



wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 09:53 pm BST

PDB ID : 4P4S
Title : GMPPCP-bound stalkless-MxA
Authors : Rennie, M.L.; McKelvie, S.A.; Bulloch, E.M.; Kingston, R.L.
Deposited on : 2014-03-13
Resolution : 3.30 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

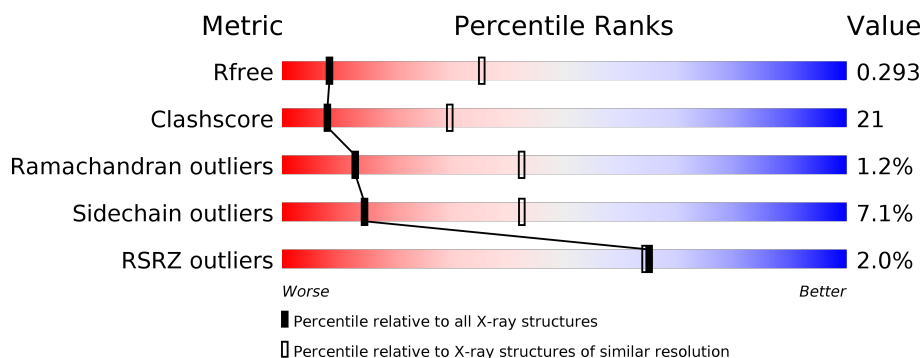
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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 on 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 $\leq 5\%$. 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	303	<div> <div>%</div> <div> <div></div> <div>68%</div> <div>30%</div> <div>.</div> </div> </div>
2	B	346	<div> <div>3%</div> <div> <div></div> <div>60%</div> <div>33%</div> <div>.</div> <div>.</div> </div> </div>

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 4962 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 Interferon-induced GTP-binding protein Mx1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	303	Total	C	N	O	S	0	0	0
			2257	1423	386	443	5			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	52	UNK	-	Remark 999	UNP P20591
A	53	UNK	-	Remark 999	UNP P20591
A	54	UNK	-	Remark 999	UNP P20591
A	55	UNK	-	Remark 999	UNP P20591
A	56	UNK	-	Remark 999	UNP P20591
A	57	UNK	-	Remark 999	UNP P20591
A	58	UNK	-	Remark 999	UNP P20591
A	59	UNK	-	Remark 999	UNP P20591
A	60	UNK	-	Remark 999	UNP P20591
A	61	UNK	-	Remark 999	UNP P20591
A	62	UNK	-	Remark 999	UNP P20591
A	63	UNK	-	Remark 999	UNP P20591
A	322	SER	CYS	engineered mutation	UNP P20591
A	336	SER	CYS	engineered mutation	UNP P20591
A	343	UNK	-	Remark 999	UNP P20591
A	344	UNK	-	Remark 999	UNP P20591
A	345	UNK	-	Remark 999	UNP P20591
A	346	UNK	-	Remark 999	UNP P20591
A	347	UNK	-	Remark 999	UNP P20591
A	348	UNK	-	Remark 999	UNP P20591
A	349	UNK	-	Remark 999	UNP P20591
A	350	UNK	-	Remark 999	UNP P20591
A	351	UNK	-	Remark 999	UNP P20591
A	352	UNK	-	Remark 999	UNP P20591
A	353	UNK	-	Remark 999	UNP P20591
A	354	UNK	-	Remark 999	UNP P20591
A	355	UNK	-	Remark 999	UNP P20591

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	356	UNK	-	Remark 999	UNP P20591
A	357	UNK	-	Remark 999	UNP P20591
A	358	UNK	-	Remark 999	UNP P20591
A	359	UNK	-	Remark 999	UNP P20591
A	360	UNK	-	Remark 999	UNP P20591

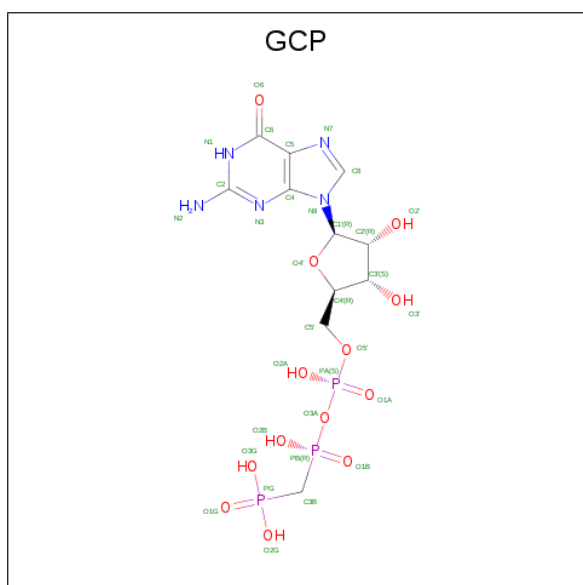
- Molecule 2 is a protein called Interferon-induced GTP-binding protein Mx1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	336	Total	C	N	O	S	0	0	0
			2622	1654	454	508	6			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	322	SER	CYS	engineered mutation	UNP P20591
B	336	SER	CYS	engineered mutation	UNP P20591
B	636	GLY	THR	engineered mutation	UNP P20591
B	637	ASP	SER	engineered mutation	UNP P20591
B	638	PRO	ASP	engineered mutation	UNP P20591
B	640	ALA	ARG	engineered mutation	UNP P20591

- Molecule 3 is PHOSPHOMETHYLPHOSPHONIC ACID GUANYLATE ESTER (three-letter code: GCP) (formula: C₁₁H₁₈N₅O₁₃P₃).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			32	11	5	13	3		
3	B	1	Total	C	N	O	P	0	0
			32	11	5	13	3		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

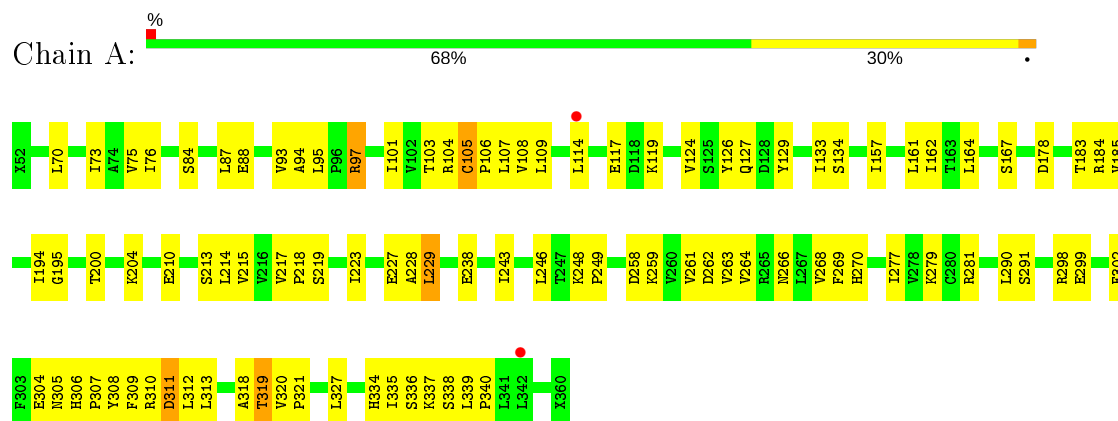
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	14	Total	O	0	0
			14	14		
5	B	3	Total	O	0	0
			3	3		

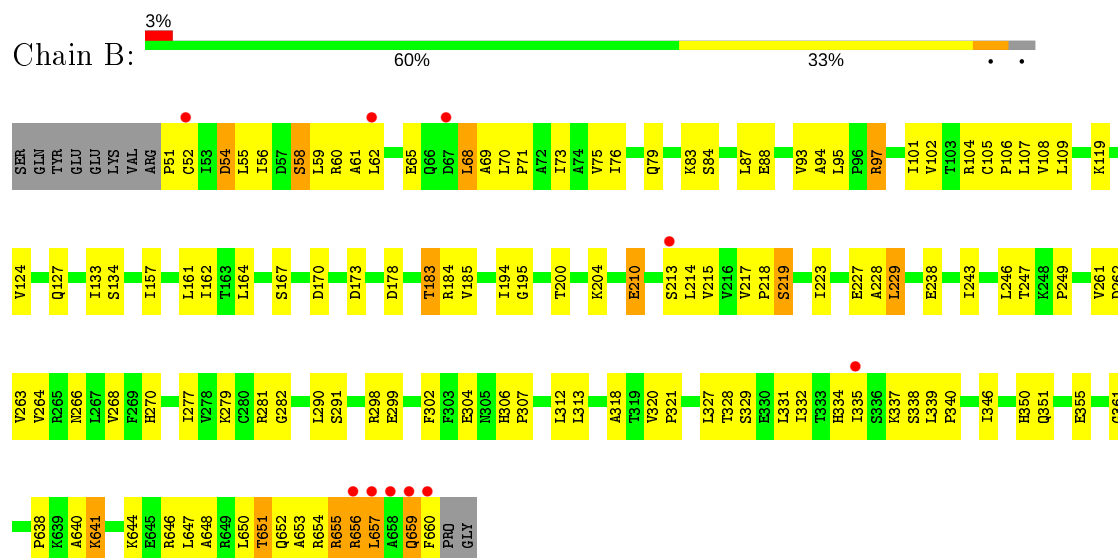
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: Interferon-induced GTP-binding protein Mx1



- Molecule 2: Interferon-induced GTP-binding protein Mx1



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	60.29 Å 66.76 Å 170.84 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.26 – 3.30 49.26 – 3.29	Depositor EDS
% Data completeness (in resolution range)	99.5 (49.26-3.30) 99.1 (49.26-3.29)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.78 (at 3.33 Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.231 , 0.305 0.229 , 0.293	Depositor DCC
R_{free} test set	529 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å ²)	75.8	Xtriage
Anisotropy	0.532	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 75.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	4962	wwPDB-VP
Average B, all atoms (Å ²)	84.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.28% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GCP, 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	0.62	0/2135	0.77	0/2888
2	B	0.57	0/2655	0.78	0/3584
All	All	0.59	0/4790	0.78	0/6472

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	2257	0	2219	68	0
2	B	2622	0	2721	143	0
3	A	32	0	14	2	0
3	B	32	0	14	2	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	14	0	0	2	0
5	B	3	0	0	1	0
All	All	4962	0	4968	212	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 21.

The worst 5 of 212 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)
2:B:646:ARG:O	2:B:650:LEU:CD1	1.79	1.28
1:A:117:GLU:OE1	1:A:119:LYS:HD2	1.34	1.25
2:B:70:LEU:HD11	2:B:335:ILE:CG1	1.68	1.24
1:A:117:GLU:OE1	1:A:119:LYS:CD	1.93	1.15
2:B:68:LEU:CD2	2:B:338:SER:CB	2.27	1.13

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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	272/303 (90%)	242 (89%)	26 (10%)	4 (2%)	10	38
2	B	334/346 (96%)	301 (90%)	30 (9%)	3 (1%)	17	48
All	All	606/649 (93%)	543 (90%)	56 (9%)	7 (1%)	13	42

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	361	GLY
1	A	127	GLN
2	B	127	GLN
1	A	319	THR
1	A	307	PRO

5.3.2 Protein sidechains [i](#)

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	239/239 (100%)	225 (94%)	14 (6%)	19	49
2	B	293/302 (97%)	269 (92%)	24 (8%)	11	36
All	All	532/541 (98%)	494 (93%)	38 (7%)	14	42

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	97	ARG
2	B	170	ASP
2	B	656	ARG
2	B	134	SER
2	B	183	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 carbohydrates 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	$\# Z > 2$
3	GCP	A	401	4	26,34,34	2.98	7 (26%)	31,54,54	1.66	6 (19%)
3	GCP	B	701	4	26,34,34	2.65	7 (26%)	31,54,54	1.74	5 (16%)

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
3	GCP	A	401	4	-	3/18/38/38	0/3/3/3
3	GCP	B	701	4	-	2/18/38/38	0/3/3/3

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	401	GCP	C4-N9	-11.76	1.32	1.47
3	B	701	GCP	C4-N9	-10.70	1.33	1.47
3	A	401	GCP	PG-O1G	5.73	1.62	1.50
3	A	401	GCP	C8-N9	-3.53	1.33	1.45
3	B	701	GCP	PG-O3G	3.52	1.63	1.54

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	701	GCP	PA-O3A-PB	-5.83	114.06	132.56
3	A	401	GCP	O4'-C1'-N9	4.18	115.26	109.04
3	A	401	GCP	PA-O3A-PB	-4.09	119.61	132.56
3	B	701	GCP	O3G-PG-C3B	3.90	115.86	106.40
3	B	701	GCP	C4-C5-N7	3.43	107.01	102.46

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	401	GCP	C2'-C1'-N9-C4
3	B	701	GCP	C2'-C1'-N9-C4
3	A	401	GCP	C3'-C4'-C5'-O5'
3	A	401	GCP	O4'-C4'-C5'-O5'

Continued on next page...

Continued from previous page...

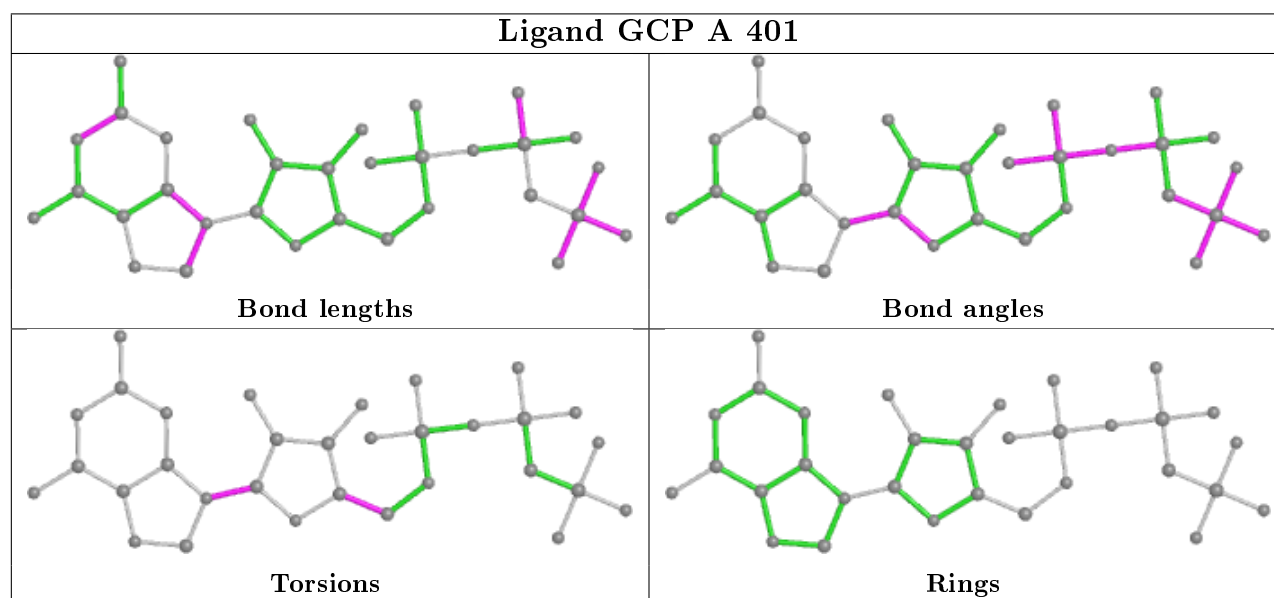
Mol	Chain	Res	Type	Atoms
3	B	701	GCP	C5'-O5'-PA-O1A

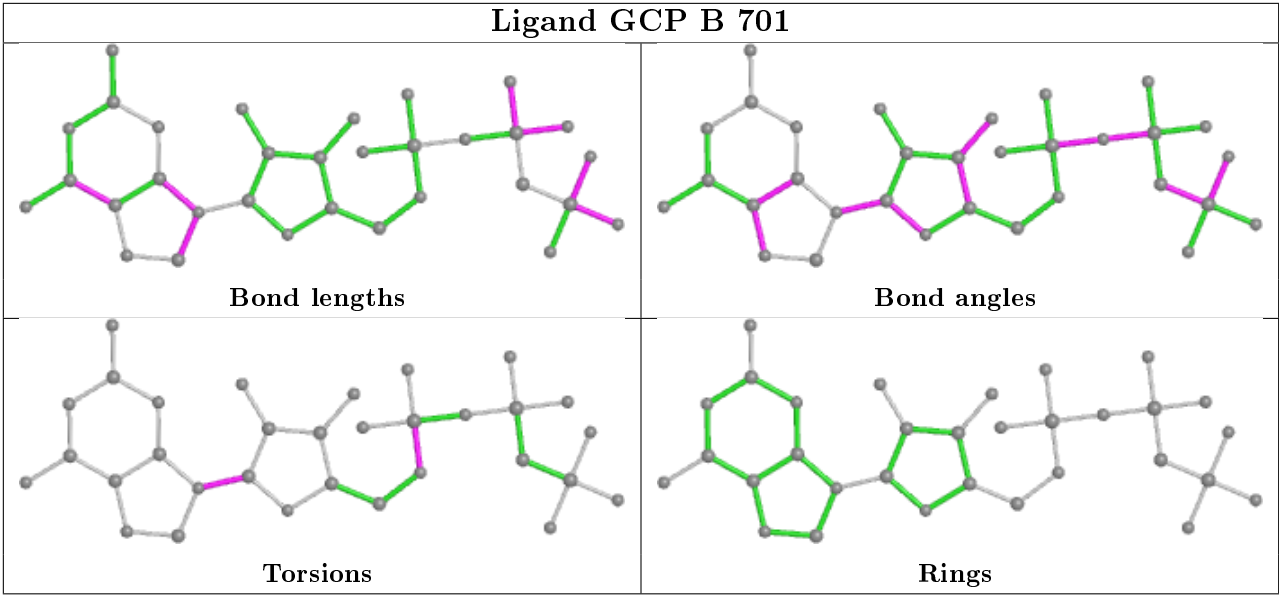
There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	401	GCP	2	0
3	B	701	GCP	2	0

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 than 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.





5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	63:UNK	C	70:LEU	N	8.19

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	273/303 (90%)	-0.01	2 (0%) 87 88	41, 73, 120, 162	0
2	B	336/346 (97%)	0.10	10 (2%) 50 49	46, 82, 142, 177	0
All	All	609/649 (93%)	0.05	12 (1%) 65 64	41, 78, 133, 177	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	660	PHE	3.1
2	B	658	ALA	2.8
2	B	657	LEU	2.8
2	B	335	ILE	2.6
2	B	52	CYS	2.6

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

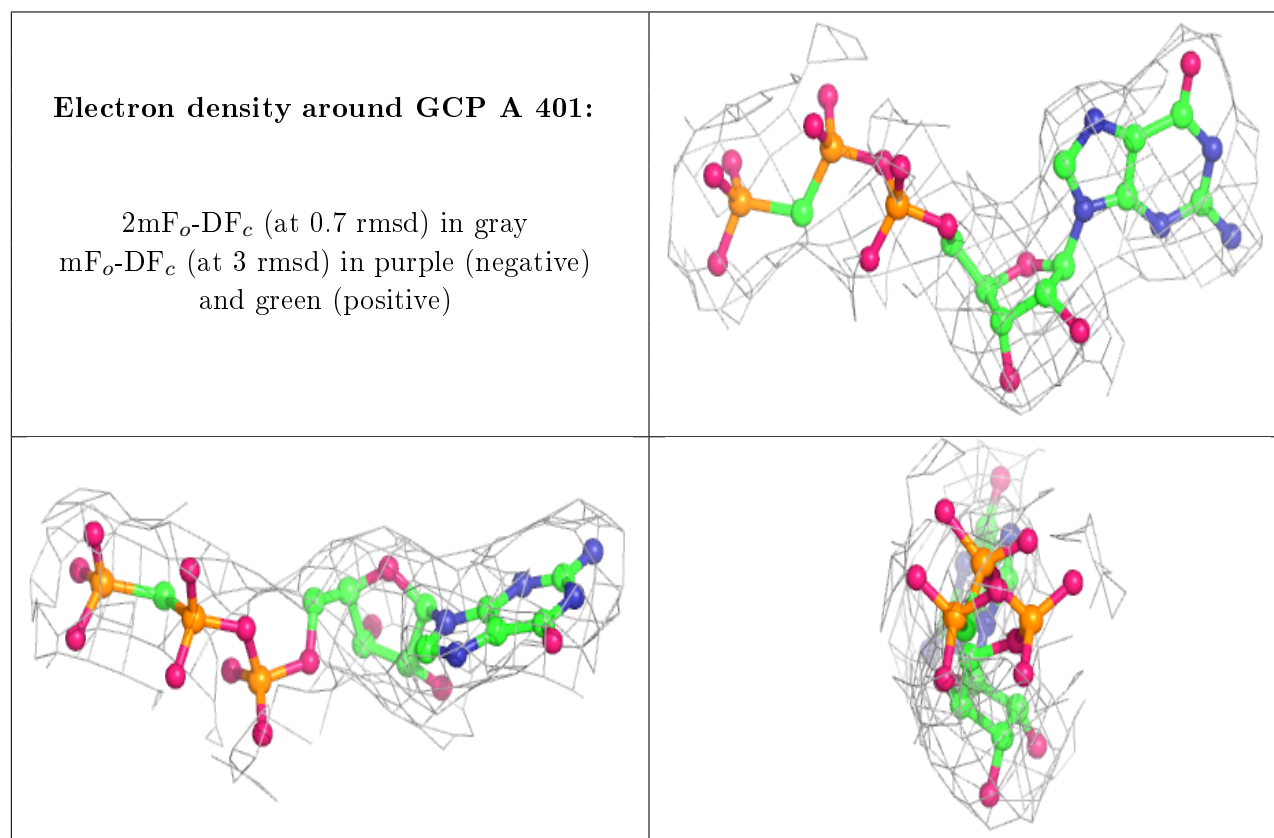
There are no carbohydrates in this entry.

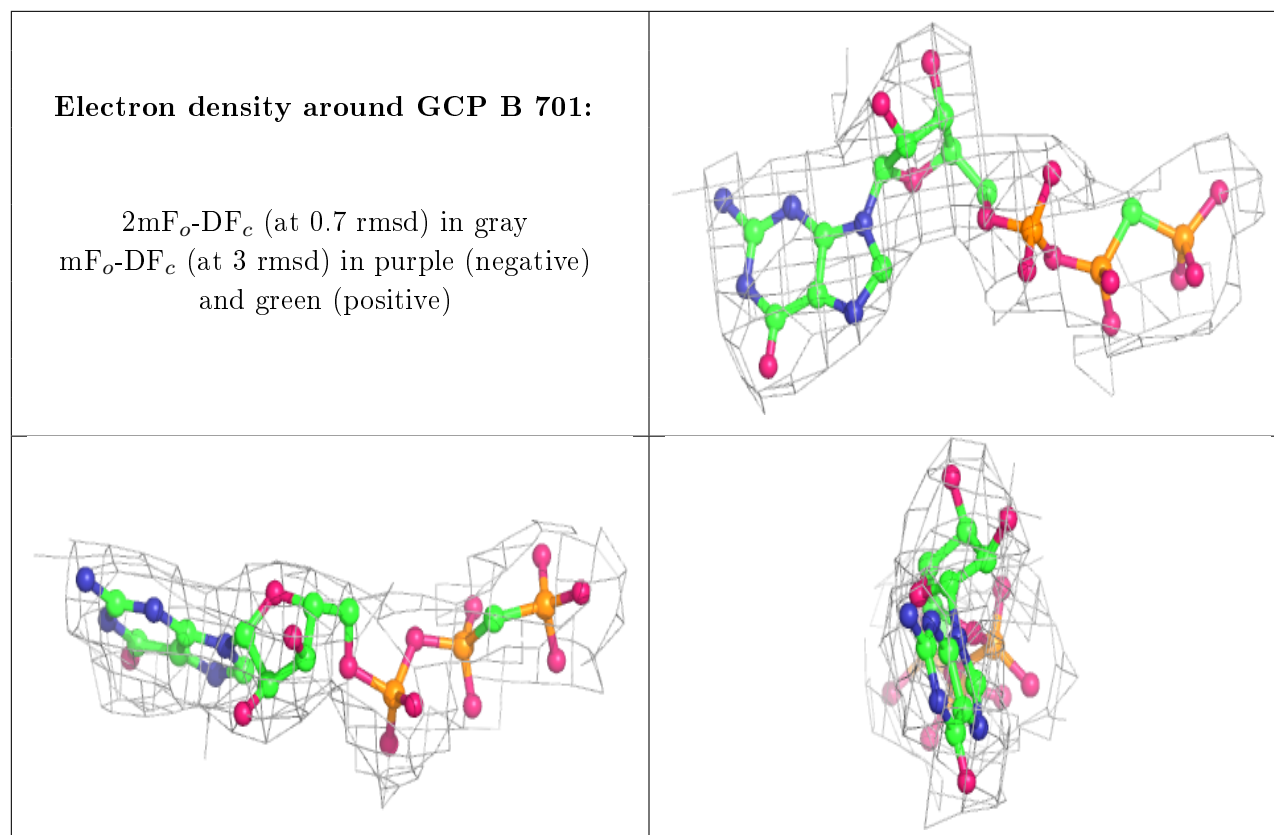
6.4 Ligands [i](#)

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(\AA^2)	Q<0.9
4	MG	B	702	1/1	0.94	0.06	43,43,43,43	0
4	MG	A	402	1/1	0.97	0.06	39,39,39,39	0
3	GCP	A	401	32/32	0.98	0.13	29,38,42,46	0
3	GCP	B	701	32/32	0.98	0.14	39,49,58,65	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.





6.5 Other polymers [i](#)

There are no such residues in this entry.