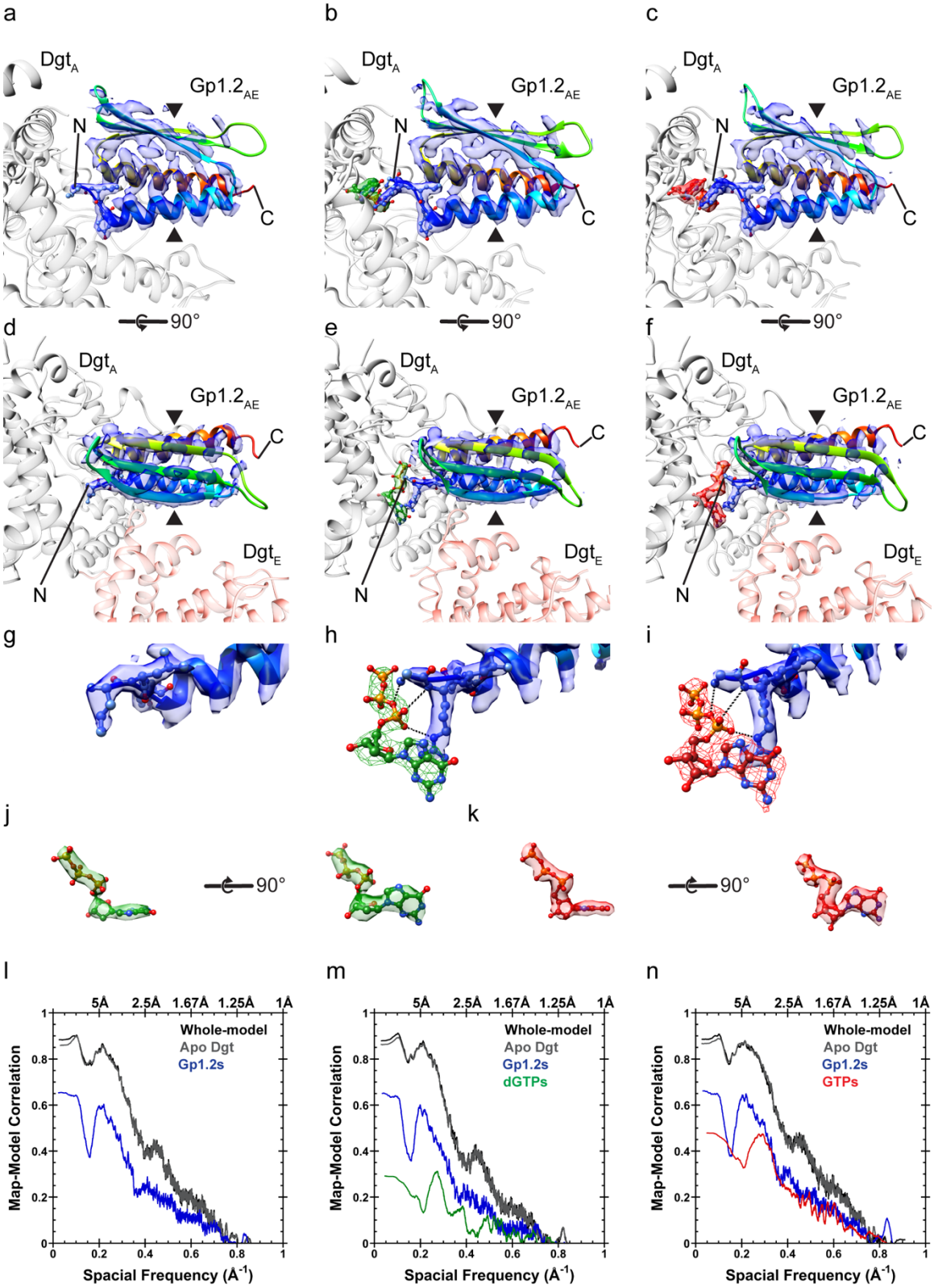

Fig. S3. Fitting Gp1.2 into each of the Cryo-EM maps. (**a–i**) Gp1.2$_{AE}$ from the three cryo-EM structures. The map is colored blue, while the model is colored as a rainbow from N-terminus (blue) to C-terminus (red). The black arrows approximately indicate the boundary between the

Gp1.2 residues that bind to Dgt (proximal, left of arrows) and those that make no contacts with Dgt (distal, right of arrows). The density is discontinuous in regions that are more distal to Dgt and in the β1–β2 and β2–β3 loops, which showed a high degree of disorder in the NMR structure (Fig. S2a). (**a,d,g**) Gp1.2$_{AE}$ from the Dgt-Gp1.2 structure. (**b,e,h**) Gp1.2$_{AE}$ from the Dgt–dGTP–Gp1.2 structure. (**c,f,i**) Gp1.2$_{AE}$ from the Dgt–GTP–Gp1.2 structure. (**j**) dGTP fitting into the Dgt–dGTP–Gp1.2 cryo-EM map. (**k**) GTP fitting into the Dgt–GTP–Gp1.2 cryo-EM map. (**l–n**) Map-Model Correlation plots broken down by component, as indicated, for each map.

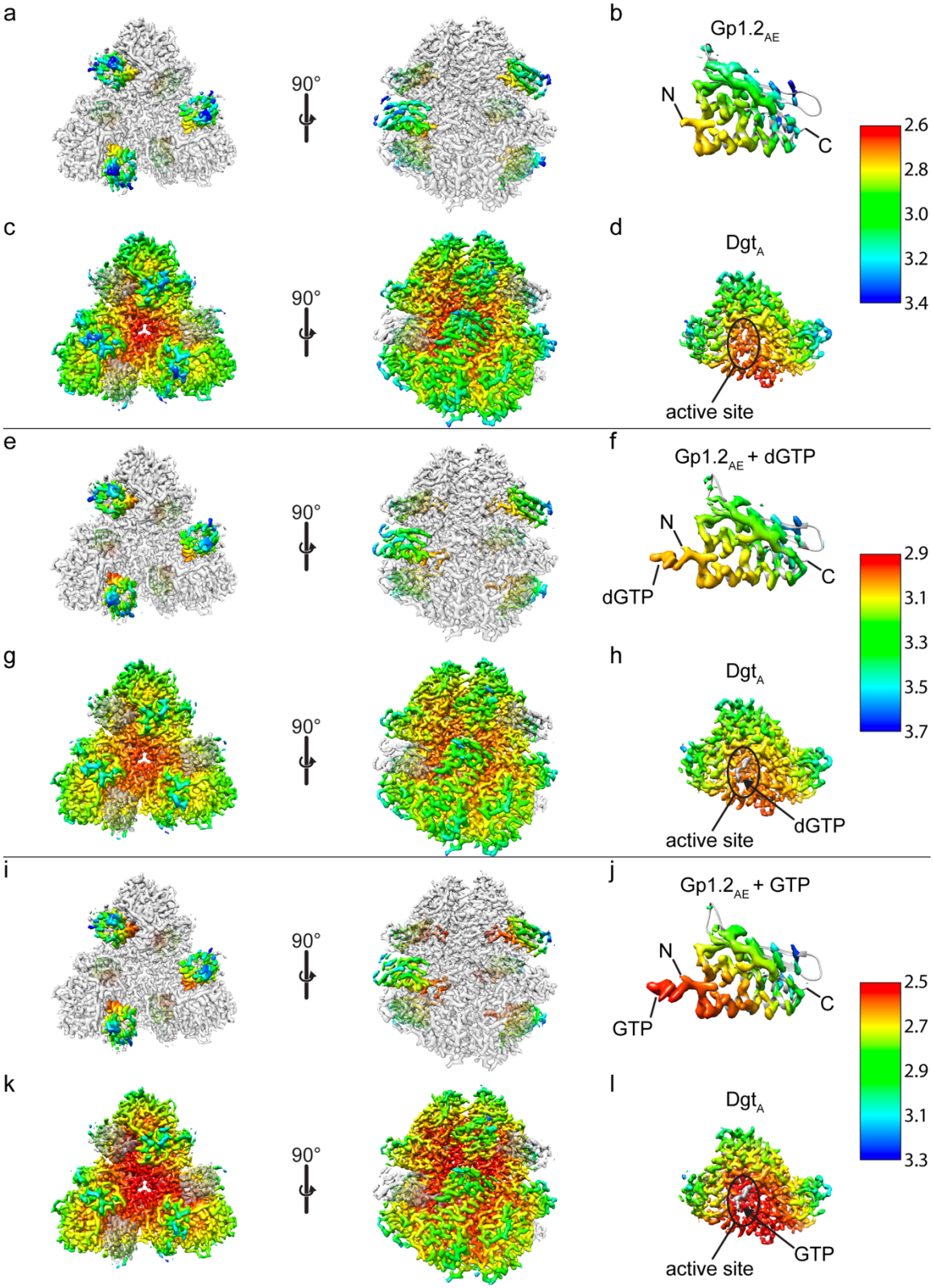

Fig. S4. Local resolution distribution across the cryo-EM maps. (**a–d**) Dgt–Gp1.2 cryo-EM structure. (**e–h**) Dgt– dGTP–Gp1.2 cryo-EM structure. (**i–l**) Dgt–GTP–Gp1.2 cryo-EM structure.

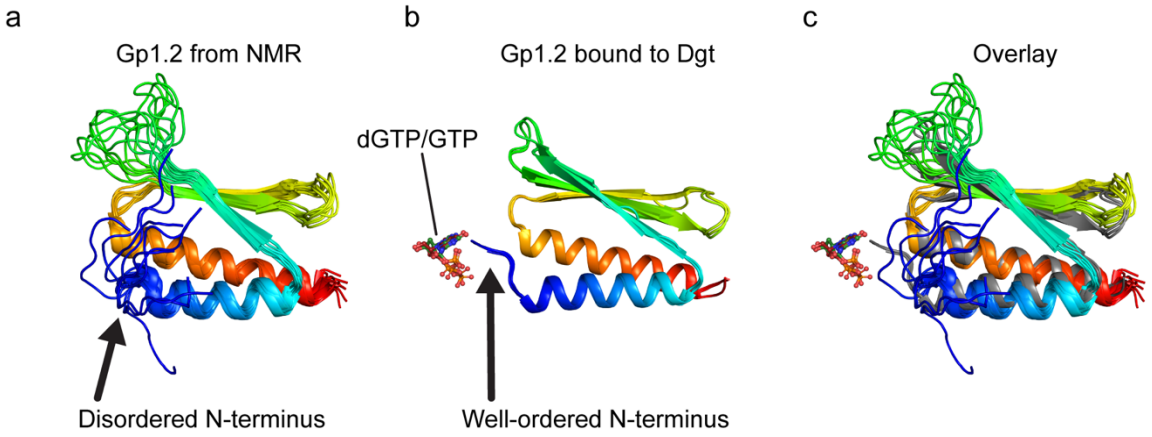

Fig. S5. Comparison of the Gp1.2 structures from NMR and cryo-EM. (**a**) All 10 states of the NMR structure. The regions that show the highest degree of disorder are the N-terminus and the β1–β2 loop, with the C-terminus and the β2–β3 also modestly disordered (compare to Fig. S2a). (**b**) $Gp1.2_{AE}$ from all three cryo-EM structures. (**c**) An overlay of the Gp1.2 models from the NMR structure (rainbow) and the cryo-EM structures (grey). The well-ordered regions of the NMR structure are mostly indistinguishable from Gp1.2 in the cryo-EM maps, while the N-terminus becomes ordered upon binding to the Dgt active site.

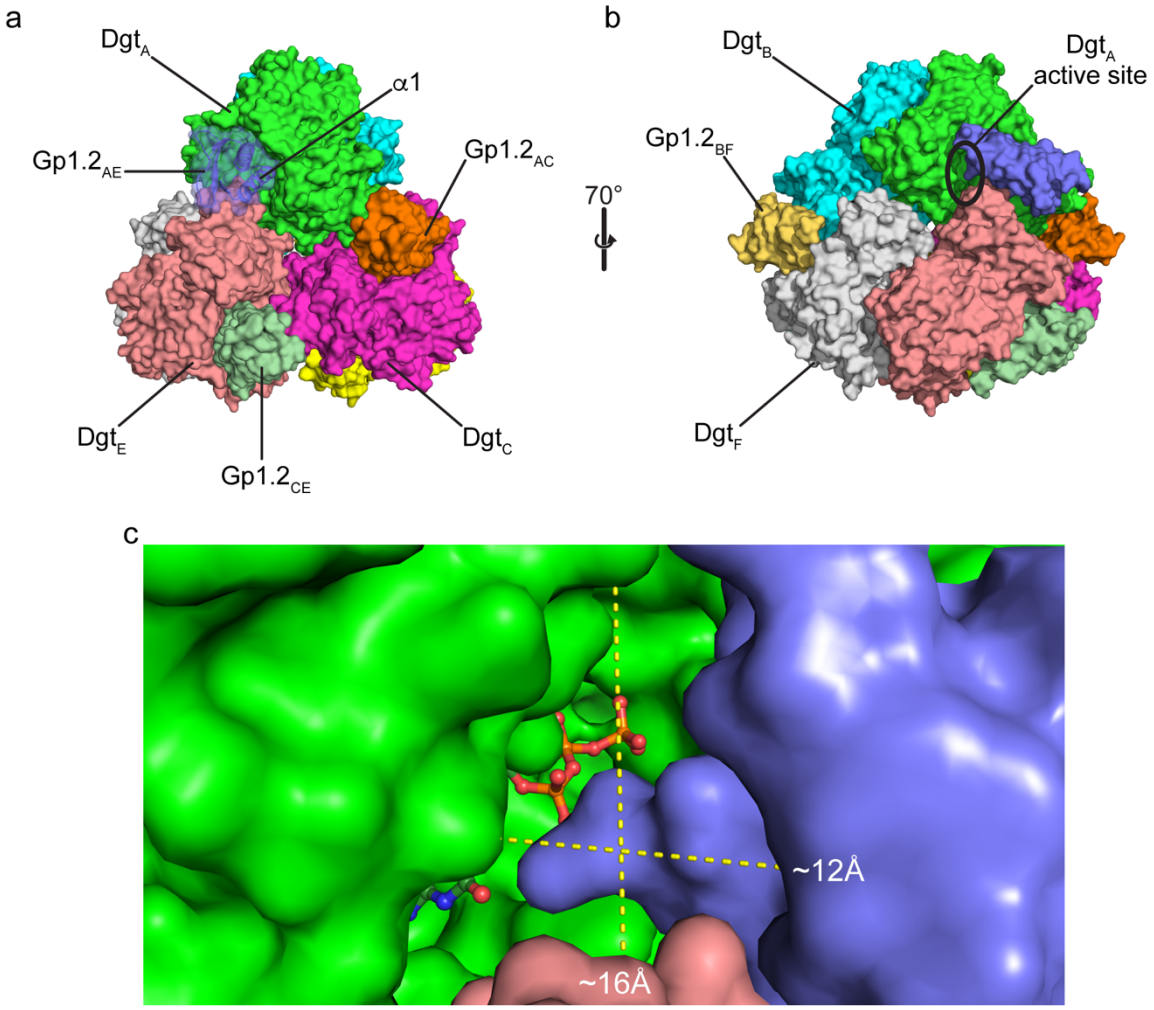

Fig. S6. Overview of the Dgt–Gp1.2 structure. Generated in PyMOL. (**a**) the Dgt–Gp1.2 structure colored by chain. The Gp1.2 bound between monomers A and E is shown with a transparent surface to highlight the binding site of α1. (**b**) the Dgt active site remains accessible when Gp1.2 is bound. (**c**) close-up of the opening to the Gp1.2-inhibited Dgt_A active site bound to dGTP, with the dimensions of the solvent-accessible opening to pocket labeled.

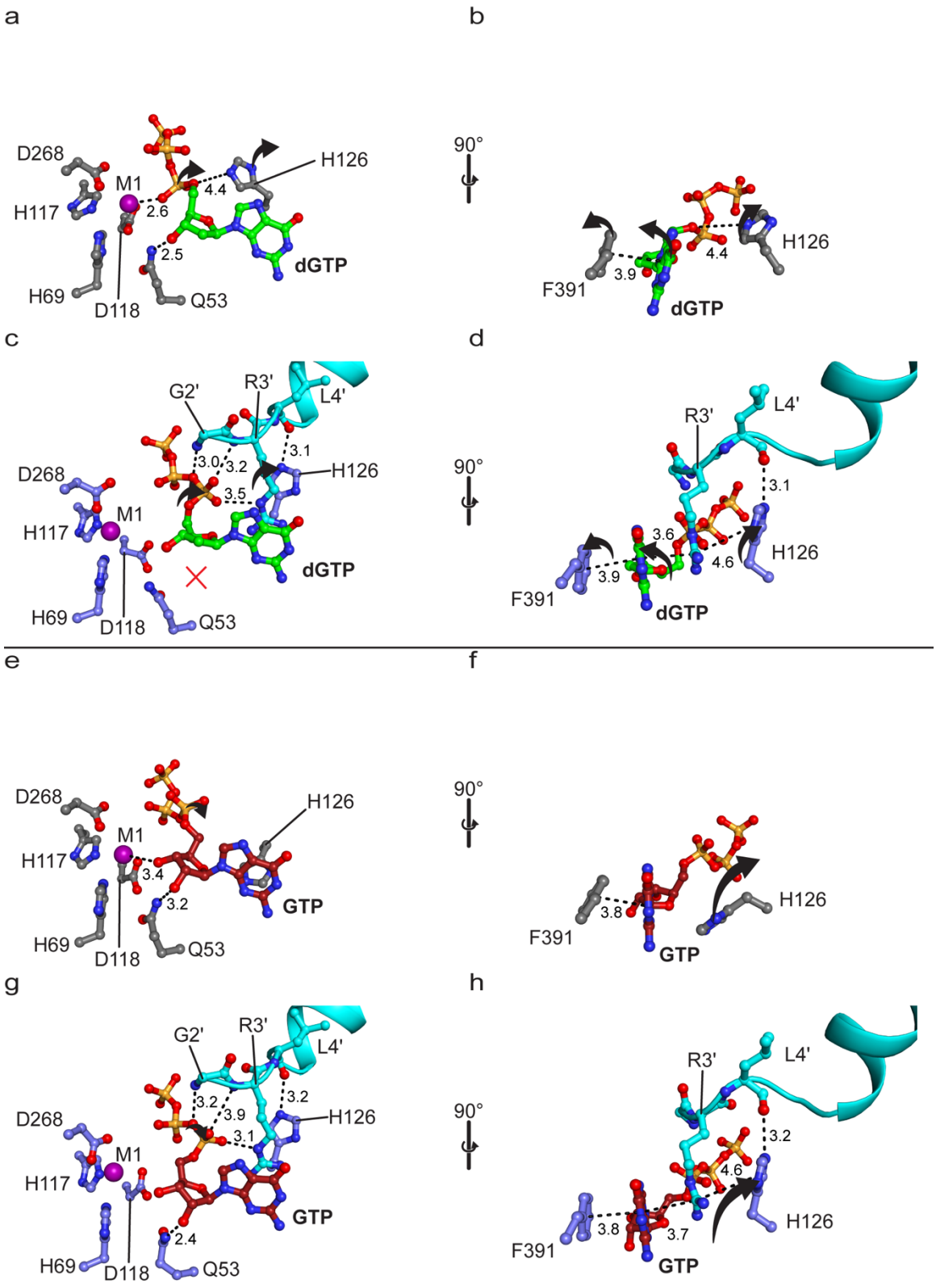

Fig. S7. Gp1.2 displaces substrate and critical Dgt residues in the active site. Generated in PyMOL. (**a–d**) displacement of the dGTP α-phosphate and catalytic acid H126 by the Gp1.2 N-terminus. The dGTP-bound Dgt structure (**a–b**) (PDB ID: 6OIY) (3) was aligned to the Dgt–

Gp1.2–dGTP structure (**c–d**) by the four HD-motif residues (r.m.s.d.=0.213 over the four C-alpha atoms). The black arrows highlight the significant Dgt/dGTP displacements that occur upon Gp1.2 binding, and they maintain their absolute positions between equivalent panels. In the upper panels (a–b, e–f) the arrows indicate the displacements that will take place upon Gp1.2 binding, while in the lower panels the arrows represent the displacements once they have taken place. The dGTP α-phosphate is displaced by >2.5 Å and the H126 imidazole ring is displaced by >3.5 Å, for instance. Black dashed lines are substrate interactions without Gp1.2, or the interactions that replace them in the ternary complex. (**b, d**) 90° rotated views from panels a, c. the dGTP guanine ring is displaced by ~2 Å, F391 by 1.6 Å, and H126 by 3.5 Å. (**e**–**h**) displacement of GTP by the Gp1.2 N-terminus is much more modest than for dGTP. The GTP-bound Dgt structure (**e–f**) (PDB ID: 6OIX) (3) was aligned to the Dgt–Gp1.2–GTP structure (**g–h**) by the four HD-motif residues (r.m.s.d.=0.286 over the four C-alpha atoms). Like in the above panels for dGTP, the black arrows highlight GTP displacement by Gp1.2 and maintain their absolute positions between equivalent panels: the α-phosphate is only displaced by 1.7 Å. Black dashed lines are GTP interactions without Gp1.2, or the interactions that replace them in the ternary complex. (**f, h**) 90° rotated views from panels e, g. The GTP guanine ring and F391 sidechain are only modestly displaced by Gp1.2 (<1 Å), while the H126 imidazole ring is displaced by 5.4 Å.

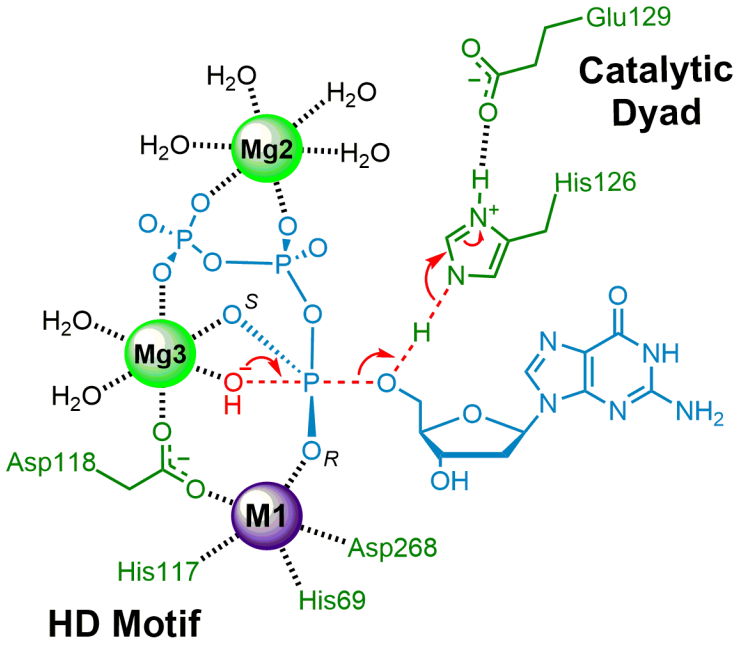

Fig. S8. Proposed Dgt catalytic mechanism (adapted to the *E. coli* active site from (4, 5)). Gp1.2 inhibits the reaction by i) displacing the phosphates from the active site metals, which prevents the alignment of the nucleophilic water/hydroxide with the scissile bond, and ii) preventing H126 from accessing the 5'-oxygen to protonate the leaving group.

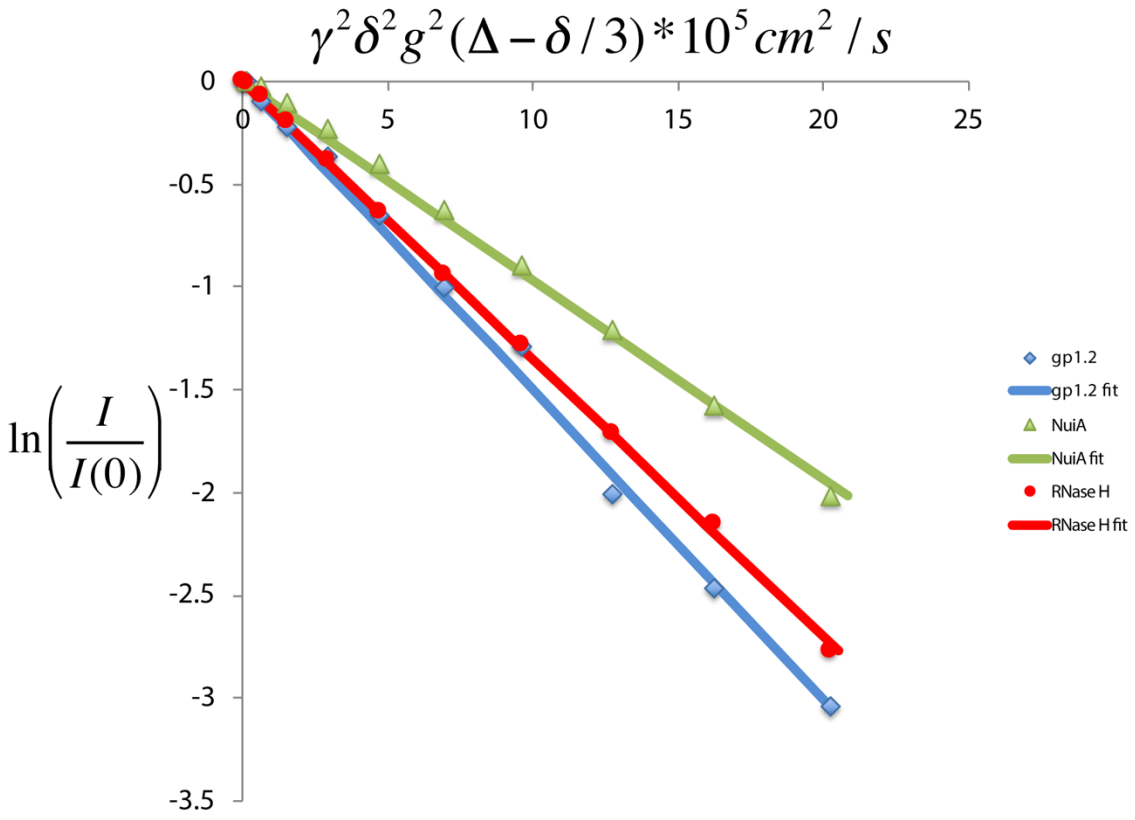

Fig. S9. Diffusion decays of three proteins. The signal decay as a function gradient strength (*g*) is plotted according to Nesmelova *et al*. (6). The proteins are gp1.2 (blue, Rg =14.5 Å), NuiA (green, Rg=16.3 Å), and RNase H (red, Rg=14.8 Å). The X-axis is the natural log of the signal intensity divided by the signal intensity with no gradient. The Y-axis is formulated so the slope is proportional to the rate of diffusion: γ is gyromagnetic ratio of ^{1}H, δ = 4 ms, and Δ=62.55 ms.

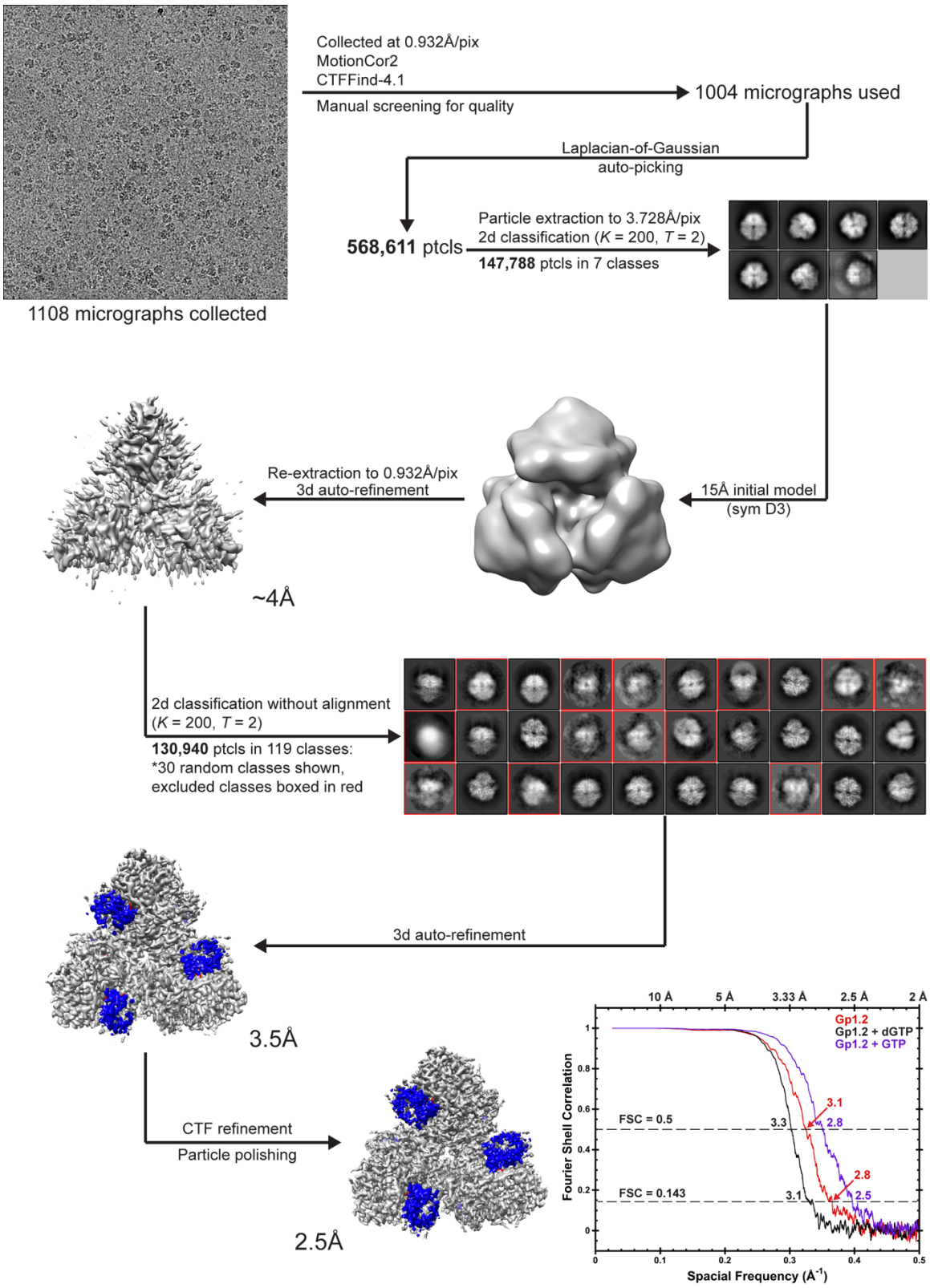

Fig. S10. Data collection/processing scheme using the Dgt–gp1.2–GTP grid as an example. All data processing steps were handled in RELION (7). The density in the maps following refinement and CTF refinement/particle polishing that corresponds with gp1.2 is colored blue, while the GTP

(mostly not visible from this point of view) is colored red. The FSC curves for the masked half-maps were generated in PHENIX using the comprehensive cryo-EM validation tool (8) and are plotted in KaleidaGraph.

Table S1. NMR statistics. Gp1.2 structure PDB ID: 2MDP, BMRB ID: 19498.

	NMR Data	Supplemental Reference
NOEs from CYANA	798	(9)
NOEs assigned manually	50	
RDC (secondary structure only)	33	
H-bonds	28	(10)
Dihedrals (TALOS+)	122	(11)

	Structure Data	Supplemental Reference
Ramachandran Plot		
Favored	96.6%	(12)
Allowed	3.4%	
Disallowed	0	
MolProbity Clash Score	9.14	(12)
(Z-score)	-0.04	
RMSD (Å)		(1)
Backbone all	2.2	
Backbone ordered	0.6	
Heavy atoms all	2.8	
Heavy atoms ordered	1.1	

Table S2. Cryo-EM data collection/processing information.

Complex	Dgt + Gp1.2	Dgt + Gp1.2 + dGTP	Dgt + Gp1.2 + GTP
EMDB ID	26360	26361	26362
PDB ID	7U65	7U66	7U67
Data Collection			
Microscope	Talos Arctica		
Voltage (keV)	200		
Camera	K2 Summit		
Magnification	45,000		
Pixel size (Å)	0.932		
Defocus range (µm, from CtfFind)	-0.50 to -2.5		
Data Processing			
Micrographs used (collected)	1404 (1638)	1037 (1206)	1004 (1108)
Particles in final class	114,127	113,934	130,940
Refinement resolution limit (post-processed resolution, both in Å)	2.96 (2.85)	3.13 (2.96)	2.81 (2.54)
RELION Sharpening B-factor ($Å^2$)	-59	-87	-53
Symmetry Imposed	D3	D3	D3
Residues built per Dgt monomer (505 total)	481	474	476
Residues built per Gp1.2 (84 total)	83	83	83
Number of ligands (type)	6 (Mg^{2+})	6 (dGTP) 6 (Mg^{2+})	6 (GTP) 6 (Mg^{2+})
RMSD bonds (Å)	0.005	0.005	0.005
RMSD angles (°)	0.648	0.667	0.581
Ramachandran: favored (%)	96.39	96.31	97.61
allowed (%)	3.61	3.51	2.39
outliers (%)	0	0	0
Rotamer outliers (%)	0.46	0.92	0.91
Cβ outliers (%)	0	0	0
Cis proline/general (%)	0 (0)	0 (0)	0 (0)
Twisted proline/general (%)	0 (0)	0 (0)	0 (0)
All atom clashscore	2.74	2.83	3.12
MolProbity score	1.30	1.32	1.19

Table S3. Biochemical parameters reported in Figure 6.

[Gp1.2] (nM)	k_{cat} (s^{-1}) *	K_M (μM) *	n^H *	v_0 vs [S] †	IC_{50} (nM)	n^H
0	2.83 ± 0.05	21.3 ± 0.9	1.6 ± 0.1	k_{cat} effect: ↓	66 ± 8	1.3 ± 0.2
28	2.0 ± 0.2	27 ± 5	1.5 ± 0.4	K_M effect: ↑	70 ± 10	2.3 ± 0.9
50	1.79 ± 0.08	47 ± 5	1.4 ± 0.2	$k_{cat,app}$ ‡	IC_{50} (nM)	n^H
100	1.11 ± 0.09	60 ± 10	1.2 ± 0.2	–GTP	160 ± 20	1.0 ± 0.1
135	0.62 ± 0.07	80 ± 20	1.5 ± 0.4	+GTP	27 ± 3	1.2 ± 0.1
200	0.56 ± 0.05	80 ± 20	1.5 ± 0.3			
300	0.51 ± 0.03	90 ± 10	1.1 ± 0.2			
2000	~0	ND§	ND§			
[GTP] (mM)	**k_{cat} (s^{-1}) ¶**	**K_M (μM) ¶**	**n^H ¶**	**v_0 vs [S] #**	**K_I (μM)**	
0	2.6 ± 0.2	13 ± 2	1.2 ± 0.3	k_{cat} effect	n/a	
0.25	2.54 ± 0.04	70 ± 3	2.0 ± 0.1	K_M effect: ↑	120 ± 50	
0.5	2.90 ± 0.05	98 ± 3	2.3 ± 0.2	$k_{cat,app}$ ‖	IC_{50} (nM)	n^H
0.75	2.53 ± 0.06	123 ± 6	2.1 ± 0.2	+Gp1.2	310 ± 50	2.1 ± 0.5
1	2.66 ± 0.06	185 ± 9	2.0 ± 0.2			
2	2.62 ± 0.07	330 ± 10	2.8 ± 0.3			

* Derived from fitting Eq. 1 to the $v_0/[E]_t$ vs [S] plots in Fig. 6a. The k_{cat} and K_M data are replotted in Fig. 6b/c.
† Derived from fitting Eq. 2 or Eq. 3 to the data in Fig. 6b/c.
‡ Derived from fitting Eq. 4 to the data in Fig. 6d.
§ ND: Not Determined.
¶ Derived from fitting Eq. 1 to the $v_0/[E]_t$ vs [S] plots in Fig. 6e. The k_{cat} and K_M data are replotted in Fig. 6f/g.
\# Derived from fitting Eq. 5 to the data in Fig. 6g.
‖ Derived from fitting Eq. 4 to the data in Fig. 6h.

SI References

1. A. Sato *et al.*, Solution structure of the *E. coli* ribosome hibernation promoting factor HPF: Implications for the relationship between structure and function. *Biochem Biophys Res Commun* **389**, 580-585 (2009).
2. A. Bhattacharya, R. Tejero, G. T. Montelione, Evaluating protein structures determined by structural genomics consortia. *Proteins* **66**, 778-795 (2007).
3. C. O. Barnes *et al.*, The crystal structure of dGTPase reveals the molecular basis of dGTP selectivity. *Proc Natl Acad Sci U S A* **116**, 9333-9339 (2019).
4. E. R. Morris *et al.*, Crystal structures of SAMHD1 inhibitor complexes reveal the mechanism of water-mediated dNTP hydrolysis. *Nat Commun* **11**, 3165 (2020).
5. B. P. Klemm *et al.*, High-resolution structures of the SAMHD1 dGTPase homolog from Leeuwenhoekiella blandensis reveal a novel mechanism of allosteric activation by dATP. *J Biol Chem* **298**, 102073 (2022).
6. I. V. Nesmelova, D. Idiyatullin, K. H. Mayo, Measuring protein self-diffusion in protein-protein mixtures using a pulsed gradient spin-echo technique with WATERGATE and isotope filtering. *J Magn Reson* **166**, 129-133 (2004).
7. J. Zivanov *et al.*, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7** (2018).
8. P. V. Afonine *et al.*, New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr D Struct Biol* **74**, 814-840 (2018).
9. T. Herrmann, P. Güntert, K. Wüthrich, Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J Mol Biol* **319**, 209-227 (2002).
10. A. Grishaev, A. Bax, An empirical backbone-backbone hydrogen-bonding potential in proteins and its applications to NMR structure refinement and validation. *J Am Chem Soc* **126**, 7281-7292 (2004).
11. Y. Shen, R. Vernon, D. Baker, A. Bax, *De novo* protein structure generation from incomplete chemical shift assignments. *J Biomol NMR* **43**, 63-78 (2009).
12. I. W. Davis *et al.*, MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**, W375-383 (2007).