Full wwPDB X-ray Structure Validation Report (i)

Nov 7, 2022 – 05:00 PM JST

PDB ID	:	7VUL
Title	:	Structure of Klebsiella Phage P560 depolymerase
Authors	:	Li, M.; Chen, R.; Zhang, S.; Zhang, W.
Deposited on	:	2021-11-02
Resolution	:	1.59 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (i) 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 (1)) were used in the production of this report:

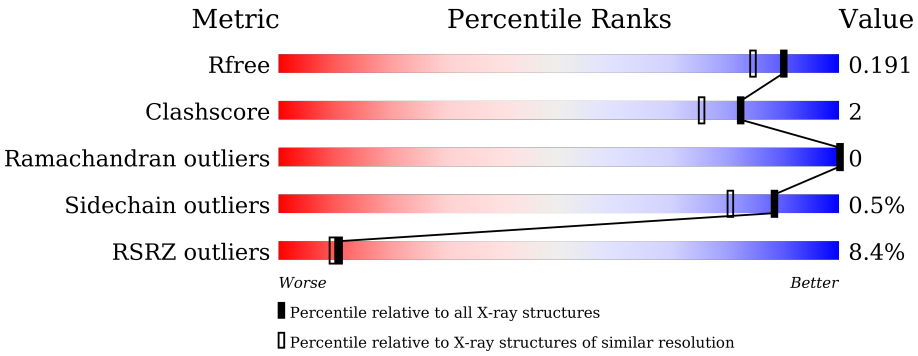
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.59 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	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)
RSRZ outliers	127900	3321 (1.60-1.60)

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 <=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	663	7% 91% · 5%
1	B	663	8% 87% 7% 6%
1	C	663	10% 89% 6% 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 15195 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 Non-contractile tail fiber protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	627	Total	C	N	O	S	0	0	0
			4674	2943	804	910	17			
1	B	622	Total	C	N	O	S	0	0	0
			4638	2920	798	903	17			
1	C	627	Total	C	N	O	S	0	0	0
			4674	2943	804	910	17			

There are 66 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP A0A7S6TZU4
A	-18	GLY	-	expression tag	UNP A0A7S6TZU4
A	-17	SER	-	expression tag	UNP A0A7S6TZU4
A	-16	SER	-	expression tag	UNP A0A7S6TZU4
A	-15	HIS	-	expression tag	UNP A0A7S6TZU4
A	-14	HIS	-	expression tag	UNP A0A7S6TZU4
A	-13	HIS	-	expression tag	UNP A0A7S6TZU4
A	-12	HIS	-	expression tag	UNP A0A7S6TZU4
A	-11	HIS	-	expression tag	UNP A0A7S6TZU4
A	-10	HIS	-	expression tag	UNP A0A7S6TZU4
A	-9	SER	-	expression tag	UNP A0A7S6TZU4
A	-8	SER	-	expression tag	UNP A0A7S6TZU4
A	-7	GLY	-	expression tag	UNP A0A7S6TZU4
A	-6	LEU	-	expression tag	UNP A0A7S6TZU4
A	-5	VAL	-	expression tag	UNP A0A7S6TZU4
A	-4	PRO	-	expression tag	UNP A0A7S6TZU4
A	-3	ARG	-	expression tag	UNP A0A7S6TZU4
A	-2	GLY	-	expression tag	UNP A0A7S6TZU4
A	-1	SER	-	expression tag	UNP A0A7S6TZU4
A	0	HIS	-	expression tag	UNP A0A7S6TZU4
A	642	LEU	-	expression tag	UNP A0A7S6TZU4
A	643	GLU	-	expression tag	UNP A0A7S6TZU4
B	-19	MET	-	initiating methionine	UNP A0A7S6TZU4

Continued on next page...

Continued from previous page...

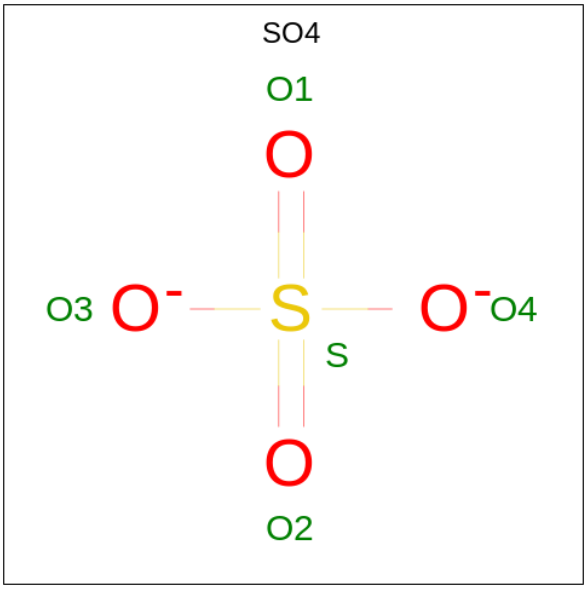
Chain	Residue	Modelled	Actual	Comment	Reference
B	-18	GLY	-	expression tag	UNP A0A7S6TZU4
B	-17	SER	-	expression tag	UNP A0A7S6TZU4
B	-16	SER	-	expression tag	UNP A0A7S6TZU4
B	-15	HIS	-	expression tag	UNP A0A7S6TZU4
B	-14	HIS	-	expression tag	UNP A0A7S6TZU4
B	-13	HIS	-	expression tag	UNP A0A7S6TZU4
B	-12	HIS	-	expression tag	UNP A0A7S6TZU4
B	-11	HIS	-	expression tag	UNP A0A7S6TZU4
B	-10	HIS	-	expression tag	UNP A0A7S6TZU4
B	-9	SER	-	expression tag	UNP A0A7S6TZU4
B	-8	SER	-	expression tag	UNP A0A7S6TZU4
B	-7	GLY	-	expression tag	UNP A0A7S6TZU4
B	-6	LEU	-	expression tag	UNP A0A7S6TZU4
B	-5	VAL	-	expression tag	UNP A0A7S6TZU4
B	-4	PRO	-	expression tag	UNP A0A7S6TZU4
B	-3	ARG	-	expression tag	UNP A0A7S6TZU4
B	-2	GLY	-	expression tag	UNP A0A7S6TZU4
B	-1	SER	-	expression tag	UNP A0A7S6TZU4
B	0	HIS	-	expression tag	UNP A0A7S6TZU4
B	642	LEU	-	expression tag	UNP A0A7S6TZU4
B	643	GLU	-	expression tag	UNP A0A7S6TZU4
C	-19	MET	-	initiating methionine	UNP A0A7S6TZU4
C	-18	GLY	-	expression tag	UNP A0A7S6TZU4
C	-17	SER	-	expression tag	UNP A0A7S6TZU4
C	-16	SER	-	expression tag	UNP A0A7S6TZU4
C	-15	HIS	-	expression tag	UNP A0A7S6TZU4
C	-14	HIS	-	expression tag	UNP A0A7S6TZU4
C	-13	HIS	-	expression tag	UNP A0A7S6TZU4
C	-12	HIS	-	expression tag	UNP A0A7S6TZU4
C	-11	HIS	-	expression tag	UNP A0A7S6TZU4
C	-10	HIS	-	expression tag	UNP A0A7S6TZU4
C	-9	SER	-	expression tag	UNP A0A7S6TZU4
C	-8	SER	-	expression tag	UNP A0A7S6TZU4
C	-7	GLY	-	expression tag	UNP A0A7S6TZU4
C	-6	LEU	-	expression tag	UNP A0A7S6TZU4
C	-5	VAL	-	expression tag	UNP A0A7S6TZU4
C	-4	PRO	-	expression tag	UNP A0A7S6TZU4
C	-3	ARG	-	expression tag	UNP A0A7S6TZU4
C	-2	GLY	-	expression tag	UNP A0A7S6TZU4
C	-1	SER	-	expression tag	UNP A0A7S6TZU4
C	0	HIS	-	expression tag	UNP A0A7S6TZU4
C	642	LEU	-	expression tag	UNP A0A7S6TZU4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	643	GLU	-	expression tag	UNP A0A7S6TZU4

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	409	Total O 409 409	0	0
3	B	394	Total O 394 394	0	0

Continued on next page...

Continued from previous page...

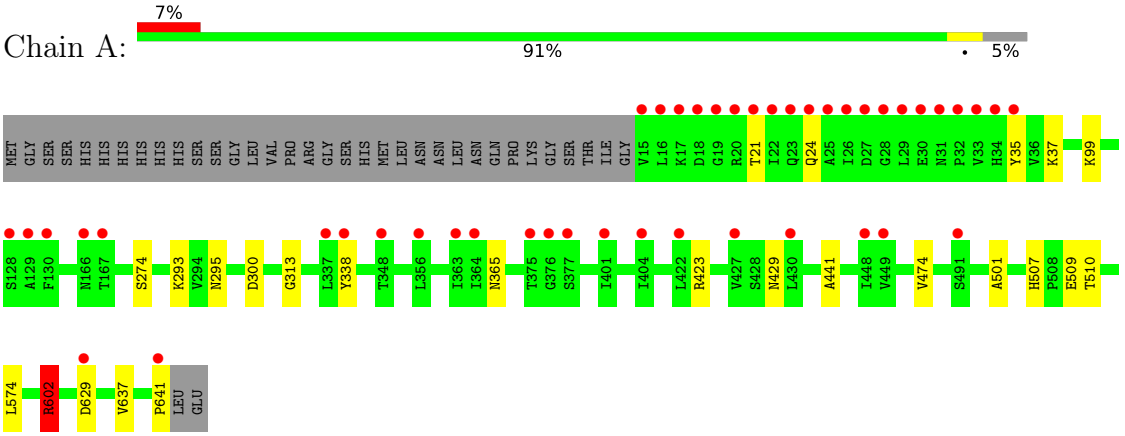
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	371	Total	O	0	0
			371	371		

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.

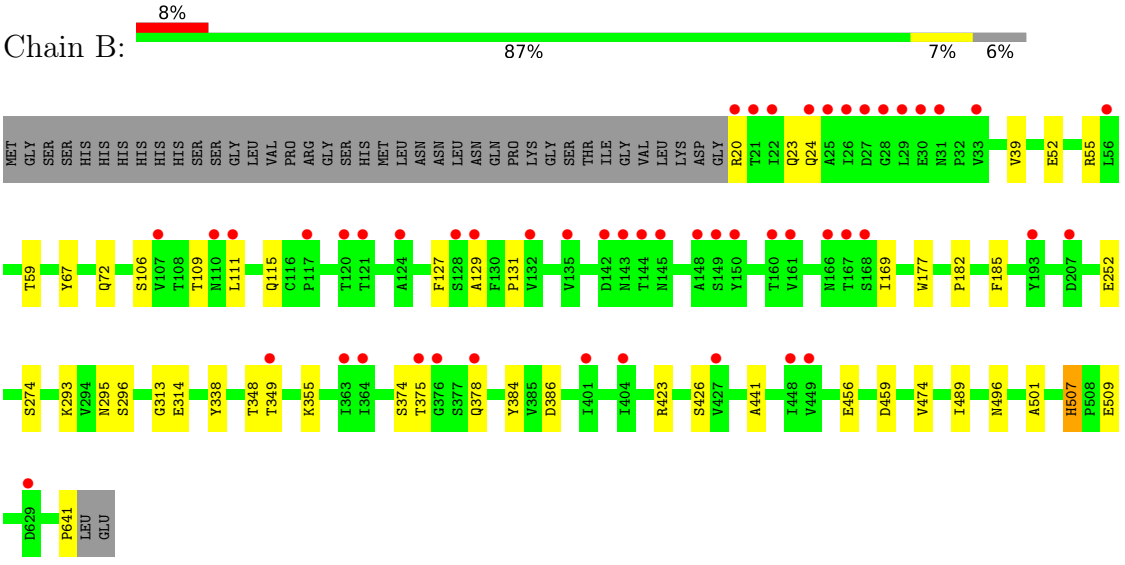
• Molecule 1: Non-contractile tail fiber protein

Chain A:



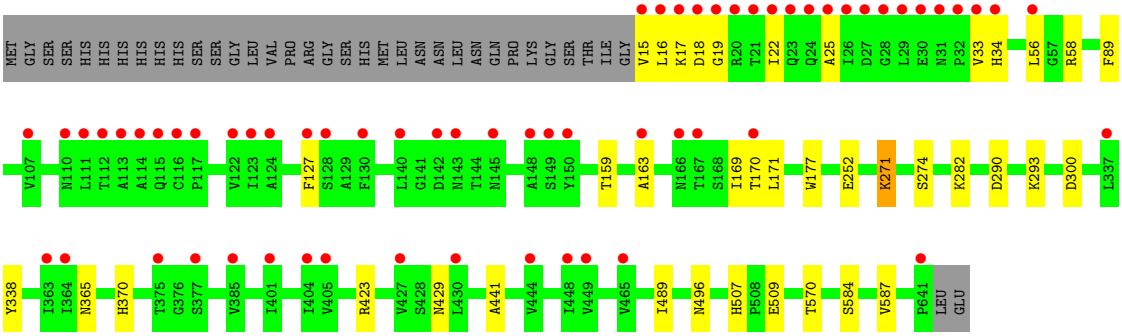
• Molecule 1: Non-contractile tail fiber protein

Chain B:



• Molecule 1: Non-contractile tail fiber protein

Chain C:


MET GLY SER SER HIS HIS HIS HIS HIS HIS SER SER GLY LEU VAL PRO ARG GLY SER HIS MET LEU ASN ASN LEU ASN GLN PRO LYS GLY SER THR ILE GLY
V15 L16 K17 D18 G19 R20 T21 I22 Q23 Q24 A25 I26 D27 G28 L29 E30 N31 P32 V33 H34 L56 G57 R58 F89
V107 N110 L111 T112 A113 A114 Q115 C116 P117 V122 I123 A124 F127 S128 A129 F130 L140 G141 D142 N143 T144 N145 A148 S149 Y150 T159 A163 N166 T167 S168 I169 T170 L171 W177 E252 K271 S274 K282 D290 K293 D300 L337
Y338 I363 I364 N365 H370 T375 G376 S377 V386 I401 I404 V405 R423 V427 S428 N429 L430 A441 V444 I448 V449 V465 I489 N496 H507 P508 E509 T570 S584 V587 P641 LEU GLU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	100.67Å 126.84Å 144.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.55 – 1.59 47.73 – 1.59	Depositor EDS
% Data completeness (in resolution range)	99.6 (47.55-1.59) 99.6 (47.73-1.59)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.10 (at 1.58Å)	Xtriage
Refinement program	PHENIX 1.19_4092	Depositor
R, R_{free}	0.168 , 0.193 0.165 , 0.191	Depositor DCC
R_{free} test set	12644 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	20.9	Xtriage
Anisotropy	0.387	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 48.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	15195	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.20% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: SO4

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	0.83	1/4775 (0.0%)	0.88	2/6511 (0.0%)
1	B	0.81	1/4739 (0.0%)	0.86	3/6463 (0.0%)
1	C	0.77	2/4775 (0.0%)	0.83	1/6511 (0.0%)
All	All	0.80	4/14289 (0.0%)	0.86	6/19485 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	602	ARG	CB-CG	-6.46	1.35	1.52
1	C	252	GLU	CB-CG	-6.06	1.40	1.52
1	B	72	GLN	CB-CG	-5.19	1.38	1.52
1	C	271	LYS	CE-NZ	5.06	1.61	1.49

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	300	ASP	CB-CG-OD1	7.25	124.83	118.30
1	B	384	TYR	CB-CG-CD2	-5.91	117.45	121.00
1	C	89	PHE	CB-CG-CD1	5.56	124.69	120.80
1	B	507	HIS	N-CA-CB	5.37	120.26	110.60
1	B	386	ASP	CB-CG-OD2	-5.32	113.52	118.30
1	A	602	ARG	CG-CD-NE	-5.00	101.30	111.80

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	4674	0	4592	20	0
1	B	4638	0	4552	32	0
1	C	4674	0	4592	22	0
2	A	10	0	0	0	0
2	B	15	0	0	1	0
2	C	10	0	0	0	0
3	A	409	0	0	5	0
3	B	394	0	0	1	0
3	C	371	0	0	2	0
All	All	15195	0	13736	69	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 2.

All (69) 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:20:ARG:HE	1:B:23:GLN:HB2	1.35	0.92
1:A:629:ASP:OD2	3:A:801:HOH:O	1.91	0.86
1:B:109:THR:HG23	1:B:111:LEU:HG	1.68	0.76
1:B:52:GLU:OE2	1:B:55:ARG:NH2	2.19	0.75
1:B:109:THR:HB	1:B:182:PRO:O	1.90	0.71
1:B:59:THR:HG23	3:C:841:HOH:O	1.91	0.69
1:C:16:LEU:HD11	1:C:19:GLY:H	1.61	0.65
1:A:507:HIS:CD2	1:A:509:GLU:H	2.16	0.64
1:A:293:LYS:NZ	3:A:802:HOH:O	2.04	0.64
1:C:15:VAL:HG13	1:C:22:ILE:HD12	1.81	0.63
1:B:109:THR:CG2	1:B:111:LEU:HG	2.28	0.62
1:B:20:ARG:NE	1:B:23:GLN:HB2	2.11	0.60
1:C:17:LYS:HE2	1:C:25:ALA:HB2	1.82	0.60
1:B:20:ARG:O	1:B:24:GLN:HG3	2.02	0.59
1:B:20:ARG:NH2	1:B:23:GLN:OE1	2.37	0.58
1:C:16:LEU:HD13	1:C:17:LYS:N	2.18	0.58
1:B:20:ARG:HE	1:B:23:GLN:CB	2.13	0.58
1:A:507:HIS:HD2	1:A:510:THR:H	1.53	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:507:HIS:NE2	1:A:509:GLU:OE2	2.36	0.56
1:A:641:PRO:O	3:A:803:HOH:O	2.18	0.56
1:A:99:LYS:HD2	3:A:1146:HOH:O	2.06	0.54
1:B:109:THR:HG23	1:B:111:LEU:H	1.72	0.54
1:B:426:SER:OG	1:C:429:ASN:OD1	2.27	0.52
1:A:637:VAL:HG11	1:B:641:PRO:HG3	1.93	0.51
1:A:602:ARG:HB2	1:C:584:SER:OG	2.11	0.50
1:B:355:LYS:HE2	3:B:1030:HOH:O	2.12	0.50
1:B:295:ASN:ND2	1:B:296:SER:OG	2.46	0.49
1:A:293:LYS:NZ	1:B:314:GLU:OE1	2.45	0.49
1:B:507:HIS:CE1	1:B:509:GLU:HB2	2.48	0.48
1:C:56:LEU:HD13	1:C:58:ARG:HH11	1.78	0.48
1:B:129:ALA:O	1:B:131:PRO:HD3	2.14	0.47
1:B:348:THR:HG22	1:B:375:THR:OG1	2.14	0.47
1:A:21:THR:OG1	1:A:24:GLN:HG3	2.15	0.47
1:C:271:LYS:HE2	1:C:290:ASP:OD2	2.16	0.46
1:B:127:PHE:CG	1:B:169:ILE:HB	2.51	0.46
1:B:106:SER:HB3	1:B:185:PHE:HB3	1.96	0.45
1:A:507:HIS:HD2	1:A:509:GLU:H	1.63	0.45
1:C:489:ILE:HD12	1:C:496:ASN:HA	1.99	0.45
1:B:252:GLU:HG2	1:C:370:HIS:CE1	2.52	0.45
1:A:423:ARG:HA	1:A:441:ALA:O	2.17	0.44
1:C:274:SER:HA	1:C:293:LYS:O	2.17	0.44
1:C:338:TYR:HA	1:C:365:ASN:O	2.17	0.44
1:B:423:ARG:HA	1:B:441:ALA:O	2.17	0.44
1:C:33:VAL:HG23	1:C:34:HIS:CE1	2.53	0.43
1:B:313:GLY:HA2	1:B:338:TYR:O	2.18	0.43
1:C:159:THR:HG21	1:C:171:LEU:HB3	2.00	0.43
1:C:127:PHE:CG	1:C:169:ILE:HB	2.53	0.43
1:B:459:ASP:OD2	2:B:702:SO4:O1	2.36	0.43
1:A:429:ASN:OD1	3:A:804:HOH:O	2.22	0.42
1:C:423:ARG:HA	1:C:441:ALA:O	2.19	0.42
1:A:274:SER:HA	1:A:293:LYS:O	2.20	0.42
1:A:295:ASN:HA	1:A:313:GLY:O	2.19	0.42
1:C:16:LEU:HD13	1:C:17:LYS:H	1.85	0.42
1:C:34:HIS:CD2	3:C:994:HOH:O	2.71	0.42
1:C:163:ALA:HB3	1:C:170:THR:HB	2.01	0.42
1:A:338:TYR:HA	1:A:365:ASN:O	2.19	0.42
1:B:109:THR:HG21	1:B:182:PRO:HG2	2.02	0.42
1:B:111:LEU:HA	1:B:115:GLN:OE1	2.20	0.42
1:B:39:VAL:HA	1:B:67:TYR:CD2	2.54	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:507:HIS:CE1	1:C:509:GLU:HB2	2.55	0.41
1:C:282:LYS:O	1:C:300:ASP:HB2	2.20	0.41
1:C:570:THR:HA	1:C:587:VAL:O	2.21	0.41
1:B:489:ILE:HD12	1:B:496:ASN:HA	2.03	0.41
1:A:35:TYR:CE2	1:A:37:LYS:HB3	2.55	0.41
1:B:274:SER:HA	1:B:293:LYS:O	2.21	0.41
1:B:474:VAL:O	1:B:501:ALA:HA	2.21	0.40
1:A:507:HIS:CD2	1:A:510:THR:H	2.34	0.40
1:A:474:VAL:O	1:A:501:ALA:HA	2.21	0.40
1:B:349:THR:HG22	1:B:378:GLN:HE22	1.87	0.40

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	625/663 (94%)	605 (97%)	20 (3%)	0	100	100
1	B	620/663 (94%)	600 (97%)	20 (3%)	0	100	100
1	C	625/663 (94%)	604 (97%)	21 (3%)	0	100	100
All	All	1870/1989 (94%)	1809 (97%)	61 (3%)	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	A	514/545 (94%)	512 (100%)	2 (0%)	91 84
1	B	510/545 (94%)	507 (99%)	3 (1%)	86 77
1	C	514/545 (94%)	512 (100%)	2 (0%)	91 84
All	All	1538/1635 (94%)	1531 (100%)	7 (0%)	88 80

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	574	LEU
1	A	602	ARG
1	B	177	TRP
1	B	374	SER
1	B	456	GLU
1	C	18	ASP
1	C	177	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	507	HIS
1	B	31	ASN
1	B	145	ASN
1	B	295	ASN
1	B	366	ASN
1	B	378	GLN
1	C	34	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)

7 ligands are modelled in this entry.

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	A	701	-	4,4,4	0.27	0	6,6,6	0.63	0
2	SO4	C	702	-	4,4,4	0.23	0	6,6,6	0.71	0
2	SO4	B	703	-	4,4,4	0.40	0	6,6,6	0.77	0
2	SO4	C	701	-	4,4,4	0.32	0	6,6,6	0.37	0
2	SO4	B	701	-	4,4,4	0.21	0	6,6,6	0.46	0
2	SO4	A	702	-	4,4,4	0.27	0	6,6,6	0.39	0
2	SO4	B	702	-	4,4,4	0.38	0	6,6,6	0.52	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 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	702	SO4	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

6.1 Protein, DNA and RNA chains

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, 95^{th} 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	627/663 (94%)	0.39	45 (7%)	15	14	16, 21, 36, 59	0
1	B	622/663 (93%)	0.49	50 (8%)	12	11	15, 21, 41, 63	0
1	C	627/663 (94%)	0.53	63 (10%)	7	6	16, 22, 41, 71	0
All	All	1876/1989 (94%)	0.47	158 (8%)	11	9	15, 21, 40, 71	0

All (158) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	15	VAL	17.6
1	C	22	ILE	12.1
1	C	16	LEU	11.3
1	C	19	GLY	9.4
1	A	15	VAL	9.0
1	B	22	ILE	8.1
1	A	22	ILE	7.7
1	A	26	ILE	7.5
1	C	33	VAL	7.2
1	C	18	ASP	7.1
1	C	21	THR	6.7
1	A	16	LEU	6.6
1	A	23	GLN	6.3
1	A	21	THR	6.1
1	A	25	ALA	5.8
1	C	29	LEU	5.7
1	C	56	LEU	5.7
1	A	27	ASP	5.6
1	C	20	ARG	5.5
1	B	21	THR	5.5
1	C	27	ASP	5.5
1	A	31	ASN	5.5
1	C	23	GLN	5.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	33	VAL	5.4
1	B	375	THR	5.2
1	A	375	THR	5.1
1	B	20	ARG	5.1
1	B	33	VAL	5.0
1	A	19	GLY	5.0
1	C	24	GLN	4.9
1	C	26	ILE	4.7
1	B	25	ALA	4.7
1	A	28	GLY	4.7
1	A	24	GLN	4.5
1	B	26	ILE	4.4
1	B	167	THR	4.3
1	B	27	ASP	4.2
1	C	128	SER	4.1
1	B	143	ASN	4.1
1	B	150	TYR	4.0
1	B	56	LEU	4.0
1	A	29	LEU	3.9
1	A	18	ASP	3.9
1	C	110	ASN	3.9
1	B	427	VAL	3.9
1	C	375	THR	3.9
1	B	148	ALA	3.8
1	A	30	GLU	3.8
1	B	145	ASN	3.8
1	B	144	THR	3.7
1	C	143	ASN	3.7
1	C	163	ALA	3.7
1	B	24	GLN	3.7
1	C	124	ALA	3.7
1	A	17	LYS	3.6
1	A	641	PRO	3.6
1	A	35	TYR	3.6
1	A	32	PRO	3.5
1	C	31	ASN	3.5
1	C	427	VAL	3.5
1	C	167	THR	3.4
1	B	129	ALA	3.4
1	B	166	ASN	3.4
1	C	145	ASN	3.4
1	B	29	LEU	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	107	VAL	3.4
1	B	142	ASP	3.3
1	B	31	ASN	3.3
1	C	17	LYS	3.3
1	C	111	LEU	3.3
1	C	641	PRO	3.2
1	C	142	ASP	3.2
1	B	168	SER	3.2
1	C	364	ILE	3.2
1	B	149	SER	3.2
1	C	32	PRO	3.2
1	A	401	ILE	3.1
1	C	107	VAL	3.1
1	C	34	HIS	3.1
1	C	112	THR	3.0
1	C	116	CYS	3.0
1	A	364	ILE	3.0
1	B	364	ILE	3.0
1	C	117	PRO	3.0
1	C	123	ILE	2.9
1	A	166	ASN	2.9
1	C	448	ILE	2.9
1	C	30	GLU	2.9
1	C	28	GLY	2.8
1	A	404	ILE	2.8
1	C	401	ILE	2.8
1	C	127	PHE	2.8
1	A	376	GLY	2.8
1	A	128	SER	2.8
1	C	404	ILE	2.7
1	A	20	ARG	2.7
1	A	363	ILE	2.7
1	B	448	ILE	2.7
1	B	128	SER	2.7
1	C	25	ALA	2.7
1	C	170	THR	2.6
1	B	401	ILE	2.6
1	A	629	ASP	2.6
1	C	150	TYR	2.6
1	C	166	ASN	2.6
1	A	129	ALA	2.6
1	C	149	SER	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	130	PHE	2.5
1	B	110	ASN	2.5
1	A	427	VAL	2.5
1	C	113	ALA	2.5
1	A	377	SER	2.4
1	A	491	SER	2.4
1	B	111	LEU	2.4
1	B	124	ALA	2.4
1	C	114	ALA	2.4
1	C	363	ILE	2.4
1	B	160	THR	2.4
1	C	115	GLN	2.4
1	A	34	HIS	2.4
1	A	337	LEU	2.4
1	A	430	LEU	2.4
1	C	337	LEU	2.4
1	C	465	VAL	2.3
1	A	167	THR	2.3
1	B	132	VAL	2.3
1	A	356	LEU	2.3
1	C	430	LEU	2.3
1	C	377	SER	2.3
1	C	444	VAL	2.3
1	C	449	VAL	2.3
1	B	404	ILE	2.3
1	A	449	VAL	2.3
1	B	28	GLY	2.2
1	A	348	THR	2.2
1	C	140	LEU	2.2
1	B	121	THR	2.2
1	B	135	VAL	2.2
1	B	349	THR	2.2
1	A	448	ILE	2.2
1	B	30	GLU	2.2
1	C	148	ALA	2.1
1	B	363	ILE	2.1
1	B	120	THR	2.1
1	B	117	PRO	2.1
1	A	422	LEU	2.1
1	B	207	ASP	2.1
1	C	130	PHE	2.1
1	A	338	TYR	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	629	ASP	2.1
1	B	449	VAL	2.1
1	B	378	GLN	2.0
1	B	193	TYR	2.0
1	C	122	VAL	2.0
1	B	376	GLY	2.0
1	B	161	VAL	2.0
1	C	385	VAL	2.0
1	C	405	VAL	2.0

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, 95^{th} 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	SO4	B	702	5/5	0.97	0.14	28,30,33,36	0
2	SO4	B	703	5/5	0.97	0.18	26,29,32,36	0
2	SO4	C	702	5/5	0.97	0.12	30,30,37,37	0
2	SO4	A	702	5/5	0.98	0.12	27,28,33,33	0
2	SO4	A	701	5/5	0.98	0.08	23,24,26,27	0
2	SO4	C	701	5/5	0.99	0.06	22,23,26,30	0
2	SO4	B	701	5/5	0.99	0.08	18,20,21,24	0

6.5 Other polymers (i)

There are no such residues in this entry.