W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

Full wwPDB X-ray Structure Validation Report (i)

May 26, 2020 – 11:48 pm BST

PDB ID	:	5VPJ
Title	:	The crystal structure of a thioesteras from Actinomadura verrucosospora.
Authors	:	Tan, K.; Joachimiak, G.; Endres, M.; Phillips Jr., G.N.; Joachmiak, A.; Midwest Center for Structural Genomics (MCSG); Enzyme Discovery for Natural Product Biosynthesis (NatPro)
Deposited on	:	2017-05-05
Resolution	:	2.35 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 following versions of software and data (see references (i)) 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	:	2.11
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)	:	2.11

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 2.35 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.
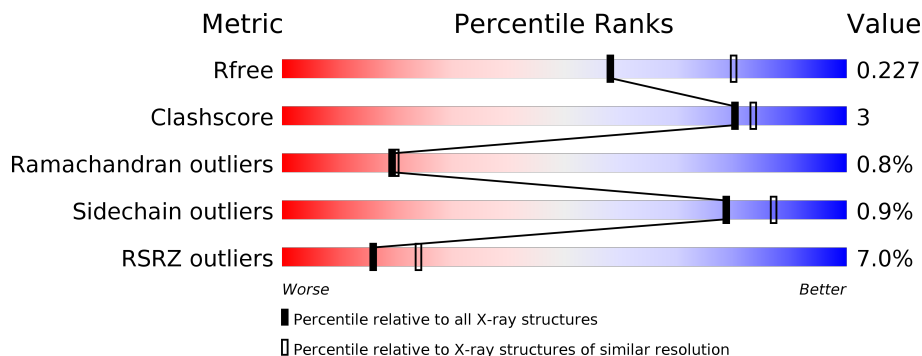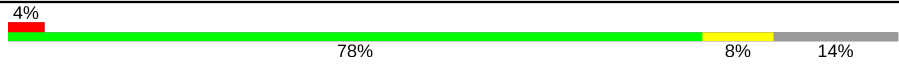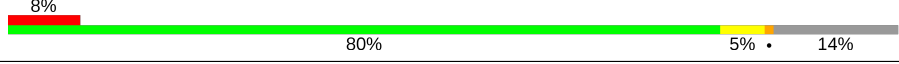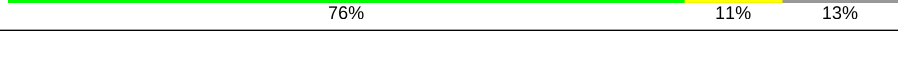
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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 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	159	3% / 81% 6% 13%
1	B	159	6% / 79% 8% 13%
1	C	159	13% / 79% 8% 13%
1	D	159	3% / 82% 5% 13%
1	E	159	9% / 75% 9% 15%
1	F	159	4% / 82% 5% 13%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	159	4% ... 78% ... 8% ... 14%
1	H	159	8% ... 80% ... 5% ... 14%
1	I	159	3% ... 77% ... 9% ... 14%
1	J	159	8% ... 77% ... 7% ... 16%
1	K	159	8% ... 73% ... 9% ... 18%
1	L	159	3% ... 76% ... 11% ... 13%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 13045 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 Thioesterase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	138	Total	C	N	O	S	Se	0	0	0
			1085	691	182	205	4	3			
1	B	138	Total	C	N	O	S	Se	0	0	0
			1079	687	183	202	4	3			
1	C	139	Total	C	N	O	S	Se	0	0	0
			1058	677	177	197	4	3			
1	D	138	Total	C	N	O	S	Se	0	0	0
			1084	690	182	205	4	3			
1	E	135	Total	C	N	O	S	Se	0	0	0
			1049	672	176	194	4	3			
1	F	139	Total	C	N	O	S	Se	0	0	0
			1084	690	181	206	4	3			
1	G	136	Total	C	N	O	S	Se	0	0	0
			1070	682	178	203	4	3			
1	H	136	Total	C	N	O	S	Se	0	0	0
			1065	681	176	201	4	3			
1	I	136	Total	C	N	O	S	Se	0	0	0
			1070	684	178	201	4	3			
1	J	133	Total	C	N	O	S	Se	0	0	0
			1025	652	174	192	4	3			
1	K	131	Total	C	N	O	S	Se	0	0	0
			1022	654	171	190	4	3			
1	L	139	Total	C	N	O	S	Se	0	0	0
			1090	693	184	206	4	3			

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	H	1	Total	Cl	0	0
			1	1		
2	G	1	Total	Cl	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Cl 1	0	0
2	K	2	Total 2	Cl 2	0	0

- Molecule 3 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).

PG4
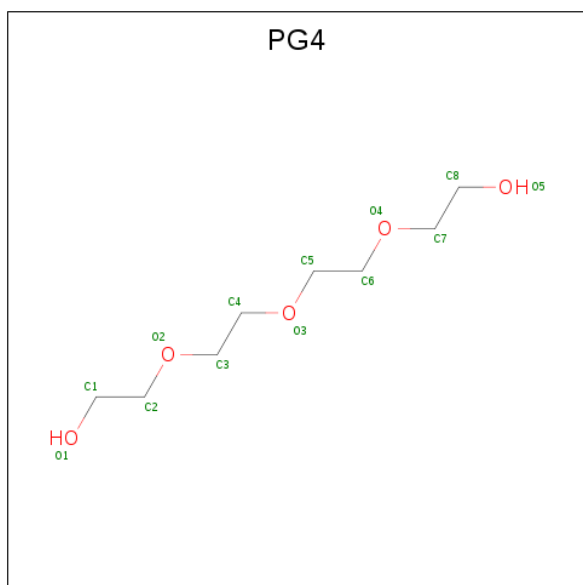
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	D	1	Total 13	C 8	O 5	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	32	Total 32	O 32	0	0
4	B	22	Total 22	O 22	0	0
4	C	25	Total 25	O 25	0	0
4	D	38	Total 38	O 38	0	0
4	E	7	Total 7	O 7	0	0
4	F	18	Total 18	O 18	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	G	34	Total 34	O 34	0	0
4	H	15	Total 15	O 15	0	0
4	I	26	Total 26	O 26	0	0
4	J	8	Total 8	O 8	0	0
4	K	10	Total 10	O 10	0	0
4	L	11	Total 11	O 11	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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: Thioesterase

Chain A:

- Molecule 1: Thioesterase

Chain B:

- Molecule 1: Thioesterase

Chain C:

- Molecule 1: Thioesterase

Chain D:

- Molecule 1: Thioesterase

Chain E:

PRO HIS GLY ALA LEU ARG PRO ARG SER GLU THR PRO PRO ARG

- Molecule 1: Thioesterase

Chain F:

4%

82% 5% 13%

SER ASN ALA MSE R2 T12 A33 A40 L56 F86 Y96 R97 A103 A119 E120 ASP G122 L123 D137 P141 ASP PRO HIS GLY ALA LEU ARG PRO ARG SER GLU THR PRO PRO ARG

- Molecule 1: Thioesterase

Chain G:

4%

78% 8% 14%

SER ASN ALA MSE ARG A3 T12 N17 L28 R35 E36 L39 M50 L72 L81 F86 V98 T99 D100 G101 P102 Q111 A119 GLU ASP GLY L123 P141 ASP PRO HIS GLY ALA LEU ARG PRO ARG SER GLU THR PRO PRO ARG

- Molecule 1: Thioesterase

Chain H:

8%

80% 5% 14%

SER ASN ALA MSE R2 E6 Y7 R8 T12 F13 A14 N17 L28 E36 D52 D53 L54 L72 F86 D100 G101 L117 R118 A119 GLU ASP GLY LEU T124 P125 V126 L136 A140 P141 ASP PRO HIS GLY ALA LEU ARG PRO ARG SER GLU THR PRO PRO

ARG

- Molecule 1: Thioesterase

Chain I:

3%

77% 9% 14%

SER ASN ALA MSE ARG A3 T12 L28 R37 F38 L39 M50 L56 C63 L72 L81 F86 T99 P102 G109 L117 R118 ALA ASP ASP GLY L123 D137 P141 D142 PRO HIS GLY ALA LEU ARG PRO ARG SER GLU THR PRO PRO ARG

- Molecule 1: Thioesterase

Chain J:

8%

77% 7% 16%

SER ASN ALA MSE ARG ALA Y4 T12 L28 R35 E36 A40 T46 V47 A48 H51 D52 ASP L54 S67 F93 D100 G101 P102 A103 R104 Q111 L117 R118 ALA ASP GLY L123 V126 E127 V128 D129 D130 E131 Y139 A140 P141 ASP PRO HIS GLY

ALA LEU ARG PRO ARG SER GLU THR PRO PRO ARG

- Molecule 1: Thioesterase

Chain K:

8%

73% 9% 18%

SER ASN ALA MSE ARG ALA Y4 F24 L28 R35 E36 L39 V47 A48 R49 M50 D53 L54 F65 L72 I84 D85 F86 H87 R97 V98 T99 D100 R104 Q111 C115 T116 L117 ARG ALA GLU ASP GLY LEU T124 P125 V126 A140 PRO ASP PRO
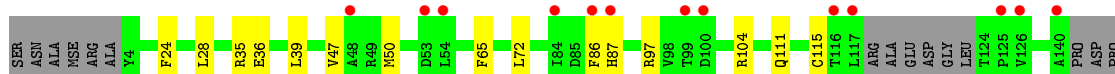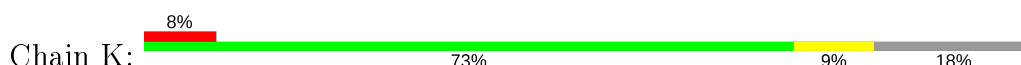
HIS GLY ALA LEU ARG PRO ARG SER GLU THR PRO PRO ARG

• Molecule 1: Thioesterase

Chain L:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	105.68Å 96.36Å 114.16Å 90.00° 99.16° 90.00°	Depositor
Resolution (Å)	33.68 – 2.35 33.68 – 2.35	Depositor EDS
% Data completeness (in resolution range)	86.8 (33.68-2.35) 86.8 (33.68-2.35)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	6.15 (at 2.34Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.186 , 0.227 0.188 , 0.227	Depositor DCC
R_{free} test set	4168 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	31.2	Xtriage
Anisotropy	0.080	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 46.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	13045	wwPDB-VP
Average B, all atoms (Å^2)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.17% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: PG4, 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	0.21	0/1106	0.36	0/1500
1	B	0.21	0/1100	0.36	0/1491
1	C	0.21	0/1079	0.38	0/1467
1	D	0.25	0/1105	0.38	0/1499
1	E	0.21	0/1070	0.37	0/1453
1	F	0.21	0/1105	0.36	0/1499
1	G	0.21	0/1091	0.36	0/1480
1	H	0.21	0/1086	0.37	0/1473
1	I	0.21	0/1091	0.36	0/1480
1	J	0.21	0/1044	0.37	0/1416
1	K	0.21	0/1042	0.36	0/1413
1	L	0.23	0/1111	0.37	0/1506
All	All	0.22	0/13030	0.37	0/17677

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	1085	0	1037	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1079	0	1031	8	0
1	C	1058	0	992	8	0
1	D	1084	0	1038	5	0
1	E	1049	0	1000	8	0
1	F	1084	0	1031	4	0
1	G	1070	0	1024	7	0
1	H	1065	0	1018	5	0
1	I	1070	0	1028	10	0
1	J	1025	0	957	7	0
1	K	1022	0	968	8	0
1	L	1090	0	1045	10	0
2	A	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
2	K	2	0	0	0	0
3	D	13	0	18	1	0
4	A	32	0	0	0	0
4	B	22	0	0	0	0
4	C	25	0	0	0	0
4	D	38	0	0	0	0
4	E	7	0	0	0	0
4	F	18	0	0	0	0
4	G	34	0	0	0	0
4	H	15	0	0	0	0
4	I	26	0	0	0	0
4	J	8	0	0	0	0
4	K	10	0	0	0	0
4	L	11	0	0	0	0
All	All	13045	0	12187	67	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.

All (67) 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:E:39:LEU:HD13	1:E:50:MSE:HE1	1.75	0.68
1:G:35:ARG:NH1	1:G:111:GLN:OE1	2.28	0.67
1:K:28:LEU:HD13	1:L:28:LEU:HG	1.77	0.67
1:G:39:LEU:HD13	1:G:50:MSE:HE1	1.79	0.64
1:I:28:LEU:HD13	1:J:28:LEU:HG	1.78	0.64

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:I:72:LEU:HD22	1:L:40:ALA:HB2	1.84	0.60
1:H:12:THR:HG23	1:H:14:ALA:H	1.67	0.59
1:I:39:LEU:HD13	1:I:50:MSE:HE1	1.84	0.59
1:C:116:THR:OG1	1:C:124:THR:N	2.37	0.57
1:G:28:LEU:HD13	1:H:28:LEU:HG	1.87	0.57
1:K:47:VAL:HA	1:K:50:MSE:HE3	1.86	0.57
1:I:50:MSE:O	1:I:118:ARG:NH2	2.41	0.54
1:D:91:MSE:SE	3:D:201:PG4:H41	2.57	0.54
1:K:39:LEU:HD13	1:K:50:MSE:HE1	1.91	0.53
1:E:35:ARG:NH1	1:E:111:GLN:OE1	2.39	0.53
1:K:35:ARG:NH1	1:K:111:GLN:OE1	2.42	0.53
1:L:35:ARG:NH1	1:L:111:GLN:OE1	2.40	0.52
1:C:117:LEU:HG	1:C:126:VAL:HG21	1.91	0.52
1:A:72:LEU:HD22	1:D:40:ALA:HB2	1.92	0.52
1:D:2:ARG:NE	1:I:137:ASP:OD2	2.44	0.51
1:E:6:GLU:OE1	1:E:8:ARG:NH1	2.44	0.50
1:E:40:ALA:HB2	1:H:72:LEU:HD22	1.93	0.49
1:J:46:THR:N	1:J:131:GLU:OE1	2.46	0.49
1:F:33:ALA:HA	1:G:12:THR:HG21	1.95	0.49
1:A:50:MSE:HE2	1:A:56:LEU:HD13	1.95	0.49
1:I:50:MSE:HE2	1:I:56:LEU:HD13	1.95	0.48
1:B:72:LEU:HD22	1:C:40:ALA:HB2	1.95	0.48
1:A:39:LEU:HD13	1:A:50:MSE:HE1	1.97	0.47
1:D:39:LEU:HD13	1:D:50:MSE:HE1	1.97	0.47
1:J:35:ARG:HG3	1:J:36:GLU:HG3	1.97	0.47
1:L:39:LEU:HD13	1:L:50:MSE:HE1	1.97	0.47
1:I:63:CYS:HA	1:I:109:GLY:HA2	1.97	0.47
1:J:67:SER:OG	1:J:104:ARG:NH2	2.48	0.47
1:J:93:PHE:HE1	1:J:111:GLN:HG2	1.79	0.46
1:C:39:LEU:HD13	1:C:50:MSE:HE1	1.99	0.45
1:E:16:THR:HG21	1:E:71:ALA:HB2	1.98	0.45
1:F:40:ALA:HB2	1:G:72:LEU:HG	1.99	0.45
1:C:81:LEU:HD23	1:C:140:ALA:HB2	1.99	0.44
1:B:81:LEU:O	1:B:141:PRO:HD3	2.18	0.44
1:L:50:MSE:HE2	1:L:56:LEU:HD13	2.00	0.44
1:K:97:ARG:HH21	1:K:104:ARG:HH11	1.65	0.43
1:B:46:THR:O	1:B:50:MSE:HG3	2.19	0.43
1:I:37:ARG:HG2	1:L:10:LEU:HD23	1.99	0.43
1:B:12:THR:HB	1:C:36:GLU:OE1	2.19	0.42
1:L:81:LEU:HD22	1:L:136:LEU:HB3	2.02	0.42
1:B:87:HIS:CD2	1:B:125:PRO:HB3	2.54	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:6:GLU:OE2	1:H:8:ARG:NH2	2.49	0.42
1:L:46:THR:HG22	1:L:50:MSE:HE3	2.02	0.42
1:J:40:ALA:HB2	1:K:72:LEU:HD22	2.01	0.42
1:B:36:GLU:OE1	1:C:12:THR:HB	2.20	0.42
1:A:36:GLU:OE1	1:D:12:THR:HB	2.20	0.41
1:G:81:LEU:O	1:G:141:PRO:HD3	2.21	0.41
1:I:81:LEU:O	1:I:141:PRO:HD3	2.20	0.41
1:I:12:THR:HB	1:L:36:GLU:OE1	2.21	0.41
1:B:39:LEU:HD13	1:B:50:MSE:HE1	2.02	0.41
1:K:24:PHE:HB2	1:K:65:PHE:CZ	2.55	0.41
1:L:124:THR:HA	1:L:125:PRO:HD3	1.96	0.41
1:A:81:LEU:O	1:A:141:PRO:HD3	2.21	0.40
1:C:117:LEU:HB2	1:C:124:THR:HB	2.03	0.40
1:E:101:GLY:HA2	1:E:102:PRO:HA	1.87	0.40
1:E:91:MSE:HE3	1:E:111:GLN:HG3	2.03	0.40
1:F:96:TYR:HB3	1:F:103:ALA:HB1	2.03	0.40
1:J:12:THR:HB	1:K:36:GLU:OE1	2.21	0.40
1:B:85:ASP:O	1:B:87:HIS:N	2.55	0.40
1:E:12:THR:HB	1:H:36:GLU:OE1	2.21	0.40
1:A:2:ARG:HB3	1:A:3:ALA:H	1.70	0.40
1:F:12:THR:HB	1:G:36:GLU:OE1	2.21	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	134/159 (84%)	130 (97%)	3 (2%)	1 (1%)	22	23
1	B	134/159 (84%)	131 (98%)	2 (2%)	1 (1%)	22	23
1	C	135/159 (85%)	128 (95%)	5 (4%)	2 (2%)	10	8
1	D	134/159 (84%)	131 (98%)	2 (2%)	1 (1%)	22	23

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	131/159 (82%)	125 (95%)	5 (4%)	1 (1%)	19	20
1	F	135/159 (85%)	131 (97%)	3 (2%)	1 (1%)	22	23
1	G	132/159 (83%)	129 (98%)	2 (2%)	1 (1%)	19	20
1	H	132/159 (83%)	126 (96%)	5 (4%)	1 (1%)	19	20
1	I	132/159 (83%)	126 (96%)	5 (4%)	1 (1%)	19	20
1	J	127/159 (80%)	123 (97%)	4 (3%)	0	100	100
1	K	127/159 (80%)	124 (98%)	2 (2%)	1 (1%)	19	20
1	L	135/159 (85%)	131 (97%)	3 (2%)	1 (1%)	22	23
All	All	1588/1908 (83%)	1535 (97%)	41 (3%)	12 (1%)	19	20

All (12) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	86	PHE
1	B	86	PHE
1	I	86	PHE
1	L	86	PHE
1	C	86	PHE
1	D	86	PHE
1	F	86	PHE
1	G	86	PHE
1	H	86	PHE
1	K	86	PHE
1	E	85	ASP
1	C	141	PRO

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	114/130 (88%)	114 (100%)	0	100	100
1	B	112/130 (86%)	112 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	106/130 (82%)	106 (100%)	0	100	100
1	D	114/130 (88%)	113 (99%)	1 (1%)	78	87
1	E	108/130 (83%)	107 (99%)	1 (1%)	78	87
1	F	113/130 (87%)	111 (98%)	2 (2%)	59	70
1	G	113/130 (87%)	112 (99%)	1 (1%)	78	87
1	H	112/130 (86%)	110 (98%)	2 (2%)	59	70
1	I	113/130 (87%)	113 (100%)	0	100	100
1	J	104/130 (80%)	104 (100%)	0	100	100
1	K	105/130 (81%)	103 (98%)	2 (2%)	57	68
1	L	114/130 (88%)	111 (97%)	3 (3%)	46	56
All	All	1328/1560 (85%)	1316 (99%)	12 (1%)	78	87

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	17	ASN
1	E	108	ARG
1	F	56	LEU
1	F	97	ARG
1	G	17	ASN
1	H	12	THR
1	H	17	ASN
1	K	87	HIS
1	K	115	CYS
1	L	2	ARG
1	L	17	ASN
1	L	68	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	K	17	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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 5 are monoatomic - leaving 1 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	PG4	D	201	-	12,12,12	0.68	0	11,11,11	0.24	0

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	PG4	D	201	-	-	4/10/10/10	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	201	PG4	O3-C5-C6-O4
3	D	201	PG4	O4-C7-C8-O5
3	D	201	PG4	O2-C3-C4-O3
3	D	201	PG4	C6-C5-O3-C4

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	201	PG4	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(Å²)	Q<0.9
1	A	135/159 (84%)	-0.19	4 (2%)	50	61	17, 32, 84, 117	0
1	B	135/159 (84%)	0.39	10 (7%)	14	22	20, 43, 94, 117	0
1	C	136/159 (85%)	0.59	20 (14%)	2	3	17, 45, 109, 150	0
1	D	134/159 (84%)	-0.08	4 (2%)	50	61	18, 32, 78, 110	0
1	E	132/159 (83%)	0.52	15 (11%)	5	7	33, 63, 113, 143	0
1	F	136/159 (85%)	0.08	6 (4%)	34	46	22, 45, 89, 121	0
1	G	133/159 (83%)	0.13	7 (5%)	26	38	22, 34, 68, 136	0
1	H	133/159 (83%)	0.46	13 (9%)	7	12	21, 55, 111, 155	0
1	I	133/159 (83%)	-0.09	4 (3%)	50	61	23, 38, 76, 127	0
1	J	130/159 (81%)	0.53	12 (9%)	9	14	29, 62, 117, 135	0
1	K	128/159 (80%)	0.64	13 (10%)	6	11	35, 68, 123, 136	0
1	L	135/159 (84%)	0.11	4 (2%)	50	61	26, 50, 88, 130	0
All	All	1600/1908 (83%)	0.26	112 (7%)	16	24	17, 47, 107, 155	0

All (112) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	123	LEU	6.2
1	H	119	ALA	6.1
1	L	141	PRO	5.7
1	C	126	VAL	5.2
1	H	117	LEU	4.9
1	E	54	LEU	4.8
1	E	86	PHE	4.8
1	C	121	ASP	4.7
1	B	119	ALA	4.7
1	G	101	GLY	4.5
1	C	117	LEU	4.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	K	116	THR	4.2
1	B	120	GLU	4.2
1	G	123	LEU	4.1
1	C	86	PHE	4.1
1	K	99	THR	4.1
1	G	119	ALA	4.0
1	C	125	PRO	4.0
1	K	140	ALA	4.0
1	K	86	PHE	4.0
1	G	100	ASP	4.0
1	C	119	ALA	3.9
1	G	99	THR	3.9
1	B	86	PHE	3.8
1	E	99	THR	3.8
1	L	119	ALA	3.8
1	B	122	GLY	3.7
1	G	102	PRO	3.7
1	C	122	GLY	3.7
1	A	141	PRO	3.7
1	H	101	GLY	3.7
1	H	126	VAL	3.5
1	C	124	THR	3.5
1	E	55	ALA	3.5
1	K	53	ASP	3.5
1	K	54	LEU	3.4
1	B	118	ARG	3.4
1	J	140	ALA	3.3
1	F	120	GLU	3.3
1	F	123	LEU	3.3
1	E	102	PRO	3.2
1	E	141	PRO	3.2
1	E	101	GLY	3.2
1	J	48	ALA	3.2
1	D	141	PRO	3.2
1	C	141	PRO	3.1
1	E	100	ASP	3.1
1	J	126	VAL	3.1
1	I	117	LEU	3.1
1	E	115	CYS	3.1
1	H	52	ASP	3.0
1	F	137	ASP	3.0
1	J	130	ASP	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	J	117	LEU	2.9
1	H	100	ASP	2.9
1	K	125	PRO	2.9
1	I	102	PRO	2.8
1	E	126	VAL	2.8
1	H	141	PRO	2.8
1	J	101	GLY	2.8
1	E	17	ASN	2.8
1	C	53	ASP	2.7
1	C	120	GLU	2.7
1	K	100	ASP	2.7
1	C	140	ALA	2.7
1	F	119	ALA	2.7
1	J	102	PRO	2.7
1	H	53	ASP	2.7
1	C	136	LEU	2.6
1	H	86	PHE	2.6
1	E	87	HIS	2.6
1	C	116	THR	2.6
1	J	54	LEU	2.6
1	K	117	LEU	2.5
1	H	136	LEU	2.5
1	D	138	ALA	2.4
1	C	54	LEU	2.4
1	E	124	THR	2.4
1	I	99	THR	2.4
1	G	98	VAL	2.3
1	D	51	HIS	2.3
1	C	48	ALA	2.3
1	D	137	ASP	2.3
1	K	84	ILE	2.3
1	J	100	ASP	2.3
1	C	85	ASP	2.2
1	E	52	ASP	2.2
1	E	140	ALA	2.2
1	K	48	ALA	2.2
1	C	101	GLY	2.2
1	L	2	ARG	2.2
1	H	54	LEU	2.2
1	H	2	ARG	2.2
1	J	128	VAL	2.2
1	J	51	HIS	2.2

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	K	87	HIS	2.2
1	B	125	PRO	2.2
1	B	130	ASP	2.2
1	H	140	ALA	2.2
1	K	126	VAL	2.1
1	F	141	PRO	2.1
1	B	101	GLY	2.1
1	L	52	ASP	2.1
1	C	118	ARG	2.1
1	A	122	GLY	2.1
1	A	140	ALA	2.1
1	C	55	ALA	2.1
1	B	121	ASP	2.1
1	F	2	ARG	2.1
1	J	139	TYR	2.1
1	A	137	ASP	2.0
1	B	117	LEU	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 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. 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(Å²)	Q<0.9
3	PG4	D	201	13/13	0.86	0.20	46,59,64,66	0
2	CL	G	201	1/1	0.88	0.44	85,85,85,85	0
2	CL	K	202	1/1	0.93	0.23	73,73,73,73	0
2	CL	A	201	1/1	0.93	0.30	84,84,84,84	0
2	CL	K	201	1/1	0.96	0.18	62,62,62,62	0
2	CL	H	201	1/1	0.98	0.15	35,35,35,35	1

6.5 Other polymers (i)

There are no such residues in this entry.