



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 27, 2022 – 10:22 pm BST

PDB ID : 7A9H
Title : Truncated 1-deoxy-D-xylulose 5-phosphate synthase (DXS) from Mycobacterium tuberculosis
Authors : Gierse, R.M.; Reddem, E.; Grooves, M.R.
Deposited on : 2020-09-02
Resolution : 1.85 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.31.2
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.31.2

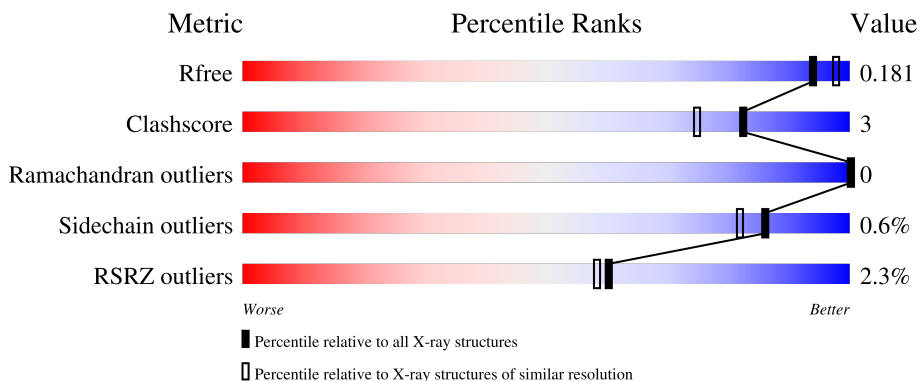
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.85 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	4003 (1.86-1.82)
Clashscore	141614	4233 (1.86-1.82)
Ramachandran outliers	138981	4185 (1.86-1.82)
Sidechain outliers	138945	4186 (1.86-1.82)
RSRZ outliers	127900	3957 (1.86-1.82)

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 $\leq 5\%$. 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	AAA	628	 2% 82% 5% 13%
1	BBB	628	 2% 82% 5% 13%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 17432 atoms, of which 8404 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 1-deoxy-D-xylulose-5-phosphate synthase,1-deoxy-D-xylulose-5-phosphate synthase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	AAA	545	8287	2594	4158	762	758	15	165	9	0
1	BBB	546	8393	2627	4214	770	765	17	167	16	1

There are 70 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	1	MET	-	initiating methionine	UNP P9WNS3
AAA	2	LYS	-	expression tag	UNP P9WNS3
AAA	3	HIS	-	expression tag	UNP P9WNS3
AAA	4	HIS	-	expression tag	UNP P9WNS3
AAA	5	HIS	-	expression tag	UNP P9WNS3
AAA	6	HIS	-	expression tag	UNP P9WNS3
AAA	7	HIS	-	expression tag	UNP P9WNS3
AAA	8	HIS	-	expression tag	UNP P9WNS3
AAA	9	PRO	-	expression tag	UNP P9WNS3
AAA	10	MET	-	expression tag	UNP P9WNS3
AAA	11	SER	-	expression tag	UNP P9WNS3
AAA	12	ASP	-	expression tag	UNP P9WNS3
AAA	13	TYR	-	expression tag	UNP P9WNS3
AAA	14	ASP	-	expression tag	UNP P9WNS3
AAA	15	ILE	-	expression tag	UNP P9WNS3
AAA	16	PRO	-	expression tag	UNP P9WNS3
AAA	17	THR	-	expression tag	UNP P9WNS3
AAA	18	THR	-	expression tag	UNP P9WNS3
AAA	19	GLU	-	expression tag	UNP P9WNS3
AAA	20	ASN	-	expression tag	UNP P9WNS3
AAA	21	LEU	-	expression tag	UNP P9WNS3
AAA	22	TYR	-	expression tag	UNP P9WNS3
AAA	23	PHE	-	expression tag	UNP P9WNS3
AAA	24	GLN	-	expression tag	UNP P9WNS3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	25	GLY	-	expression tag	UNP P9WNS3
AAA	26	ALA	-	expression tag	UNP P9WNS3
AAA	27	MET	-	expression tag	UNP P9WNS3
AAA	28	GLY	-	expression tag	UNP P9WNS3
AAA	218	GLY	-	linker	UNP P9WNS3
AAA	219	GLY	-	linker	UNP P9WNS3
AAA	220	GLY	-	linker	UNP P9WNS3
AAA	221	GLY	-	linker	UNP P9WNS3
AAA	222	GLY	-	linker	UNP P9WNS3
AAA	223	GLY	-	linker	UNP P9WNS3
AAA	224	GLY	-	linker	UNP P9WNS3
BBB	1	MET	-	initiating methionine	UNP P9WNS3
BBB	2	LYS	-	expression tag	UNP P9WNS3
BBB	3	HIS	-	expression tag	UNP P9WNS3
BBB	4	HIS	-	expression tag	UNP P9WNS3
BBB	5	HIS	-	expression tag	UNP P9WNS3
BBB	6	HIS	-	expression tag	UNP P9WNS3
BBB	7	HIS	-	expression tag	UNP P9WNS3
BBB	8	HIS	-	expression tag	UNP P9WNS3
BBB	9	PRO	-	expression tag	UNP P9WNS3
BBB	10	MET	-	expression tag	UNP P9WNS3
BBB	11	SER	-	expression tag	UNP P9WNS3
BBB	12	ASP	-	expression tag	UNP P9WNS3
BBB	13	TYR	-	expression tag	UNP P9WNS3
BBB	14	ASP	-	expression tag	UNP P9WNS3
BBB	15	ILE	-	expression tag	UNP P9WNS3
BBB	16	PRO	-	expression tag	UNP P9WNS3
BBB	17	THR	-	expression tag	UNP P9WNS3
BBB	18	THR	-	expression tag	UNP P9WNS3
BBB	19	GLU	-	expression tag	UNP P9WNS3
BBB	20	ASN	-	expression tag	UNP P9WNS3
BBB	21	LEU	-	expression tag	UNP P9WNS3
BBB	22	TYR	-	expression tag	UNP P9WNS3
BBB	23	PHE	-	expression tag	UNP P9WNS3
BBB	24	GLN	-	expression tag	UNP P9WNS3
BBB	25	GLY	-	expression tag	UNP P9WNS3
BBB	26	ALA	-	expression tag	UNP P9WNS3
BBB	27	MET	-	expression tag	UNP P9WNS3
BBB	28	GLY	-	expression tag	UNP P9WNS3
BBB	218	GLY	-	linker	UNP P9WNS3
BBB	219	GLY	-	linker	UNP P9WNS3
BBB	220	GLY	-	linker	UNP P9WNS3

Continued on next page...

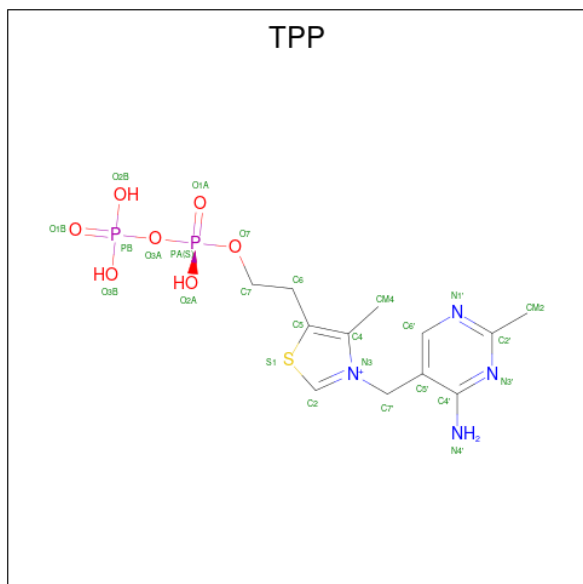
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	221	GLY	-	linker	UNP P9WNS3
BBB	222	GLY	-	linker	UNP P9WNS3
BBB	223	GLY	-	linker	UNP P9WNS3
BBB	224	GLY	-	linker	UNP P9WNS3

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	AAA	1	Total Mg 1 1	0	0
2	BBB	1	Total Mg 1 1	0	0

- Molecule 3 is THIAMINE DIPHOSPHATE (three-letter code: TPP) (formula: C₁₂H₁₉N₄O₇P₂S) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	
3	AAA	1	Total	C	H	N	O	P	S	2	0
			42	12	16	4	7	2	1		
3	BBB	1	Total	C	H	N	O	P	S	2	0
			42	12	16	4	7	2	1		

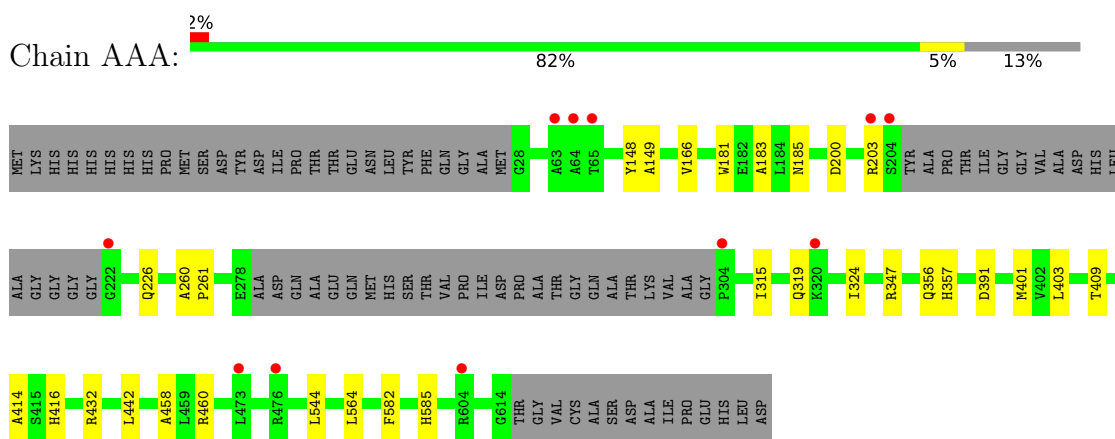
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	320	Total 320	O 320	0	1
4	BBB	346	Total 346	O 346	0	1

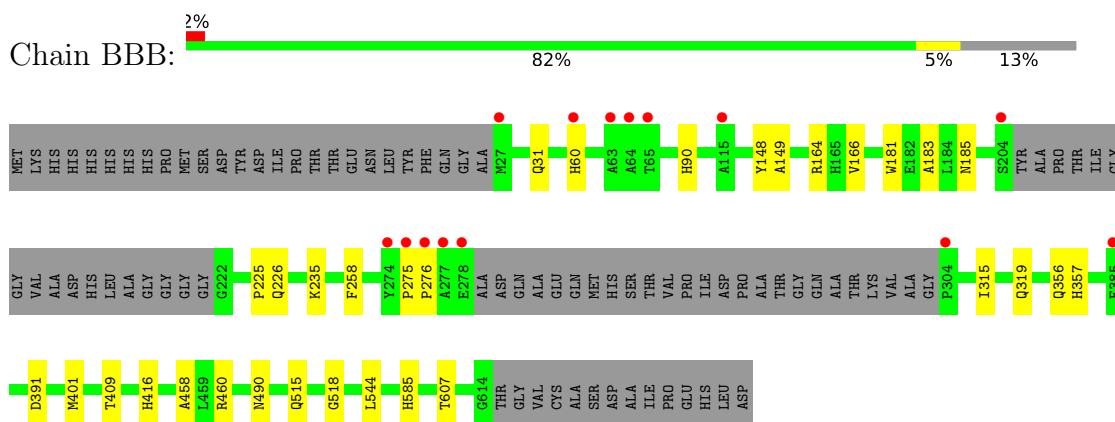
3 Residue-property plots [i](#)

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: 1-deoxy-D-xylulose-5-phosphate synthase,1-deoxy-D-xylulose-5-phosphate synthase



- Molecule 1: 1-deoxy-D-xylulose-5-phosphate synthase,1-deoxy-D-xylulose-5-phosphate synthase



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	62.99Å 76.10Å 79.15Å 108.96° 106.53° 107.59°	Depositor
Resolution (Å)	43.83 – 1.85 43.79 – 1.85	Depositor EDS
% Data completeness (in resolution range)	90.8 (43.83-1.85) 90.9 (43.79-1.85)	Depositor EDS
R_{merge}	0.02	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.00 (at 1.84Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.151 , 0.174 0.162 , 0.181	Depositor DCC
R_{free} test set	2101 reflections (2.25%)	wwPDB-VP
Wilson B-factor (Å ²)	18.8	Xtrriage
Anisotropy	0.216	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.006 for -h,-k,h+k+l	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	17432	wwPDB-VP
Average B, all atoms (Å ²)	20.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.29% 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: TPP, 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	AAA	0.68	0/4237	0.76	0/5750
1	BBB	0.68	0/4309	0.77	0/5844
All	All	0.68	0/8546	0.76	0/11594

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	AAA	4129	4158	4142	29	0
1	BBB	4179	4214	4203	30	0
2	AAA	1	0	0	0	0
2	BBB	1	0	0	0	0
3	AAA	26	16	16	0	0
3	BBB	26	16	16	0	0
4	AAA	320	0	0	1	0
4	BBB	346	0	0	1	0
All	All	9028	8404	8377	42	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 3.

The worst 5 of 42 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:AAA:226:GLN:H	1:BBB:185[B]:ASN:HD21	1.16	0.91
1:AAA:185[A]:ASN:HD21	1:BBB:226:GLN:H	1.18	0.88
1:AAA:226:GLN:H	1:BBB:185[B]:ASN:ND2	1.84	0.75
1:AAA:185[A]:ASN:ND2	1:BBB:226:GLN:H	1.85	0.74
1:AAA:356[B]:GLN:NE2	1:BBB:356[B]:GLN:NE2	2.39	0.70

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	AAA	548/628 (87%)	537 (98%)	11 (2%)	0	100	100
1	BBB	556/628 (88%)	544 (98%)	12 (2%)	0	100	100
All	All	1104/1256 (88%)	1081 (98%)	23 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	AAA	423/477 (89%)	421 (100%)	2 (0%)	88	85

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	BBB	431/477 (90%)	427 (99%)	4 (1%)	78	71
All	All	854/954 (90%)	848 (99%)	6 (1%)	86	78

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	BBB	60[A]	HIS
1	BBB	60[B]	HIS
1	BBB	460	ARG
1	AAA	460	ARG
1	AAA	432	ARG

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

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

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	TPP	BBB	702	2	22,27,27	0.60	0	29,40,40	0.97	2 (6%)
3	TPP	AAA	702	2	22,27,27	0.58	0	29,40,40	0.94	1 (3%)

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	TPP	BBB	702	2	-	4/16/17/17	0/2/2/2
3	TPP	AAA	702	2	-	4/16/17/17	0/2/2/2

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	AAA	702	TPP	C5-C4-N3	2.41	112.40	107.57
3	BBB	702	TPP	C5-C4-N3	2.41	112.40	107.57
3	BBB	702	TPP	C7'-N3-C2	-2.08	121.59	125.35

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

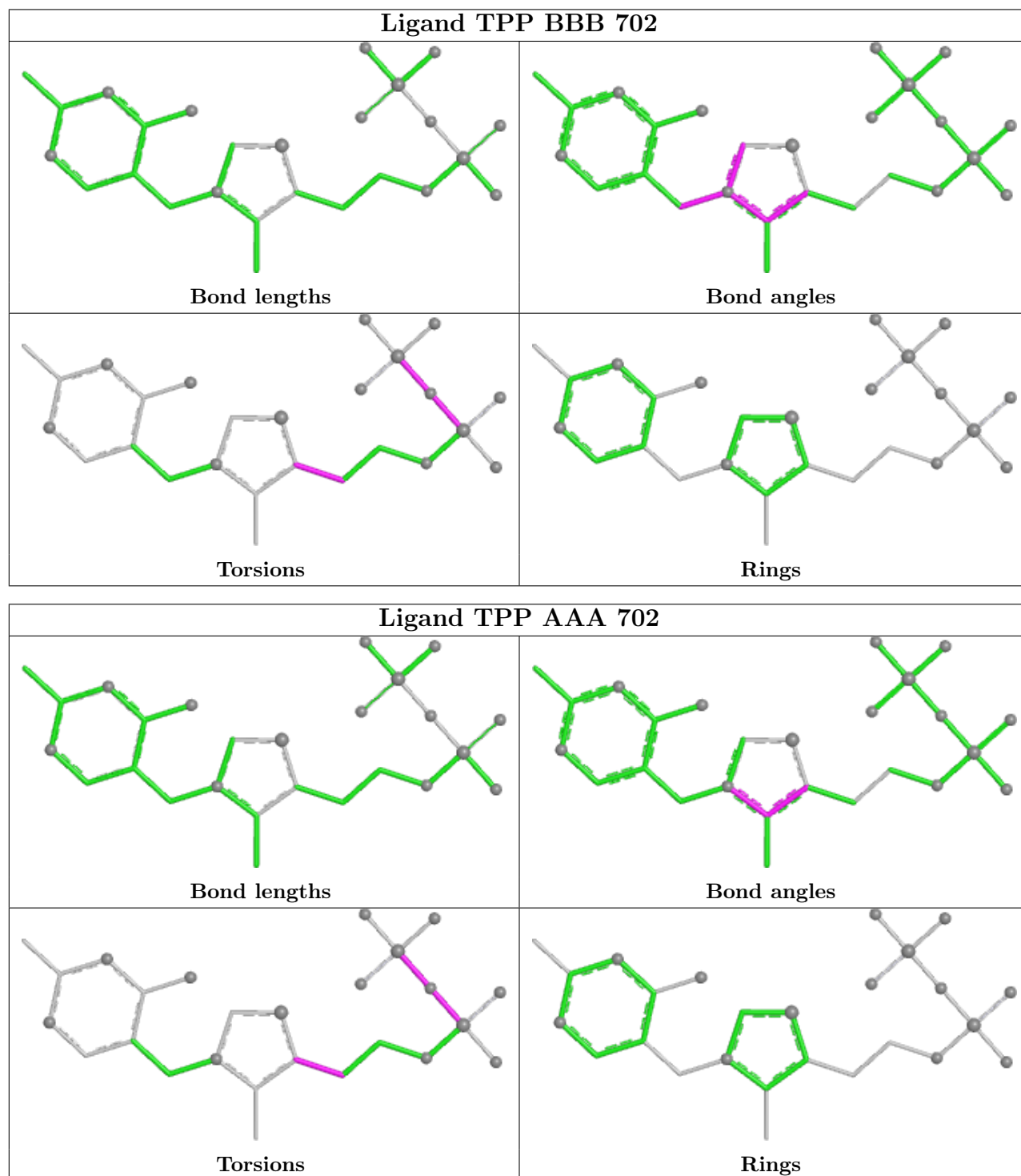
Mol	Chain	Res	Type	Atoms
3	AAA	702	TPP	PA-O3A-PB-O2B
3	BBB	702	TPP	PA-O3A-PB-O2B
3	AAA	702	TPP	C4-C5-C6-C7
3	BBB	702	TPP	C4-C5-C6-C7
3	BBB	702	TPP	PB-O3A-PA-O2A

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 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 [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

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	AAA	545/628 (86%)	-0.30	11 (2%) 65 64	11, 19, 38, 63	0
1	BBB	546/628 (86%)	-0.32	14 (2%) 56 53	11, 17, 35, 60	0
All	All	1091/1256 (86%)	-0.31	25 (2%) 60 58	11, 18, 37, 63	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	65	THR	4.5
1	AAA	63	ALA	4.1
1	BBB	63	ALA	4.0
1	AAA	65	THR	3.9
1	BBB	304	PRO	3.8

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](#)

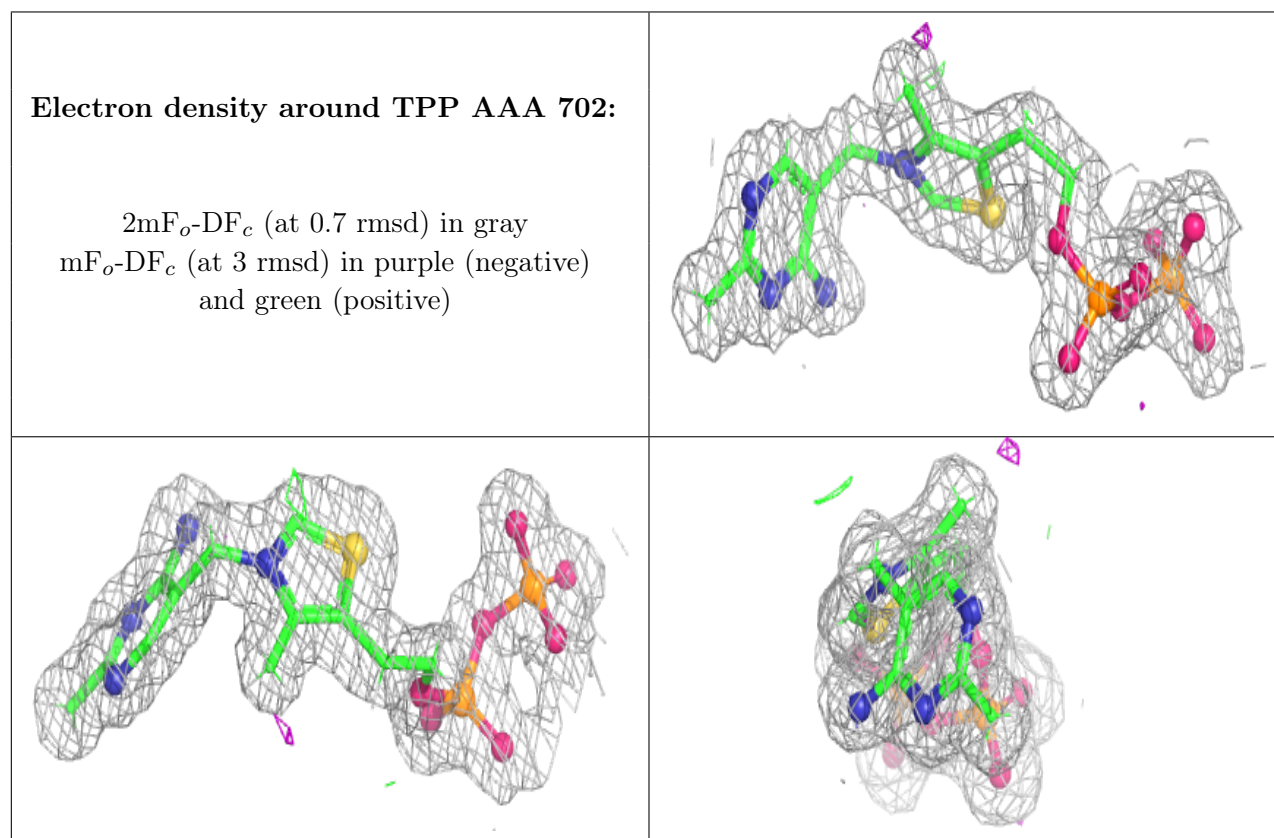
There are no monosaccharides in this entry.

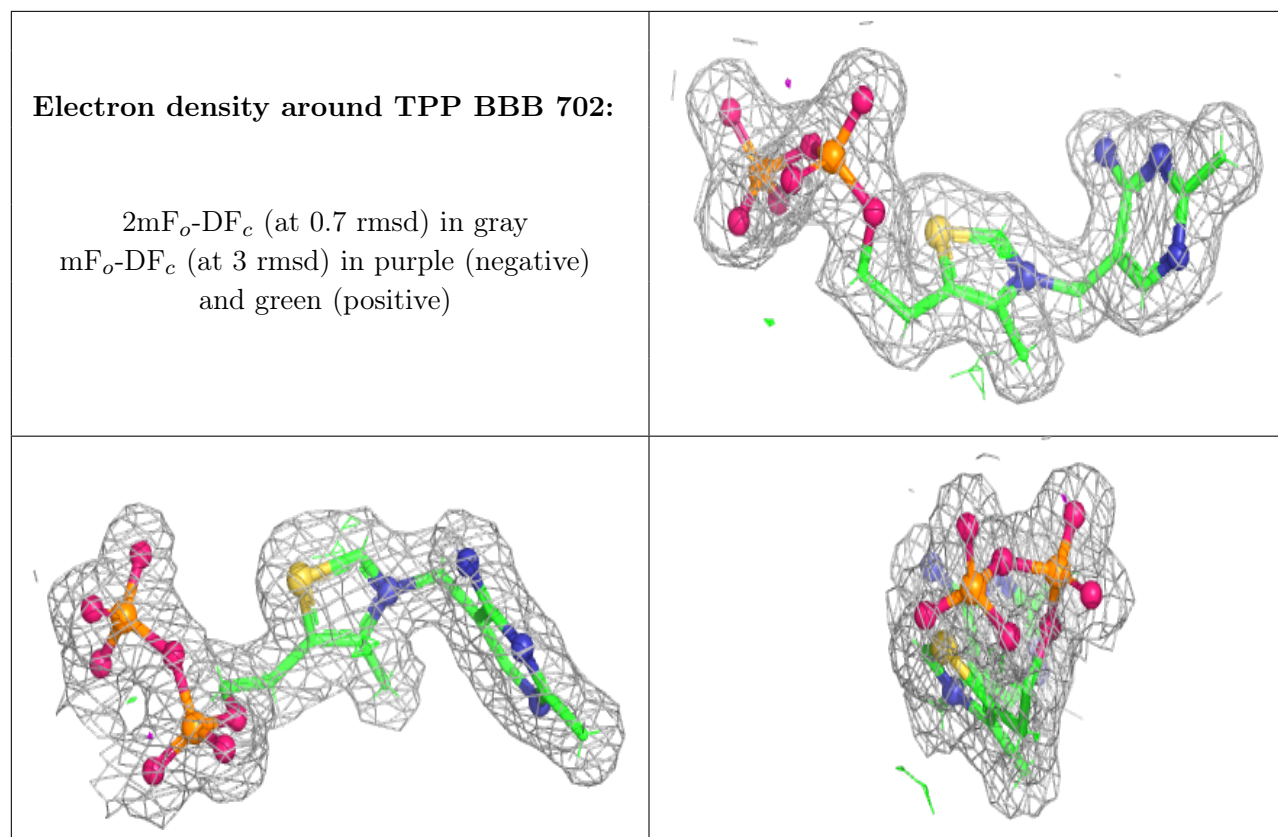
6.4 Ligands [i](#)

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
2	MG	BBB	701	1/1	0.99	0.06	11,11,11,11	0
3	TPP	AAA	702	26/26	0.99	0.07	12,14,17,22	2
3	TPP	BBB	702	26/26	0.99	0.08	12,13,16,18	2
2	MG	AAA	701	1/1	1.00	0.05	12,12,12,12	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.