



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 25, 2023 – 07:05 pm BST

PDB ID : 8B4M
Title : Crystal structure of a hydroxyphenyl synthase (M75L variant) in its closed conformation
Authors : Helmer, C.P.O.; Driller, R.; Loll, B.
Deposited on : 2022-09-20
Resolution : 3.04 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 : 2.35.1
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) : 2.35.1

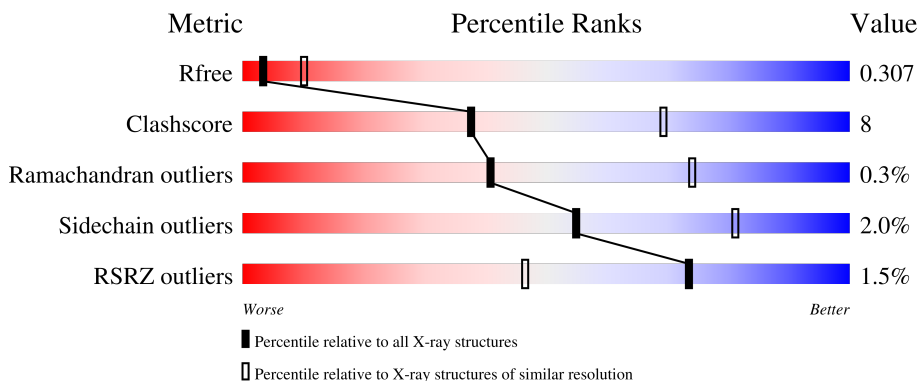
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.04 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.



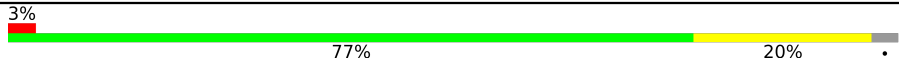
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2752 (3.08-3.00)
Clashscore	141614	3096 (3.08-3.00)
Ramachandran outliers	138981	2986 (3.08-3.00)
Sidechain outliers	138945	2988 (3.08-3.00)
RSRZ outliers	127900	2636 (3.08-3.00)

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	A	322	 77% 18% 5% 0%
1	B	322	 79% 16% 5% 0%
1	C	322	 75% 18% 6% 0%
1	D	322	 76% 19% 5% 0%
1	E	322	 72% 23% 5% 2%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	322	 <p>A horizontal bar chart representing the quality of the chain. The bar is divided into three segments: a small red segment on the left labeled '3%', a large green segment in the middle labeled '77%', and a yellow segment on the right labeled '20%'. A small grey dot is visible at the far right end of the bar.</p>

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 14633 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 Terpene synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	308	Total 2425	C 1521	N 441	O 451	S 12	0	0	0
1	B	307	Total 2433	C 1528	N 446	O 448	S 11	0	0	0
1	C	302	Total 2388	C 1499	N 435	O 442	S 12	0	0	0
1	D	305	Total 2418	C 1519	N 444	O 444	S 11	0	0	0
1	E	308	Total 2443	C 1532	N 445	O 454	S 12	0	1	0
1	F	311	Total 2458	C 1540	N 450	O 456	S 12	0	0	0

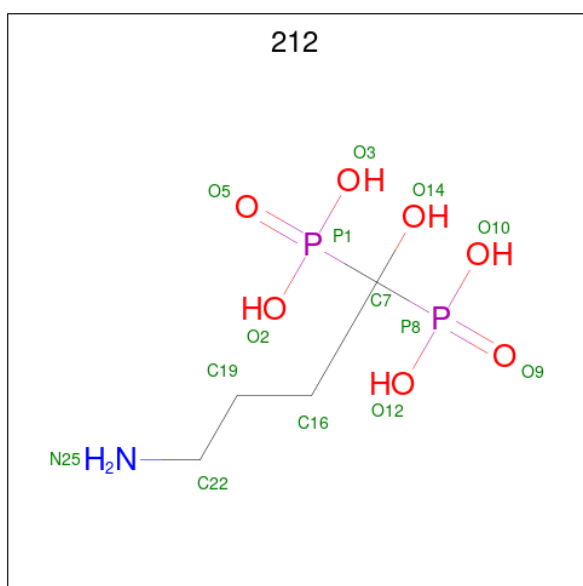
There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	ASP	-	expression tag	UNP D5SK09
A	75	LEU	MET	engineered mutation	UNP D5SK09
B	0	ASP	-	expression tag	UNP D5SK09
B	75	LEU	MET	engineered mutation	UNP D5SK09
C	0	ASP	-	expression tag	UNP D5SK09
C	75	LEU	MET	engineered mutation	UNP D5SK09
D	0	ASP	-	expression tag	UNP D5SK09
D	75	LEU	MET	engineered mutation	UNP D5SK09
E	0	ASP	-	expression tag	UNP D5SK09
E	75	LEU	MET	engineered mutation	UNP D5SK09
F	0	ASP	-	expression tag	UNP D5SK09
F	75	LEU	MET	engineered mutation	UNP D5SK09

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	3	Total Mg 3 3	0	0
2	D	3	Total Mg 3 3	0	0
2	E	3	Total Mg 3 3	0	0
2	F	3	Total Mg 3 3	0	0

- Molecule 3 is ALENDRONATE (three-letter code: 212) (formula: $C_4H_{13}NO_7P_2$).

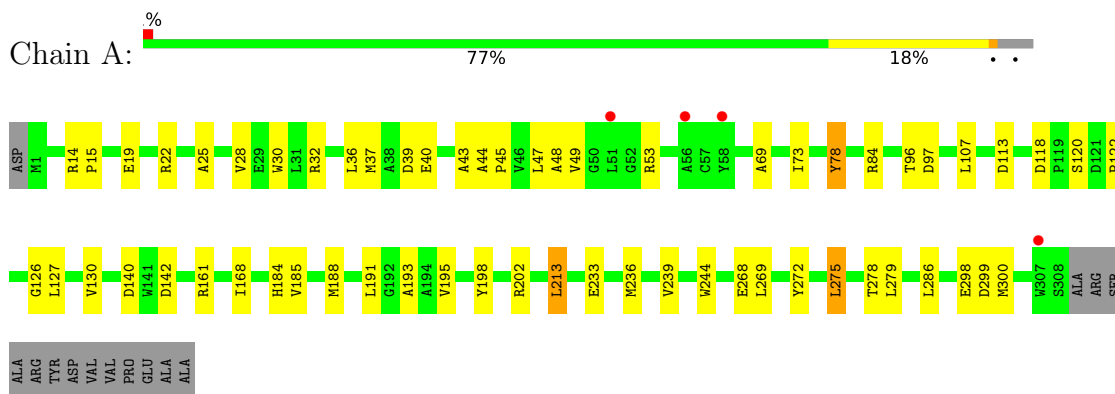


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total C N O P 14 4 1 7 2	0	0
3	D	1	Total C N O P 14 4 1 7 2	0	0
3	E	1	Total C N O P 14 4 1 7 2	0	0
3	F	1	Total C N O P 14 4 1 7 2	0	0

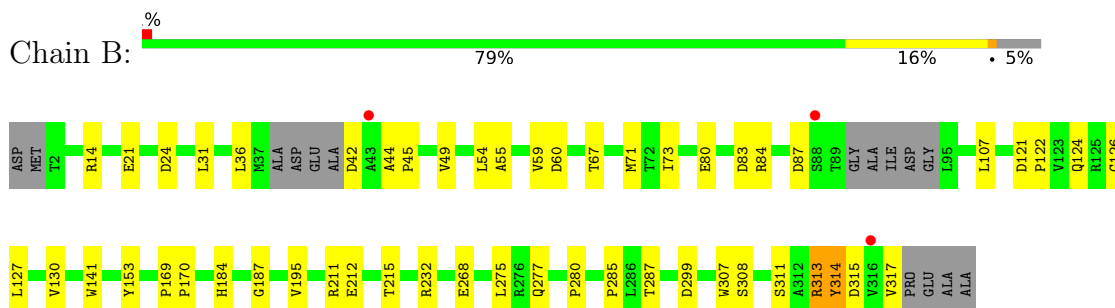
3 Residue-property plots [i](#)

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.

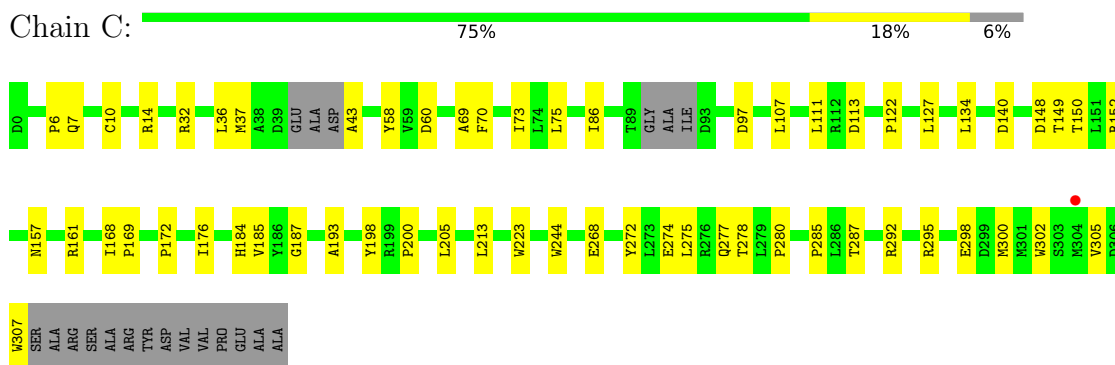
- Molecule 1: Terpene synthase




- Molecule 1: Terpene synthase

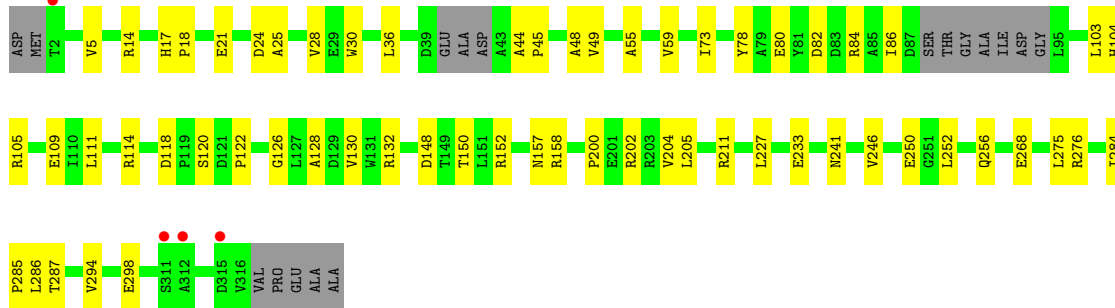


- Molecule 1: Terpene synthase



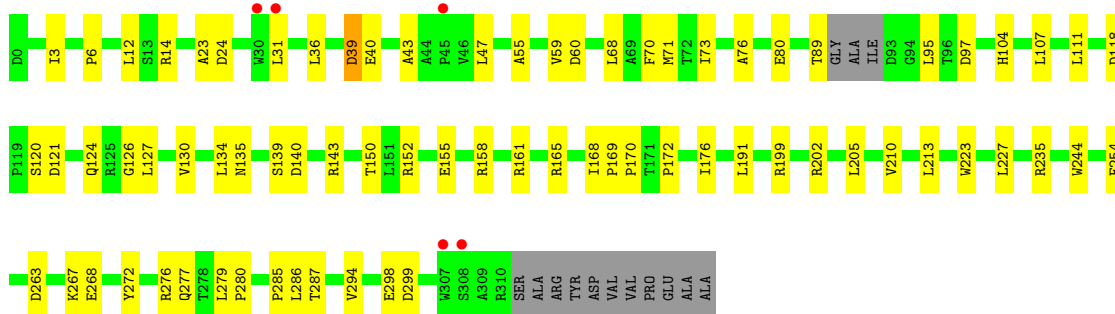
- Molecule 1: Terpene synthase

Chain D:  76% 19% 5%




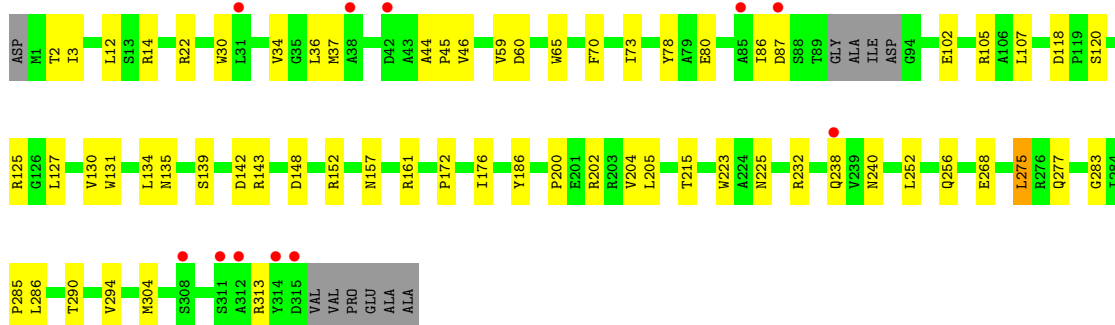
• Molecule 1: Terpene synthase

Chain E:  72% 23% 5%



• Molecule 1: Terpene synthase

Chain F:  77% 20% 3%



4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	106.72Å 82.32Å 148.29Å 90.00° 92.31° 90.00°	Depositor
Resolution (Å)	49.54 – 3.04 49.54 – 3.04	Depositor EDS
% Data completeness (in resolution range)	96.8 (49.54-3.04) 96.9 (49.54-3.04)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.16 (at 3.07Å)	Xtrriage
Refinement program	PHENIX 1.14_3260	Depositor
R, R_{free}	0.252 , 0.307 0.252 , 0.307	Depositor DCC
R_{free} test set	2101 reflections (4.35%)	wwPDB-VP
Wilson B-factor (Å ²)	64.7	Xtrriage
Anisotropy	0.165	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 54.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	0.031 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	14633	wwPDB-VP
Average B, all atoms (Å ²)	58.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.98% 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: 212, 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	0.26	0/2486	0.45	0/3390
1	B	0.27	0/2493	0.46	0/3397
1	C	0.28	0/2447	0.45	0/3334
1	D	0.27	0/2478	0.46	0/3376
1	E	0.27	0/2504	0.49	0/3412
1	F	0.27	0/2519	0.48	0/3432
All	All	0.27	0/14927	0.46	0/20341

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	2425	0	2331	35	0
1	B	2433	0	2342	30	0
1	C	2388	0	2294	36	0
1	D	2418	0	2326	39	0
1	E	2443	0	2340	48	0
1	F	2458	0	2360	38	0
2	B	3	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	3	0	0	0	0
2	E	3	0	0	0	0
2	F	3	0	0	0	0
3	B	14	0	9	0	0
3	D	14	0	9	1	0
3	E	14	0	9	0	0
3	F	14	0	9	0	0
All	All	14633	0	14029	222	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 8.

The worst 5 of 222 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:14:ARG:NH2	1:B:59:VAL:O	2.07	0.86
1:E:140:ASP:O	1:E:199:ARG:NH1	2.21	0.73
1:E:14:ARG:NH2	1:E:59:VAL:O	2.23	0.68
1:C:223:TRP:CZ3	1:C:268:GLU:HG3	2.29	0.67
1:D:36:LEU:HD13	1:D:80:GLU:HG2	1.76	0.67

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	306/322 (95%)	292 (95%)	11 (4%)	3 (1%)	15 49
1	B	301/322 (94%)	288 (96%)	13 (4%)	0	100 100
1	C	296/322 (92%)	287 (97%)	8 (3%)	1 (0%)	41 74
1	D	299/322 (93%)	291 (97%)	8 (3%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	305/322 (95%)	294 (96%)	11 (4%)	0	100	100
1	F	307/322 (95%)	295 (96%)	11 (4%)	1 (0%)	41	74
All	All	1814/1932 (94%)	1747 (96%)	62 (3%)	5 (0%)	41	74

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	39	ASP
1	A	140	ASP
1	F	139	SER
1	C	113	ASP
1	A	113	ASP

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	246/256 (96%)	241 (98%)	5 (2%)	55	81
1	B	248/256 (97%)	241 (97%)	7 (3%)	43	75
1	C	243/256 (95%)	238 (98%)	5 (2%)	53	80
1	D	245/256 (96%)	241 (98%)	4 (2%)	62	85
1	E	248/256 (97%)	243 (98%)	5 (2%)	55	81
1	F	249/256 (97%)	245 (98%)	4 (2%)	62	85
All	All	1479/1536 (96%)	1449 (98%)	30 (2%)	55	81

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	292	ARG
1	F	78	TYR
1	D	82	ASP
1	F	277	GLN
1	E	235	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	238	GLN
1	C	225	ASN
1	E	104	HIS
1	E	157	ASN

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 16 ligands modelled in this entry, 12 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	212	E	404	2	13,13,13	1.92	3 (23%)	19,21,21	1.48	3 (15%)
3	212	D	404	2	13,13,13	1.55	2 (15%)	19,21,21	1.51	2 (10%)
3	212	B	404	2	13,13,13	1.64	3 (23%)	19,21,21	1.51	2 (10%)
3	212	F	404	2	13,13,13	1.80	3 (23%)	19,21,21	1.43	2 (10%)

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	212	E	404	2	-	3/23/23/23	-
3	212	D	404	2	-	3/23/23/23	-
3	212	B	404	2	-	3/23/23/23	-
3	212	F	404	2	-	3/23/23/23	-

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	404	212	P1-C7	3.75	1.87	1.85
3	E	404	212	P8-C7	3.28	1.87	1.85
3	B	404	212	O14-C7	-3.28	1.40	1.44
3	D	404	212	O14-C7	-3.25	1.40	1.44
3	F	404	212	P1-C7	3.24	1.87	1.85

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	404	212	C19-C16-C7	-4.01	108.72	116.07
3	B	404	212	C19-C16-C7	-3.86	108.98	116.07
3	E	404	212	C19-C16-C7	-3.45	109.74	116.07
3	F	404	212	C19-C16-C7	-2.73	111.06	116.07
3	E	404	212	O2-P1-O5	-2.44	107.57	113.06

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	404	212	C19-C16-C7-P1
3	B	404	212	C19-C16-C7-P8
3	B	404	212	C19-C16-C7-O14
3	D	404	212	C19-C16-C7-P1
3	D	404	212	C19-C16-C7-P8

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	404	212	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	308/322 (95%)	-0.03	4 (1%) 77 51	39, 61, 90, 118	0
1	B	307/322 (95%)	-0.03	3 (0%) 82 59	33, 60, 92, 127	0
1	C	302/322 (93%)	-0.15	1 (0%) 94 83	33, 50, 75, 100	0
1	D	305/322 (94%)	0.05	4 (1%) 77 51	29, 57, 84, 104	0
1	E	308/322 (95%)	0.04	5 (1%) 72 44	32, 52, 84, 115	0
1	F	311/322 (96%)	0.07	11 (3%) 44 19	29, 53, 90, 120	0
All	All	1841/1932 (95%)	-0.01	28 (1%) 73 46	29, 55, 87, 127	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	315	ASP	4.1
1	F	312	ALA	4.1
1	F	314	TYR	3.3
1	F	308	SER	3.2
1	E	308	SER	3.0

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(Å ²)	Q<0.9
2	MG	F	403	1/1	0.57	0.20	81,81,81,81	0
2	MG	E	403	1/1	0.73	0.11	65,65,65,65	0
2	MG	B	403	1/1	0.77	0.17	66,66,66,66	0
3	212	F	404	14/14	0.86	0.20	58,79,92,107	0
2	MG	B	402	1/1	0.89	0.10	73,73,73,73	0
3	212	E	404	14/14	0.90	0.19	72,85,98,100	0
2	MG	D	401	1/1	0.90	0.07	79,79,79,79	0
2	MG	D	402	1/1	0.91	0.13	54,54,54,54	0
2	MG	F	401	1/1	0.91	0.11	66,66,66,66	0
2	MG	F	402	1/1	0.91	0.11	71,71,71,71	0
2	MG	B	401	1/1	0.92	0.14	74,74,74,74	0
2	MG	E	401	1/1	0.92	0.09	86,86,86,86	0
2	MG	D	403	1/1	0.93	0.15	68,68,68,68	0
3	212	B	404	14/14	0.94	0.14	55,72,92,96	0
3	212	D	404	14/14	0.94	0.14	65,78,85,85	0
2	MG	E	402	1/1	0.96	0.18	64,64,64,64	0

6.5 Other polymers [i](#)

There are no such residues in this entry.