



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:26 AM GMT

PDB ID : 2AG1
Title : Crystal structure of Benzaldehyde lyase (BAL)- SeMet
Authors : Mosbacher, T.G.; Mueller, M.; Schulz, G.E.
Deposited on : 2005-07-26
Resolution : 2.58 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
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

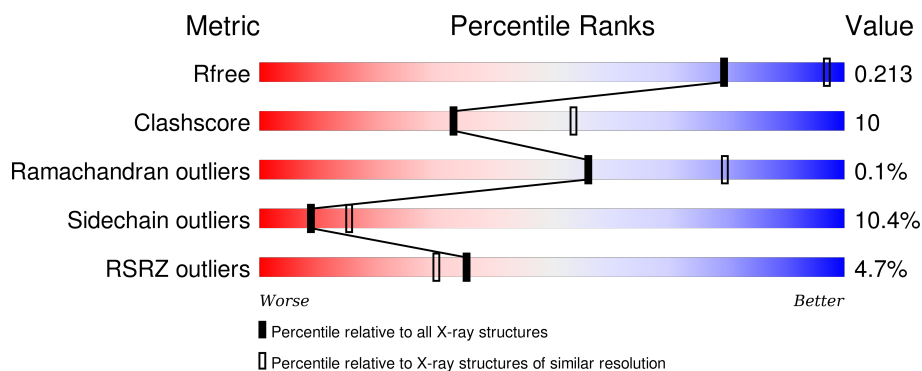
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.58 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}	91344	2636 (2.60-2.56)
Clashscore	102246	3003 (2.60-2.56)
Ramachandran outliers	100387	2956 (2.60-2.56)
Sidechain outliers	100360	2956 (2.60-2.56)
RSRZ outliers	91569	2642 (2.60-2.56)

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. 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	563	<div> <div>9%</div> <div> <div></div> <div>73%</div> <div>21%</div> <div>• •</div> </div> </div>
1	B	563	<div> <div>7%</div> <div> <div></div> <div>76%</div> <div>19%</div> <div>• •</div> </div> </div>
1	C	563	<div> <div>2%</div> <div> <div></div> <div>75%</div> <div>20%</div> <div>• •</div> </div> </div>
1	D	563	<div> <div>%</div> <div> <div></div> <div>74%</div> <div>21%</div> <div>• •</div> </div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 16808 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 benzaldehyde lyase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	553	Total	C	N	O	S	Se	0	0	0
			4074	2573	722	763	5	11			
1	B	553	Total	C	N	O	S	Se	0	0	0
			4074	2573	722	763	5	11			
1	C	553	Total	C	N	O	S	Se	0	0	0
			4074	2573	722	763	5	11			
1	D	553	Total	C	N	O	S	Se	0	0	0
			4074	2573	722	763	5	11			

There are 52 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	GB 9965498
A	3	MSE	MET	MODIFIED RESIDUE	GB 9965498
A	121	MSE	MET	MODIFIED RESIDUE	GB 9965498
A	133	MSE	MET	MODIFIED RESIDUE	GB 9965498
A	143	MSE	MET	MODIFIED RESIDUE	GB 9965498
A	167	MSE	MET	MODIFIED RESIDUE	GB 9965498
A	244	MSE	MET	MODIFIED RESIDUE	GB 9965498
A	252	MSE	MET	MODIFIED RESIDUE	GB 9965498
A	275	MSE	MET	MODIFIED RESIDUE	GB 9965498
A	404	MSE	MET	MODIFIED RESIDUE	GB 9965498
A	421	MSE	MET	MODIFIED RESIDUE	GB 9965498
A	473	MSE	MET	MODIFIED RESIDUE	GB 9965498
A	559	MSE	MET	MODIFIED RESIDUE	GB 9965498
B	1	MSE	MET	MODIFIED RESIDUE	GB 9965498
B	3	MSE	MET	MODIFIED RESIDUE	GB 9965498
B	121	MSE	MET	MODIFIED RESIDUE	GB 9965498
B	133	MSE	MET	MODIFIED RESIDUE	GB 9965498
B	143	MSE	MET	MODIFIED RESIDUE	GB 9965498
B	167	MSE	MET	MODIFIED RESIDUE	GB 9965498
B	244	MSE	MET	MODIFIED RESIDUE	GB 9965498
B	252	MSE	MET	MODIFIED RESIDUE	GB 9965498

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	275	MSE	MET	MODIFIED RESIDUE	GB 9965498
B	404	MSE	MET	MODIFIED RESIDUE	GB 9965498
B	421	MSE	MET	MODIFIED RESIDUE	GB 9965498
B	473	MSE	MET	MODIFIED RESIDUE	GB 9965498
B	559	MSE	MET	MODIFIED RESIDUE	GB 9965498
C	1	MSE	MET	MODIFIED RESIDUE	GB 9965498
C	3	MSE	MET	MODIFIED RESIDUE	GB 9965498
C	121	MSE	MET	MODIFIED RESIDUE	GB 9965498
C	133	MSE	MET	MODIFIED RESIDUE	GB 9965498
C	143	MSE	MET	MODIFIED RESIDUE	GB 9965498
C	167	MSE	MET	MODIFIED RESIDUE	GB 9965498
C	244	MSE	MET	MODIFIED RESIDUE	GB 9965498
C	252	MSE	MET	MODIFIED RESIDUE	GB 9965498
C	275	MSE	MET	MODIFIED RESIDUE	GB 9965498
C	404	MSE	MET	MODIFIED RESIDUE	GB 9965498
C	421	MSE	MET	MODIFIED RESIDUE	GB 9965498
C	473	MSE	MET	MODIFIED RESIDUE	GB 9965498
C	559	MSE	MET	MODIFIED RESIDUE	GB 9965498
D	1	MSE	MET	MODIFIED RESIDUE	GB 9965498
D	3	MSE	MET	MODIFIED RESIDUE	GB 9965498
D	121	MSE	MET	MODIFIED RESIDUE	GB 9965498
D	133	MSE	MET	MODIFIED RESIDUE	GB 9965498
D	143	MSE	MET	MODIFIED RESIDUE	GB 9965498
D	167	MSE	MET	MODIFIED RESIDUE	GB 9965498
D	244	MSE	MET	MODIFIED RESIDUE	GB 9965498
D	252	MSE	MET	MODIFIED RESIDUE	GB 9965498
D	275	MSE	MET	MODIFIED RESIDUE	GB 9965498
D	404	MSE	MET	MODIFIED RESIDUE	GB 9965498
D	421	MSE	MET	MODIFIED RESIDUE	GB 9965498
D	473	MSE	MET	MODIFIED RESIDUE	GB 9965498
D	559	MSE	MET	MODIFIED RESIDUE	GB 9965498

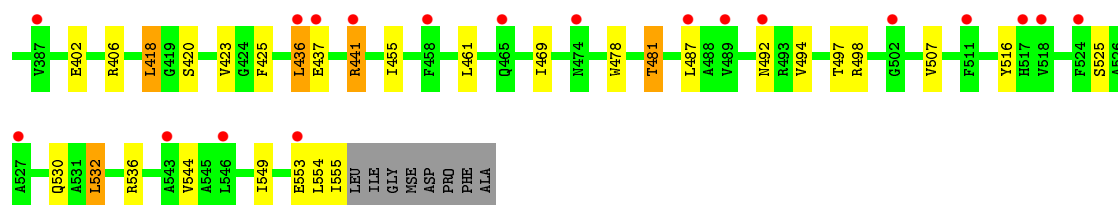
- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Mg 1 1	0	0
2	A	1	Total Mg 1 1	0	0
2	D	1	Total Mg 1 1	0	0
2	C	1	Total Mg 1 1	0	0

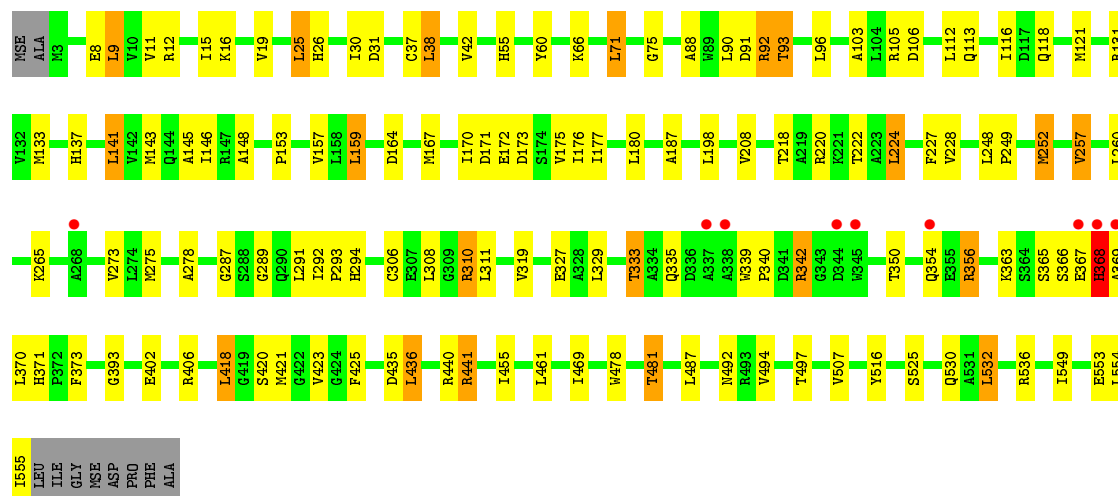
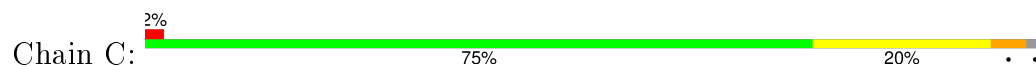
-
- Chemical structure of TPP (Thiamine Pyrophosphate) showing the pyrophosphate group, thiazolium ring, and pyrimidine ring with atom labels.

- Molecule 4 is water.

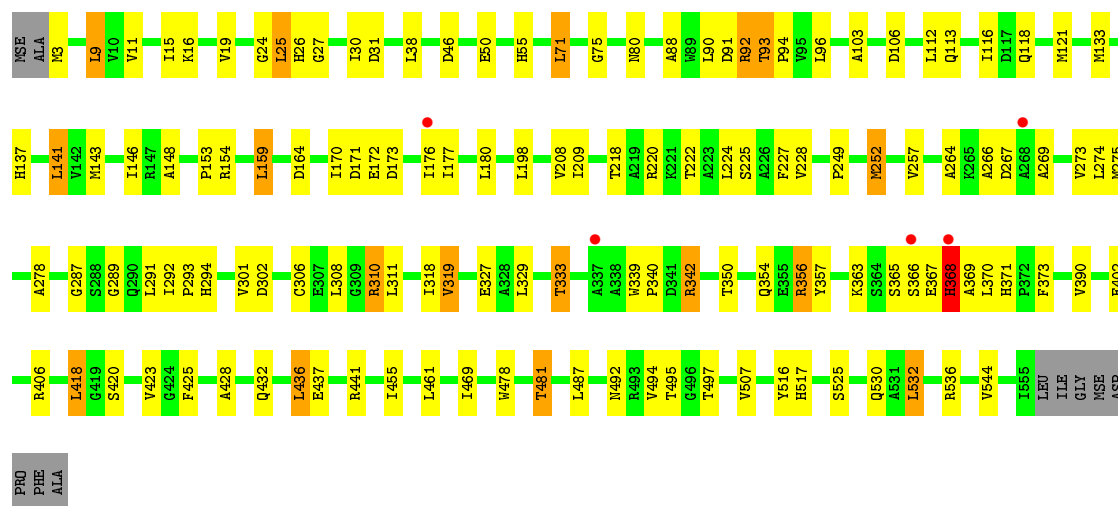
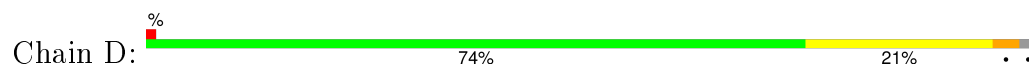
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	82	Total O 82 82	0	0
4	B	62	Total O 62 62	0	0
4	C	114	Total O 114 114	0	0
4	D	146	Total O 146 146	0	0



• Molecule 1: benzaldehyde lyase



• Molecule 1: benzaldehyde lyase



4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	150.16Å 150.16Å 195.61Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	24.82 – 2.58 24.62 – 2.58	Depositor EDS
% Data completeness (in resolution range)	100.0 (24.82-2.58) 100.0 (24.62-2.58)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	127.44 (at 2.57Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.212 , 0.245 0.215 , 0.213	Depositor DCC
R_{free} test set	4048 reflections (5.29%)	DCC
Wilson B-factor (Å ²)	37.8	Xtriage
Anisotropy	0.095	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 14.7	EDS
Estimated twinning fraction	0.079 for -h,-k,l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.38$, $\langle L^2 \rangle = 0.20$	Xtriage
Outliers	0 of 80506 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	16808	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.97% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: MG, TPP

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	0.58	0/4147	0.71	3/5640 (0.1%)
1	B	0.57	1/4147 (0.0%)	0.72	2/5640 (0.0%)
1	C	0.65	1/4147 (0.0%)	0.76	4/5640 (0.1%)
1	D	0.70	1/4147 (0.0%)	0.76	2/5640 (0.0%)
All	All	0.63	3/16588 (0.0%)	0.74	11/22560 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
1	B	0	1
1	C	0	2
1	D	0	2
All	All	0	7

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	370	LEU	C-N	-5.94	1.20	1.34
1	D	370	LEU	C-N	-5.58	1.21	1.34
1	C	370	LEU	C-N	-5.24	1.22	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	92	ARG	NE-CZ-NH1	7.52	124.06	120.30
1	A	92	ARG	NE-CZ-NH1	6.69	123.64	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	25	LEU	CA-CB-CG	6.46	130.16	115.30
1	D	92	ARG	NE-CZ-NH1	6.43	123.52	120.30
1	D	25	LEU	CA-CB-CG	5.92	128.91	115.30

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	365	SER	Peptide
1	A	368	HIS	Peptide
1	B	368	HIS	Peptide
1	C	365	SER	Peptide
1	C	368	HIS	Peptide

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	4074	0	4064	95	3
1	B	4074	0	4064	74	1
1	C	4074	0	4064	84	0
1	D	4074	0	4064	90	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	26	0	16	4	0
3	B	26	0	16	1	0
3	C	26	0	16	1	0
3	D	26	0	16	1	0
4	A	82	0	0	11	0
4	B	62	0	0	7	0
4	C	114	0	0	7	0
4	D	146	0	0	14	0
All	All	16808	0	16320	336	3

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 10.

The worst 5 of 336 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:C:167:MSE:SE	1:C:167:MSE:CE	2.14	1.45
1:D:133:MSE:SE	1:D:133:MSE:CE	2.16	1.43
1:A:133:MSE:CE	1:A:133:MSE:SE	2.19	1.40
1:A:167:MSE:SE	1:A:167:MSE:CE	2.19	1.40
1:B:133:MSE:CE	1:B:133:MSE:SE	2.18	1.39

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:265:LYS:O	1:A:265:LYS:O[6_555]	1.92	0.28
1:A:406:ARG:NH2	1:B:267:ASP:OD1[5_565]	2.09	0.11
1:A:348:LYS:CD	1:A:355:GLU:OE2[6_555]	2.10	0.10

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	551/563 (98%)	529 (96%)	21 (4%)	1 (0%)	52	75
1	B	551/563 (98%)	528 (96%)	22 (4%)	1 (0%)	52	75
1	C	551/563 (98%)	530 (96%)	20 (4%)	1 (0%)	52	75
1	D	551/563 (98%)	530 (96%)	21 (4%)	0	100	100
All	All	2204/2252 (98%)	2117 (96%)	84 (4%)	3 (0%)	56	79

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	554	LEU
1	C	554	LEU
1	B	554	LEU

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	412/406 (102%)	370 (90%)	42 (10%)	9	16
1	B	412/406 (102%)	369 (90%)	43 (10%)	9	15
1	C	412/406 (102%)	369 (90%)	43 (10%)	9	15
1	D	412/406 (102%)	368 (89%)	44 (11%)	8	14
All	All	1648/1624 (102%)	1476 (90%)	172 (10%)	9	15

5 of 172 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	481	THR
1	C	175	VAL
1	D	425	PHE
1	B	494	VAL
1	C	25	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 58 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	492	ASN
1	C	118	GLN
1	D	331	GLN
1	B	530	GLN
1	C	29	HIS

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 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 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	A	602	2	20,27,27	1.95	3 (15%)	31,40,40	2.39	12 (38%)
3	TPP	B	612	2	20,27,27	2.31	4 (20%)	31,40,40	2.95	9 (29%)
3	TPP	C	622	2	20,27,27	2.40	3 (15%)	31,40,40	2.13	10 (32%)
3	TPP	D	632	2	20,27,27	2.13	3 (15%)	31,40,40	2.91	9 (29%)

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	A	602	2	-	0/16/17/17	0/2/2/2
3	TPP	B	612	2	-	0/16/17/17	0/2/2/2
3	TPP	C	622	2	-	0/16/17/17	0/2/2/2
3	TPP	D	632	2	-	0/16/17/17	0/2/2/2

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	622	TPP	CM4-C4	-8.14	1.32	1.49
3	A	602	TPP	CM4-C4	-7.60	1.33	1.49
3	B	612	TPP	CM4-C4	-7.50	1.33	1.49
3	D	632	TPP	CM4-C4	-7.15	1.34	1.49
3	C	622	TPP	C4-N3	-5.65	1.34	1.39

The worst 5 of 40 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	632	TPP	C6-C5-C4	-11.12	117.59	127.56
3	B	612	TPP	C6-C5-C4	-9.48	119.07	127.56
3	C	622	TPP	C6-C5-C4	-6.00	122.18	127.56
3	A	602	TPP	C6-C5-C4	-5.59	122.56	127.56
3	A	602	TPP	O3A-PA-O7	-3.95	92.46	102.94

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	602	TPP	4	0
3	B	612	TPP	1	0
3	C	622	TPP	1	0
3	D	632	TPP	1	0

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	A	542/563 (96%)	0.56	49 (9%) 12 9	18, 49, 92, 123	0
1	B	542/563 (96%)	0.57	38 (7%) 19 15	21, 50, 83, 103	0
1	C	542/563 (96%)	0.09	9 (1%) 73 69	13, 31, 59, 77	0
1	D	542/563 (96%)	-0.02	5 (0%) 85 84	14, 28, 44, 63	0
All	All	2168/2252 (96%)	0.30	101 (4%) 35 30	13, 37, 81, 123	0

The worst 5 of 101 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	265	LYS	7.2
1	A	368	HIS	6.7
1	B	368	HIS	6.1
1	C	337	ALA	5.4
1	C	368	HIS	5.3

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
3	TPP	D	632	26/26	0.98	0.14	0.27	18,21,26,27	0
3	TPP	C	622	26/26	0.97	0.14	-0.20	25,35,41,42	0
3	TPP	A	602	26/26	0.91	0.17	-0.24	47,54,61,61	0
3	TPP	B	612	26/26	0.94	0.13	-0.98	30,42,50,51	0
2	MG	D	631	1/1	0.96	0.09	-1.05	20,20,20,20	0
2	MG	C	621	1/1	0.97	0.08	-1.99	33,33,33,33	0
2	MG	A	601	1/1	0.98	0.08	-2.22	55,55,55,55	0
2	MG	B	611	1/1	0.94	0.04	-5.09	44,44,44,44	0

6.5 Other polymers [i](#)

There are no such residues in this entry.