Full wwPDB X-ray Structure Validation Report (i)

Jan 10, 2022 – 12:26 pm GMT

PDB ID : 7BBN

Title : Crystal structure of an ancient sequence-reconstructed Elongation Factor Tu (node 317)

Authors : Majumdar, S.; Bergfors, T.; Sanyal, S.

Deposited on : 2020-12-18

Resolution : 1.68 Å(reported)

1   Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.68 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.
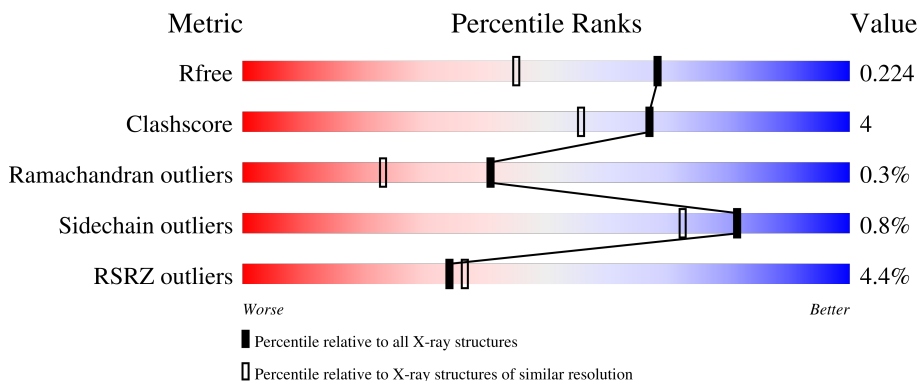
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	6780 (1.70-1.66)
Clashscore	141614	7310 (1.70-1.66)
Ramachandran outliers	138981	7173 (1.70-1.66)
Sidechain outliers	138945	7172 (1.70-1.66)
RSRZ outliers	127900	6661 (1.70-1.66)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

| Mol	Chain	Length	Quality of chain
1	A	393	3% / 83% / 8% · 8%
1	B	393	5% / 86% / 6% · 7%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6015 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Elongation Factor Tu.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	361	Total 2785	C 1765	N 470	O 537	S 13	0	5	0
1	B	366	Total 2809	C 1780	N 476	O 540	S 13	0	4	0

- Molecule 2 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: $C_{10}H_{15}N_5O_{11}P_2$) (labeled as "Ligand of Interest" by depositor).
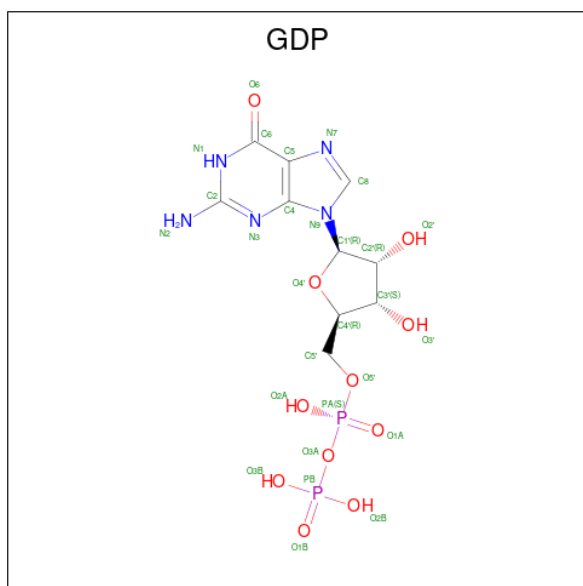
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 28	C 10	N 5	O 11	P 2	0	0
2	B	1	Total 28	C 10	N 5	O 11	P 2	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Mg 1	0	0
3	B	1	Total 1	Mg 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	191	Total 191	O 191	0	0
4	B	172	Total 172	O 172	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Elongation Factor Tu

Chain A:

- Molecule 1: Elongation Factor Tu

Chain B:
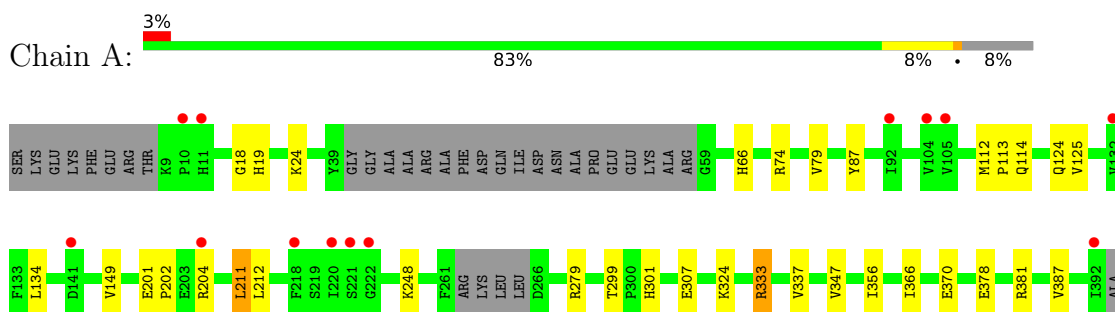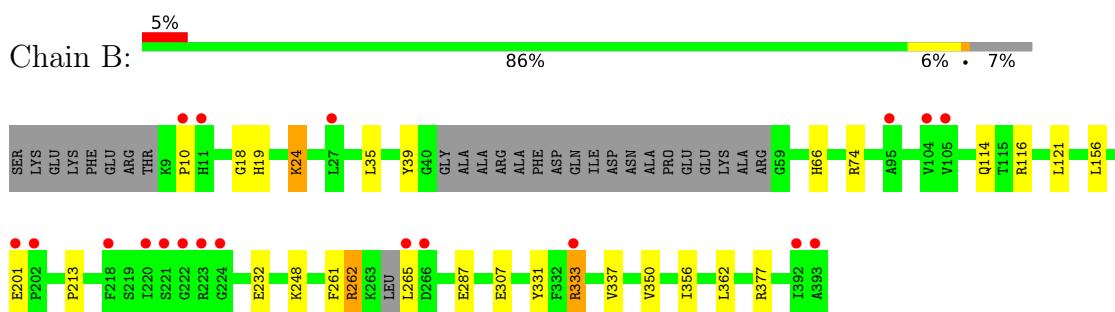
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	97.34Å 100.29Å 160.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.44 – 1.68 29.42 – 1.68	Depositor EDS
% Data completeness (in resolution range)	98.8 (29.44-1.68) 98.8 (29.42-1.68)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.03 (at 1.68Å)	Xtriage
Refinement program	REFMAC 5.8.0257	Depositor
R, R_{free}	0.191 , 0.220 0.201 , 0.224	Depositor DCC
R_{free} test set	4447 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	22.6	Xtriage
Anisotropy	0.141	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.039 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6015	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.83% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.
[2] Theoretical values of $< \|L\| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: GDP, MG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$>5	RMSZ	#$|Z|$>5
1	A	0.78	0/2852	0.95	3/3874 (0.1%)
1	B	0.77	0/2873	0.98	5/3900 (0.1%)
All	All	0.78	0/5725	0.97	8/7774 (0.1%)

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	333	ARG	NE-CZ-NH2	-13.98	113.31	120.30
1	B	333[A]	ARG	NE-CZ-NH2	-12.61	113.99	120.30
1	B	333[B]	ARG	NE-CZ-NH2	-12.61	113.99	120.30
1	A	333	ARG	NE-CZ-NH1	9.73	125.17	120.30
1	B	74	ARG	NE-CZ-NH2	-6.16	117.22	120.30
1	B	333[A]	ARG	CB-CG-CD	-6.00	96.00	111.60
1	B	333[B]	ARG	CB-CG-CD	-6.00	96.00	111.60
1	A	74	ARG	NE-CZ-NH2	-5.85	117.37	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2785	0	2768	21	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	2809	0	2784	24	0
2	A	28	0	12	0	0
2	B	28	0	12	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	191	0	0	3	0
4	B	172	0	0	3	0
All	All	6015	0	5576	44	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 4.

All (44) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:333[B]:ARG:NH1	4:B:501:HOH:O	1.96	0.99
1:B:331[A]:TYR:CE2	1:B:377:ARG:HD3	2.11	0.84
1:A:125:VAL:HG12	1:A:387:VAL:HG11	1.68	0.74
1:B:19:HIS:HD2	1:B:114:GLN:H	1.40	0.70
1:B:331[A]:TYR:CD2	1:B:377:ARG:HD3	2.26	0.68
1:B:331[A]:TYR:CE1	1:B:377:ARG:NH1	2.61	0.68
1:A:19:HIS:HD2	1:A:114:GLN:H	1.41	0.67
1:B:307:GLU:OE1	4:B:502:HOH:O	2.14	0.64
1:B:66[B]:HIS:HD2	4:B:650:HOH:O	1.82	0.63
1:A:248:LYS:HE3	4:A:593:HOH:O	2.04	0.57
1:A:24:LYS:HD2	1:A:87:TYR:OH	2.04	0.56
1:B:331[A]:TYR:CZ	1:B:377:ARG:HD3	2.42	0.54
1:B:10:PRO:HB3	1:B:201:GLU:CD	2.29	0.53
1:A:347:VAL:HG11	1:A:356:ILE:HD11	1.91	0.53
1:A:370:GLU:OE2	4:A:501:HOH:O	2.19	0.51
1:B:116:ARG:HG2	1:B:156:LEU:HD11	1.93	0.51
1:A:337:VAL:HG21	1:A:366:ILE:HD11	1.93	0.50
1:B:350:VAL:CG2	1:B:356:ILE:HD13	2.42	0.49
1:B:350:VAL:HG21	1:B:356:ILE:HD13	1.95	0.49
1:A:18:GLY:O	1:A:24:LYS:HE2	2.14	0.48
1:B:337:VAL:HG13	1:B:362:LEU:HD22	1.96	0.48
1:A:112:MET:HB3	1:A:113:PRO:HD2	1.95	0.47
1:B:213:PRO:HB3	1:B:333[A]:ARG:HD3	1.97	0.47
1:A:211:LEU:HD13	1:A:212:LEU:N	2.29	0.46
1:A:112:MET:HB3	1:A:113:PRO:CD	2.45	0.45
1:A:381:ARG:NH2	1:B:116:ARG:NH2	2.65	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:121:LEU:CD1	1:B:307:GLU:HG2	2.48	0.44
1:B:35:LEU:HD22	1:B:39:TYR:CD2	2.53	0.44
1:B:248:LYS:HE3	1:B:287:GLU:OE2	2.17	0.44
1:A:124:GLN:HB3	1:A:307:GLU:HG3	2.00	0.43
1:A:347:VAL:HG11	1:A:356:ILE:CD1	2.47	0.43
1:B:18:GLY:O	1:B:24:LYS:HE3	2.19	0.43
1:A:66[B]:HIS:CD2	1:A:79:VAL:HG22	2.53	0.43
1:A:378:GLU:OE1	1:A:381:ARG:NH2	2.52	0.42
1:B:35:LEU:HD22	1:B:39:TYR:CE2	2.54	0.42
1:A:134:LEU:HD11	1:A:149:VAL:HG11	2.01	0.42
1:A:279:ARG:NH2	4:A:510:HOH:O	2.51	0.42
1:A:19:HIS:CD2	1:A:114:GLN:H	2.29	0.41
1:B:19:HIS:CD2	1:B:114:GLN:H	2.28	0.41
1:A:202:PRO:HG2	1:A:204:ARG:HH21	1.86	0.41
1:A:299:THR:OG1	1:A:301:HIS:HE1	2.04	0.41
1:B:232:GLU:OE1	1:B:333[B]:ARG:NH1	2.54	0.40
1:B:261:PHE:O	1:B:262:ARG:HB2	2.21	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	360/393 (92%)	350 (97%)	9 (2%)	1 (0%)	41	23
1	B	364/393 (93%)	355 (98%)	8 (2%)	1 (0%)	41	23
All	All	724/786 (92%)	705 (97%)	17 (2%)	2 (0%)	41	23

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	262	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	333	ARG

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	300/325 (92%)	297 (99%)	3 (1%)	76	65
1	B	298/325 (92%)	296 (99%)	2 (1%)	84	76
All	All	598/650 (92%)	593 (99%)	5 (1%)	81	72

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	201	GLU
1	A	211	LEU
1	A	324	LYS
1	B	24	LYS
1	B	265	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	19	HIS
1	A	301	HIS
1	A	319	HIS
1	B	13	ASN
1	B	19	HIS
1	B	97	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
2	GDP	B	401	3	24,30,30	1.10	1 (4%)	31,47,47	1.58	6 (19%)
2	GDP	A	401	3	24,30,30	0.97	1 (4%)	31,47,47	1.68	6 (19%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GDP	B	401	3	-	1/12/32/32	0/3/3/3
2	GDP	A	401	3	-	2/12/32/32	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	401	GDP	C5-C6	3.43	1.47	1.41
2	A	401	GDP	C5-C6	2.51	1.45	1.41

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	401	GDP	C5-C6-N1	-4.29	117.56	123.43
2	B	401	GDP	C5-C6-N1	-4.16	117.74	123.43
2	A	401	GDP	C2-N1-C6	3.64	121.71	115.93
2	A	401	GDP	C4-C5-C6	-3.53	117.42	120.80
2	A	401	GDP	N3-C2-N1	-2.80	123.49	127.22
2	B	401	GDP	C1'-N9-C4	-2.79	121.74	126.64
2	B	401	GDP	C4-C5-N7	-2.69	106.59	109.40
2	B	401	GDP	C2-N1-C6	2.64	120.12	115.93
2	B	401	GDP	C3'-C2'-C1'	2.63	104.93	100.98
2	A	401	GDP	C1'-N9-C4	-2.28	122.63	126.64
2	A	401	GDP	PA-O3A-PB	-2.22	125.22	132.83
2	B	401	GDP	C4-C5-C6	-2.18	118.71	120.80

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	401	GDP	PA-O3A-PB-O2B
2	B	401	GDP	PA-O3A-PB-O2B
2	A	401	GDP	PA-O3A-PB-O1B

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand GDP B 401

Bond lengths

Bond angles

Torsions

Rings

Ligand GDP A 401

Bond lengths

Bond angles

Torsions

Rings
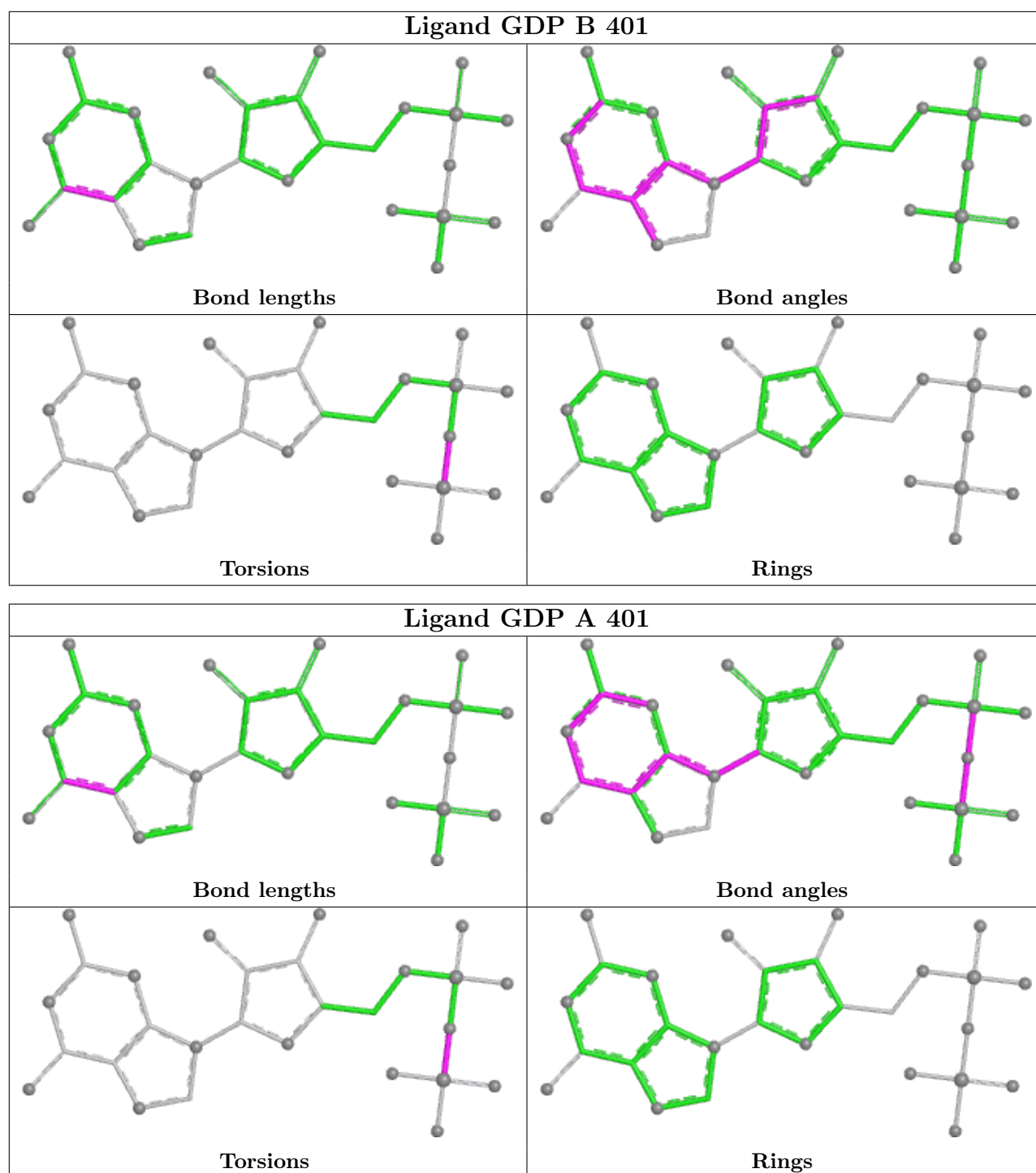
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	361/393 (91%)	0.40	13 (3%)	42	46	13, 23, 43, 62	0
1	B	366/393 (93%)	0.44	19 (5%)	27	28	13, 25, 45, 62	0
All	All	727/786 (92%)	0.42	32 (4%)	34	36	13, 24, 44, 62	0

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	393	ALA	5.2
1	B	265	LEU	5.0
1	A	204	ARG	4.8
1	A	220	ILE	4.4
1	B	222	GLY	4.3
1	A	222	GLY	4.0
1	B	95	ALA	3.5
1	B	221	SER	3.3
1	A	221	SER	3.3
1	A	104	VAL	3.3
1	B	392	ILE	3.2
1	B	220	ILE	3.1
1	A	11	HIS	3.1
1	B	223	ARG	3.0
1	B	202	PRO	3.0
1	A	10	PRO	2.9
1	B	218	PHE	2.8
1	B	224	GLY	2.6
1	B	104	VAL	2.5
1	B	11	HIS	2.5
1	B	105	VAL	2.5
1	A	218	PHE	2.5
1	B	10	PRO	2.3
1	B	201	GLU	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	266	ASP	2.2
1	A	105	VAL	2.2
1	A	141	ASP	2.2
1	B	333[A]	ARG	2.1
1	A	392	ILE	2.1
1	A	92	ILE	2.1
1	A	132	VAL	2.1
1	B	27	LEU	2.1

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

| Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	GDP	A	401	28/28	0.98	0.07	14,16,22,24	0
2	GDP	B	401	28/28	0.98	0.07	14,16,19,22	0
3	MG	A	402	1/1	0.98	0.04	16,16,16,16	0
3	MG	B	402	1/1	0.99	0.06	17,17,17,17	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around GDP A 401:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
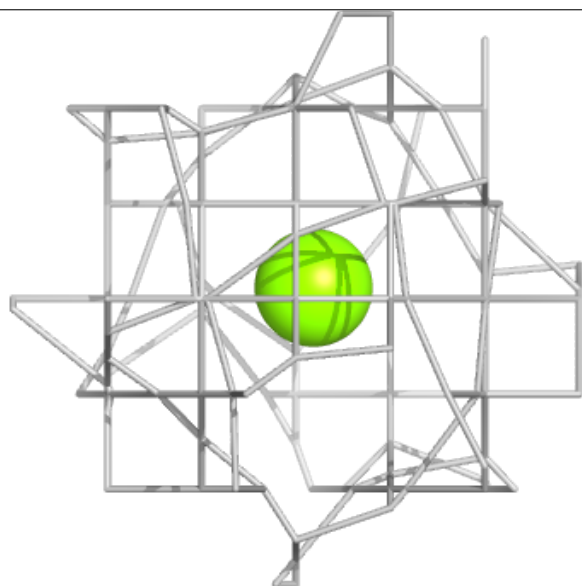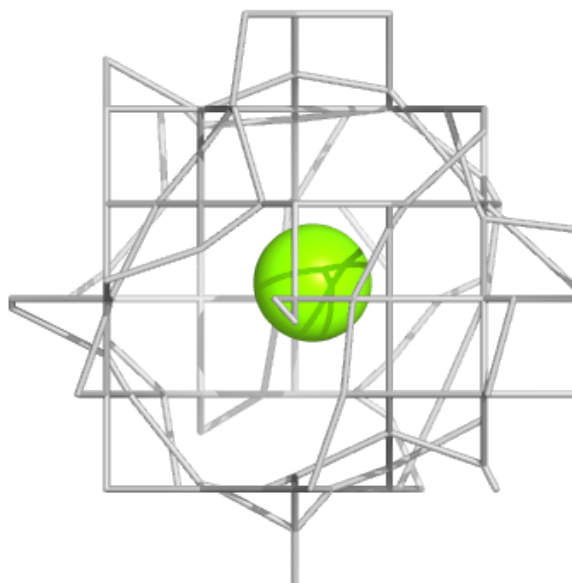mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
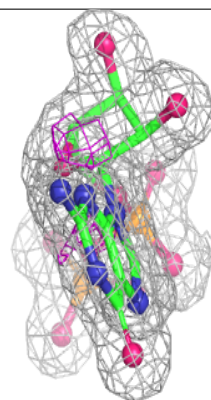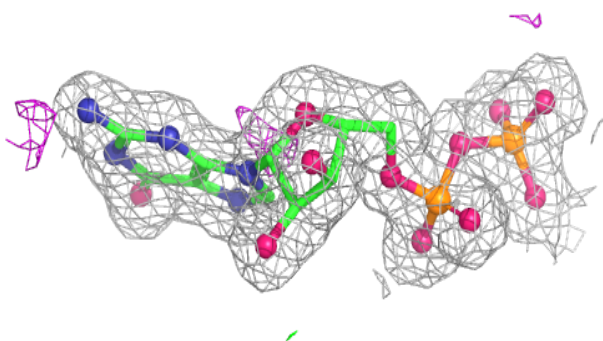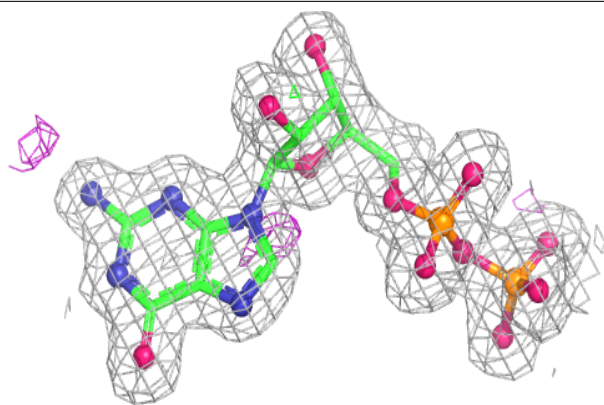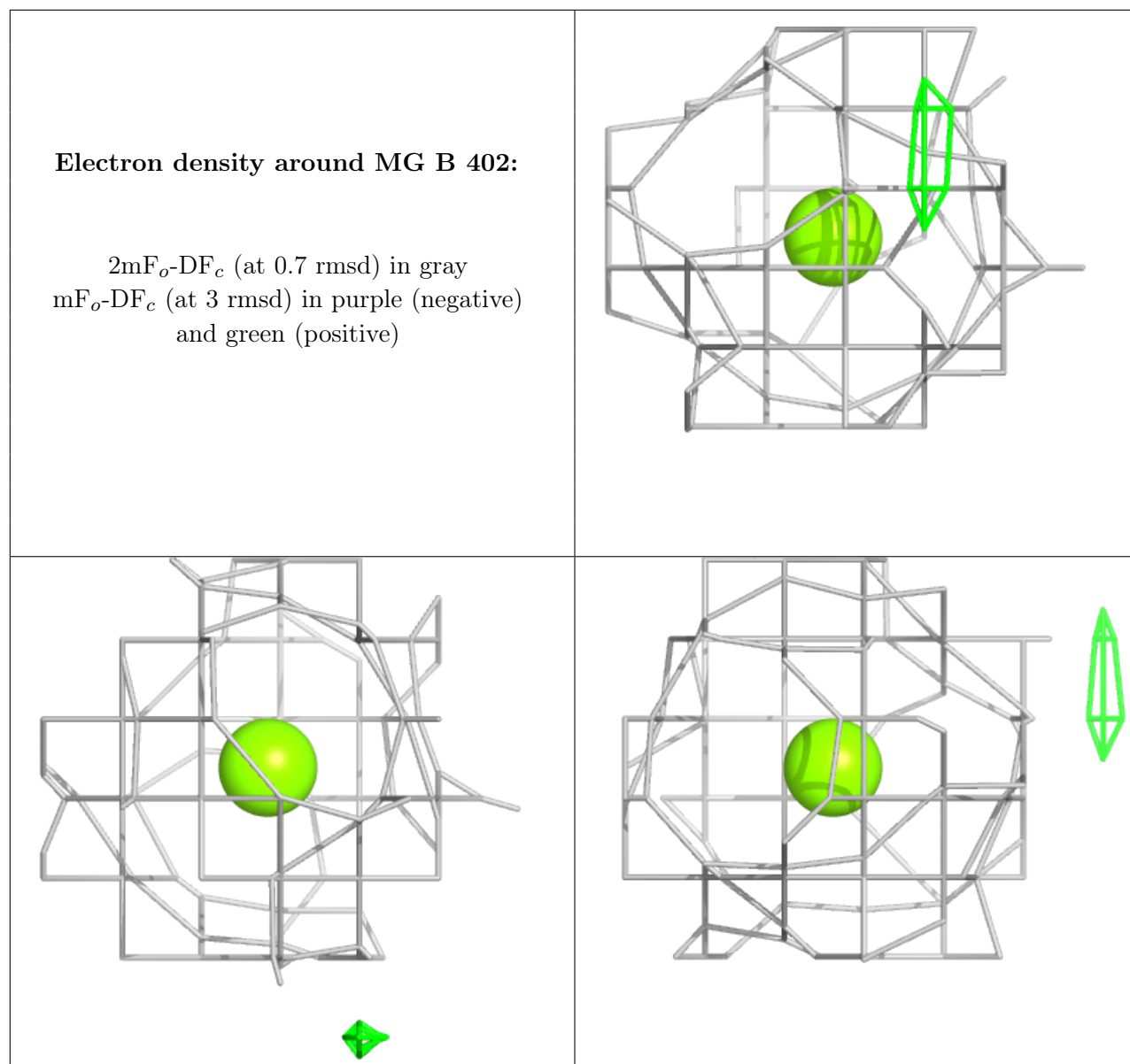

Electron density around GDP B 401:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
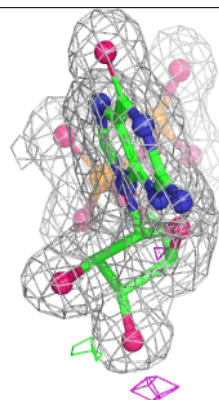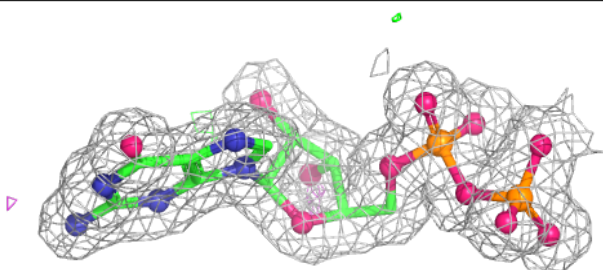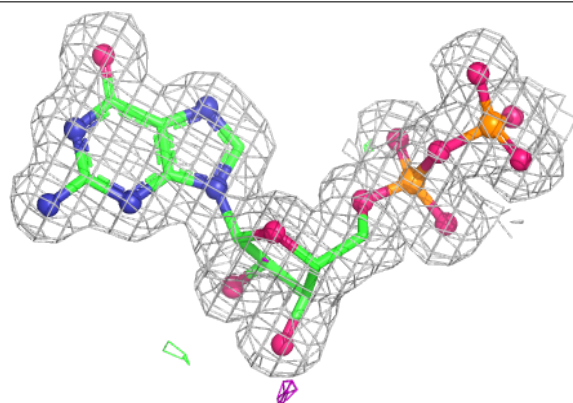
Electron density around MG A 402:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around MG B 402:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

6.5 Other polymers ⓘ

There are no such residues in this entry.