



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 3, 2021 – 04:13 AM EDT

PDB ID : 3L7J
Title : Structure of the Wall Teichoic Acid Polymerase TagF, H444N variant
Authors : Strynadka, N.C.J.; Lovering, A.L.
Deposited on : 2009-12-28
Resolution : 2.81 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 : 2.23.2
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) : 2.23.2

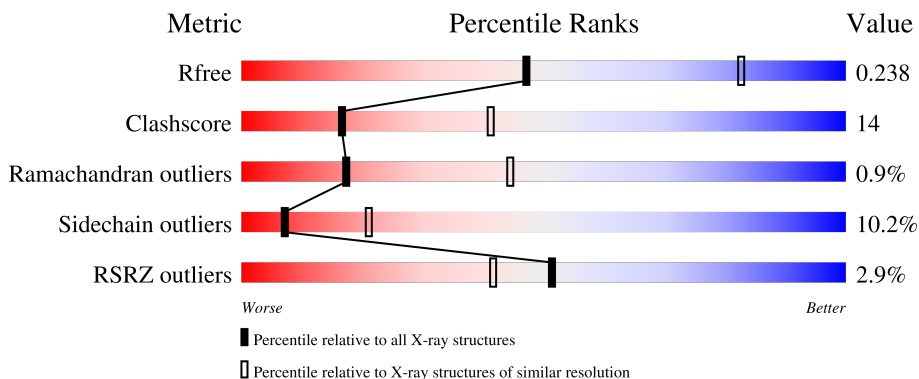
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.81 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	3617 (2.84-2.80)
Clashscore	141614	4060 (2.84-2.80)
Ramachandran outliers	138981	3978 (2.84-2.80)
Sidechain outliers	138945	3980 (2.84-2.80)
RSRZ outliers	127900	3552 (2.84-2.80)

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	729	<div> <div>2%</div> <div>39%15%•44%</div> </div>
1	B	729	<div> <div>2%</div> <div>35%18%•43%</div> </div>
1	C	729	<div> <div>2%</div> <div>38%16%•44%</div> </div>
1	D	729	<div> <div>2%</div> <div>37%16%•44%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	SO4	B	730	-	-	X	-
2	SO4	C	732	-	-	-	X
3	CL	B	732	-	-	X	-
3	CL	B	735	-	-	-	X
3	CL	B	736	-	-	X	-
3	CL	C	733	-	-	X	-
3	CL	D	731	-	-	X	-

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 13970 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 Teichoic acid biosynthesis protein F.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	410	Total	C	N	O	S	0	0	0
			3462	2224	577	650	11			
1	B	412	Total	C	N	O	S	0	0	0
			3477	2233	581	652	11			
1	C	411	Total	C	N	O	S	0	0	0
			3467	2227	578	651	11			
1	D	411	Total	C	N	O	S	0	0	0
			3472	2230	580	651	11			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	444	ASN	HIS	engineered mutation	UNP Q5HLM5
A	722	LEU	-	expression tag	UNP Q5HLM5
A	723	GLU	-	expression tag	UNP Q5HLM5
A	724	HIS	-	expression tag	UNP Q5HLM5
A	725	HIS	-	expression tag	UNP Q5HLM5
A	726	HIS	-	expression tag	UNP Q5HLM5
A	727	HIS	-	expression tag	UNP Q5HLM5
A	728	HIS	-	expression tag	UNP Q5HLM5
A	729	HIS	-	expression tag	UNP Q5HLM5
B	444	ASN	HIS	engineered mutation	UNP Q5HLM5
B	722	LEU	-	expression tag	UNP Q5HLM5
B	723	GLU	-	expression tag	UNP Q5HLM5
B	724	HIS	-	expression tag	UNP Q5HLM5
B	725	HIS	-	expression tag	UNP Q5HLM5
B	726	HIS	-	expression tag	UNP Q5HLM5
B	727	HIS	-	expression tag	UNP Q5HLM5
B	728	HIS	-	expression tag	UNP Q5HLM5
B	729	HIS	-	expression tag	UNP Q5HLM5
C	444	ASN	HIS	engineered mutation	UNP Q5HLM5
C	722	LEU	-	expression tag	UNP Q5HLM5
C	723	GLU	-	expression tag	UNP Q5HLM5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	724	HIS	-	expression tag	UNP Q5HLM5
C	725	HIS	-	expression tag	UNP Q5HLM5
C	726	HIS	-	expression tag	UNP Q5HLM5
C	727	HIS	-	expression tag	UNP Q5HLM5
C	728	HIS	-	expression tag	UNP Q5HLM5
C	729	HIS	-	expression tag	UNP Q5HLM5
D	444	ASN	HIS	engineered mutation	UNP Q5HLM5
D	722	LEU	-	expression tag	UNP Q5HLM5
D	723	GLU	-	expression tag	UNP Q5HLM5
D	724	HIS	-	expression tag	UNP Q5HLM5
D	725	HIS	-	expression tag	UNP Q5HLM5
D	726	HIS	-	expression tag	UNP Q5HLM5
D	727	HIS	-	expression tag	UNP Q5HLM5
D	728	HIS	-	expression tag	UNP Q5HLM5
D	729	HIS	-	expression tag	UNP Q5HLM5

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	C	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		

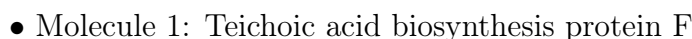
- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

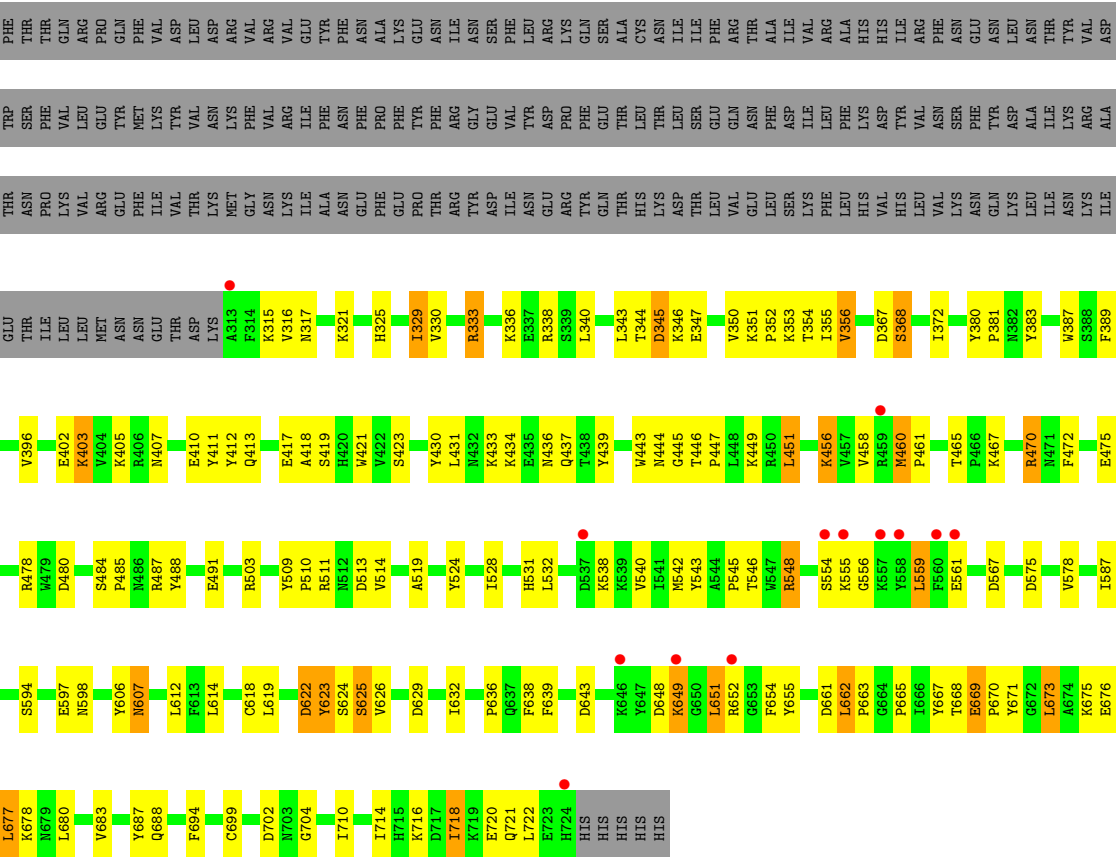
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	11	Total	Cl	0	0
			11	11		
3	B	8	Total	Cl	0	0
			8	8		
3	C	7	Total	Cl	0	0
			7	7		
3	D	6	Total	Cl	0	0
			6	6		

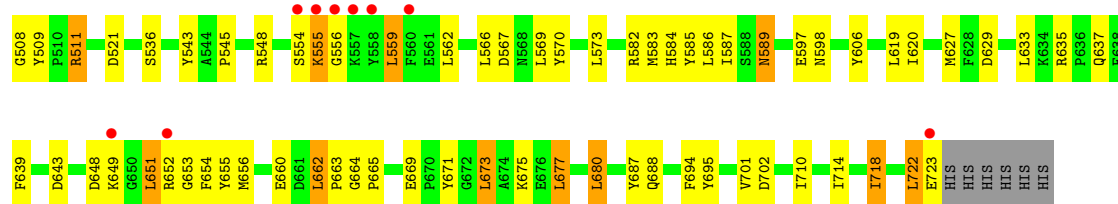
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	8	Total	O	0	0
			8	8		
4	B	5	Total	O	0	0
			5	5		
4	C	3	Total	O	0	0
			3	3		
4	D	4	Total	O	0	0
			4	4		

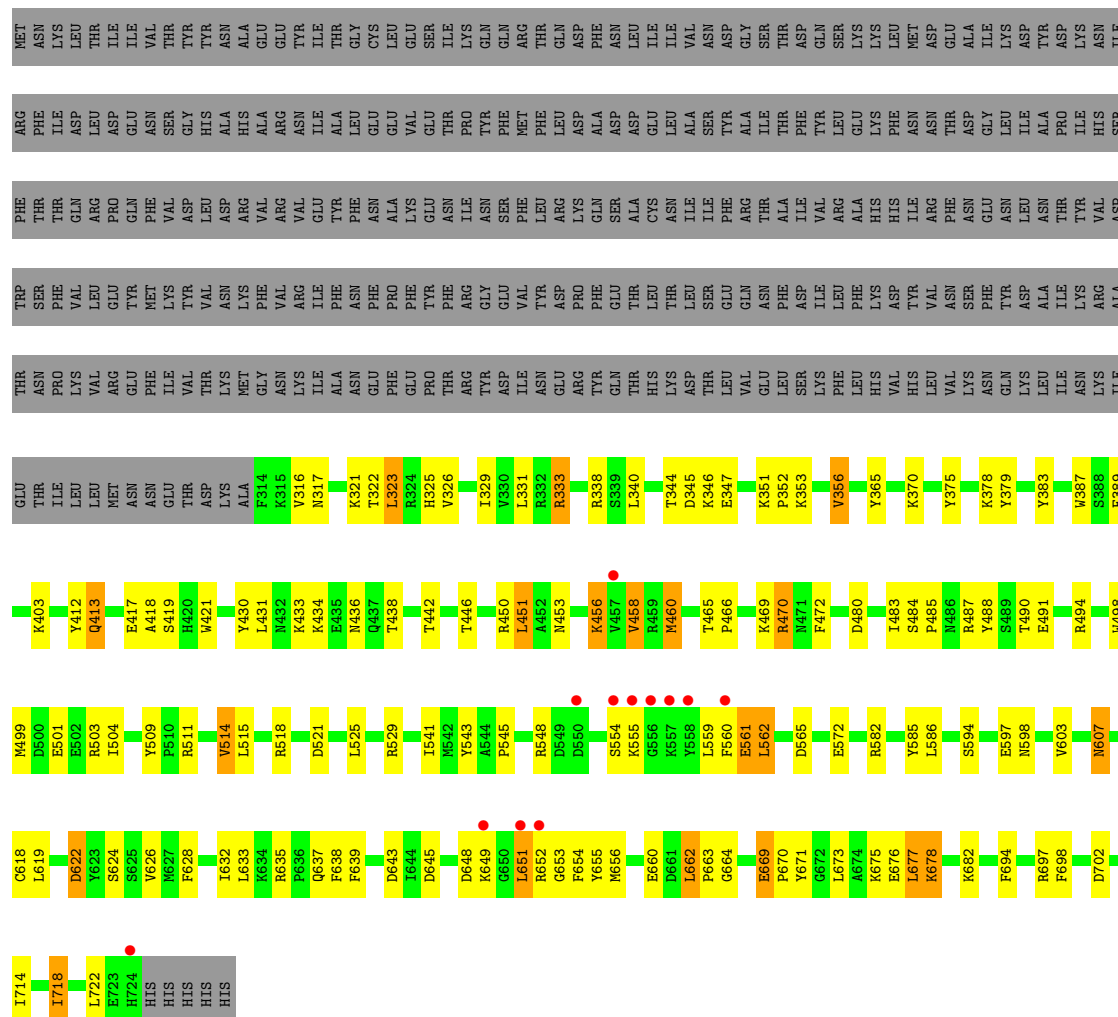
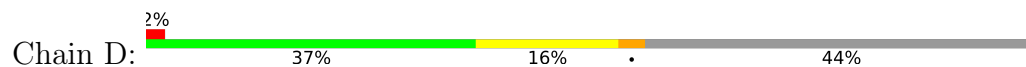
- Molecule 1: Teichoic acid biosynthesis protein F







● Molecule 1: Teichoic acid biosynthesis protein F



4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α , β , γ	223.60Å 223.60Å 100.57Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.80 – 2.81 19.80 – 2.81	Depositor EDS
% Data completeness (in resolution range)	97.9 (19.80-2.81) 97.9 (19.80-2.81)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.64 (at 2.79Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, R_{free}	0.197 , 0.246 0.191 , 0.238	Depositor DCC
R_{free} test set	3046 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	72.4	Xtriage
Anisotropy	0.048	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 29.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	13970	wwPDB-VP
Average B, all atoms (Å ²)	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.40% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, CL

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	0.48	0/3551	0.62	0/4798
1	B	0.49	0/3567	0.65	0/4820
1	C	0.46	0/3556	0.61	0/4805
1	D	0.46	0/3562	0.61	0/4813
All	All	0.47	0/14236	0.62	0/19236

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 [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	3462	0	3369	86	0
1	B	3477	0	3381	105	0
1	C	3467	0	3374	104	0
1	D	3472	0	3376	96	0
2	A	15	0	0	0	0
2	B	5	0	0	2	0
2	C	15	0	0	0	0
2	D	5	0	0	0	0
3	A	11	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	8	0	0	6	0
3	C	7	0	0	2	0
3	D	6	0	0	4	0
4	A	8	0	0	1	0
4	B	5	0	0	0	0
4	C	3	0	0	0	0
4	D	4	0	0	0	0
All	All	13970	0	13500	389	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 14.

The worst 5 of 389 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:A:333:ARG:HG2	1:A:333:ARG:HH11	1.02	1.19
1:D:470:ARG:HG2	1:D:470:ARG:HH11	1.04	1.10
1:C:333:ARG:HG2	1:C:333:ARG:HH11	1.28	0.99
1:B:333:ARG:HH11	1:B:333:ARG:HG2	1.28	0.98
1:C:470:ARG:HH11	1:C:470:ARG:HG2	1.27	0.97

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	408/729 (56%)	384 (94%)	21 (5%)	3 (1%)	22 51
1	B	410/729 (56%)	378 (92%)	30 (7%)	2 (0%)	29 59
1	C	409/729 (56%)	370 (90%)	36 (9%)	3 (1%)	22 51
1	D	409/729 (56%)	376 (92%)	26 (6%)	7 (2%)	9 27

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1636/2916 (56%)	1508 (92%)	113 (7%)	15 (1%)	17	44

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	647	TYR
1	D	561	GLU
1	A	535	PRO
1	B	345	ASP
1	B	649	LYS

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	379/675 (56%)	341 (90%)	38 (10%)	7	22
1	B	380/675 (56%)	334 (88%)	46 (12%)	5	14
1	C	379/675 (56%)	343 (90%)	36 (10%)	8	24
1	D	380/675 (56%)	345 (91%)	35 (9%)	9	26
All	All	1518/2700 (56%)	1363 (90%)	155 (10%)	7	21

5 of 155 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	677	LEU
1	D	645	ASP
1	C	722	LEU
1	D	460	MET
1	D	677	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	436	ASN
1	C	531	HIS
1	D	686	GLN
1	D	413	GLN
1	D	531	HIS

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 40 ligands modelled in this entry, 32 are monoatomic - leaving 8 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	SO4	C	732	-	4,4,4	0.13	0	6,6,6	0.10	0
2	SO4	A	731	-	4,4,4	0.08	0	6,6,6	0.35	0
2	SO4	B	730	-	4,4,4	0.11	0	6,6,6	0.15	0
2	SO4	A	730	-	4,4,4	0.17	0	6,6,6	0.18	0
2	SO4	C	730	-	4,4,4	0.15	0	6,6,6	0.23	0
2	SO4	C	731	-	4,4,4	0.12	0	6,6,6	0.14	0
2	SO4	D	730	-	4,4,4	0.14	0	6,6,6	0.10	0
2	SO4	A	732	-	4,4,4	0.17	0	6,6,6	0.06	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	730	SO4	2	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	410/729 (56%)	-0.28	8 (1%) 65 56	50, 73, 143, 221	0
1	B	412/729 (56%)	-0.24	13 (3%) 47 37	51, 72, 158, 239	0
1	C	411/729 (56%)	-0.19	14 (3%) 45 35	56, 79, 158, 237	0
1	D	411/729 (56%)	-0.18	12 (2%) 51 41	57, 78, 159, 221	0
All	All	1644/2916 (56%)	-0.22	47 (2%) 51 41	50, 76, 154, 239	0

The worst 5 of 47 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	558	TYR	5.6
1	D	556	GLY	5.4
1	B	652	ARG	5.3
1	D	724	HIS	5.2
1	D	652	ARG	5.2

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	SO4	C	732	5/5	0.61	0.47	177,177,178,178	0
3	CL	B	735	1/1	0.77	0.47	103,103,103,103	0
3	CL	B	734	1/1	0.78	0.17	104,104,104,104	0
2	SO4	D	730	5/5	0.78	0.29	188,188,189,190	0
3	CL	A	743	1/1	0.80	0.16	99,99,99,99	0
3	CL	A	733	1/1	0.81	0.41	97,97,97,97	0
3	CL	B	732	1/1	0.81	0.31	99,99,99,99	0
2	SO4	B	730	5/5	0.84	0.26	174,174,176,176	0
3	CL	D	731	1/1	0.84	0.08	90,90,90,90	0
2	SO4	A	732	5/5	0.85	0.42	155,155,156,157	0
3	CL	D	733	1/1	0.86	0.30	96,96,96,96	0
3	CL	D	735	1/1	0.86	0.13	103,103,103,103	0
3	CL	A	734	1/1	0.87	0.35	108,108,108,108	0
3	CL	C	737	1/1	0.88	0.29	92,92,92,92	0
3	CL	A	740	1/1	0.91	0.11	95,95,95,95	0
3	CL	B	733	1/1	0.91	0.40	111,111,111,111	0
3	CL	D	734	1/1	0.92	0.14	93,93,93,93	0
3	CL	C	739	1/1	0.92	0.08	99,99,99,99	0
3	CL	D	736	1/1	0.92	0.08	93,93,93,93	0
3	CL	C	733	1/1	0.93	0.17	105,105,105,105	0
3	CL	A	741	1/1	0.93	0.18	91,91,91,91	0
3	CL	C	738	1/1	0.94	0.33	87,87,87,87	0
3	CL	D	732	1/1	0.94	0.27	86,86,86,86	0
3	CL	A	742	1/1	0.94	0.16	95,95,95,95	0
3	CL	A	735	1/1	0.95	0.64	98,98,98,98	0
2	SO4	C	731	5/5	0.95	0.12	118,119,121,121	0
3	CL	C	736	1/1	0.95	0.10	92,92,92,92	0
3	CL	A	736	1/1	0.96	0.13	82,82,82,82	0
3	CL	C	735	1/1	0.96	0.22	81,81,81,81	0
3	CL	A	737	1/1	0.96	0.49	93,93,93,93	0
3	CL	B	736	1/1	0.96	0.32	103,103,103,103	0
3	CL	C	734	1/1	0.97	0.21	93,93,93,93	0
2	SO4	C	730	5/5	0.97	0.11	113,114,115,116	0
3	CL	B	738	1/1	0.97	0.09	97,97,97,97	0
3	CL	A	738	1/1	0.97	0.31	98,98,98,98	0
3	CL	B	737	1/1	0.98	0.10	86,86,86,86	0
3	CL	A	739	1/1	0.98	0.19	68,68,68,68	0
2	SO4	A	730	5/5	0.98	0.12	99,100,102,103	0
3	CL	B	731	1/1	0.98	0.08	50,50,50,50	0
2	SO4	A	731	5/5	0.98	0.14	87,87,88,91	0

6.5 Other polymers [i](#)

There are no such residues in this entry.