Full wwPDB X-ray Structure Validation Report (i)

Oct 20, 2025 – 10:37 AM JST

PDB ID	:	9K7J / pdb_00009k7j
Title	:	Gamma-glutamyl peptidase 1 from Arabidopsis thaliana (ligand-free)
Authors	:	Sone, K.; Kashima, T.; Miyanaga, A.; Fushinobu, S.
Deposited on	:	2024-10-23
Resolution	:	1.90 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:

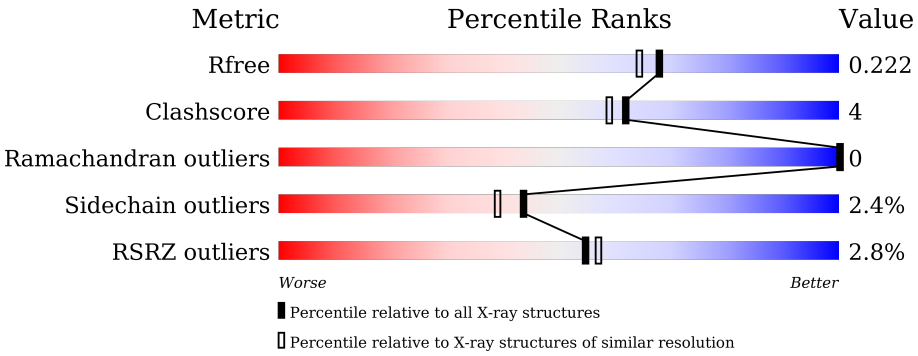
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.46

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 1.90 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}	164625	7293 (1.90-1.90)
Clashscore	180529	8090 (1.90-1.90)
Ramachandran outliers	177936	8022 (1.90-1.90)
Sidechain outliers	177891	8022 (1.90-1.90)
RSRZ outliers	164620	7292 (1.90-1.90)

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	260	5% 81% 11% •• 6%
1	B	260	% 82% 12% • 5%

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	ACT	A	301	-	-	X	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4379 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.

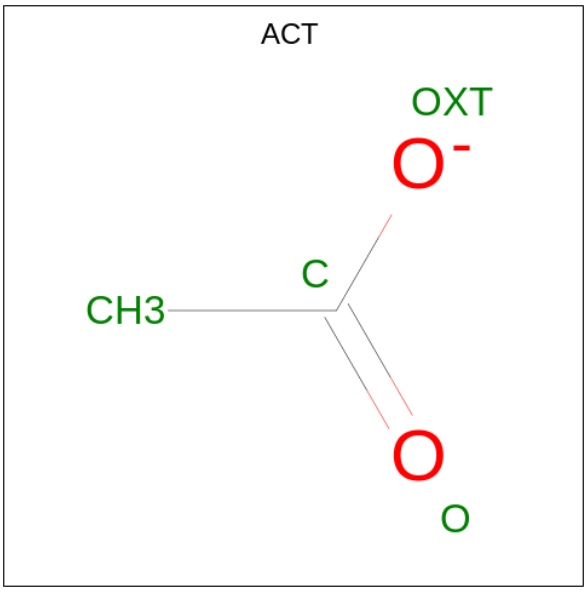
- Molecule 1 is a protein called Gamma-glutamyl peptidase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	245	Total	C	N	O	S	0	7	0
			2017	1297	332	380	8			
1	B	247	Total	C	N	O	S	0	3	0
			2001	1282	332	379	8			

There are 20 discrepancies between the modelled and reference sequences:

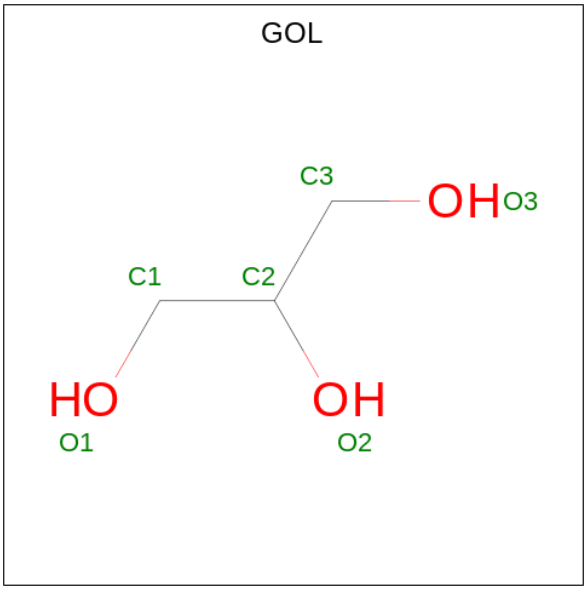
Chain	Residue	Modelled	Actual	Comment	Reference
A	-9	MET	-	initiating methionine	UNP Q9M0A7
A	-8	SER	-	expression tag	UNP Q9M0A7
A	-7	TYR	-	expression tag	UNP Q9M0A7
A	-6	TYR	-	expression tag	UNP Q9M0A7
A	-5	HIS	-	expression tag	UNP Q9M0A7
A	-4	HIS	-	expression tag	UNP Q9M0A7
A	-3	HIS	-	expression tag	UNP Q9M0A7
A	-2	HIS	-	expression tag	UNP Q9M0A7
A	-1	HIS	-	expression tag	UNP Q9M0A7
A	0	HIS	-	expression tag	UNP Q9M0A7
B	-9	MET	-	initiating methionine	UNP Q9M0A7
B	-8	SER	-	expression tag	UNP Q9M0A7
B	-7	TYR	-	expression tag	UNP Q9M0A7
B	-6	TYR	-	expression tag	UNP Q9M0A7
B	-5	HIS	-	expression tag	UNP Q9M0A7
B	-4	HIS	-	expression tag	UNP Q9M0A7
B	-3	HIS	-	expression tag	UNP Q9M0A7
B	-2	HIS	-	expression tag	UNP Q9M0A7
B	-1	HIS	-	expression tag	UNP Q9M0A7
B	0	HIS	-	expression tag	UNP Q9M0A7

- Molecule 2 is ACETATE ION (CCD ID: ACT) (formula: $C_2H_3O_2$).



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

- Molecule 3 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total C O 6 3 3	0	0

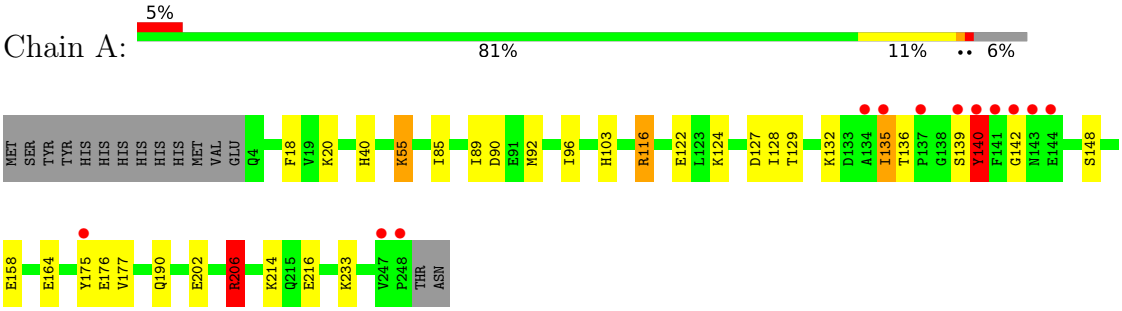
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	158	Total O 158 158	0	0
4	B	189	Total O 189 189	0	0

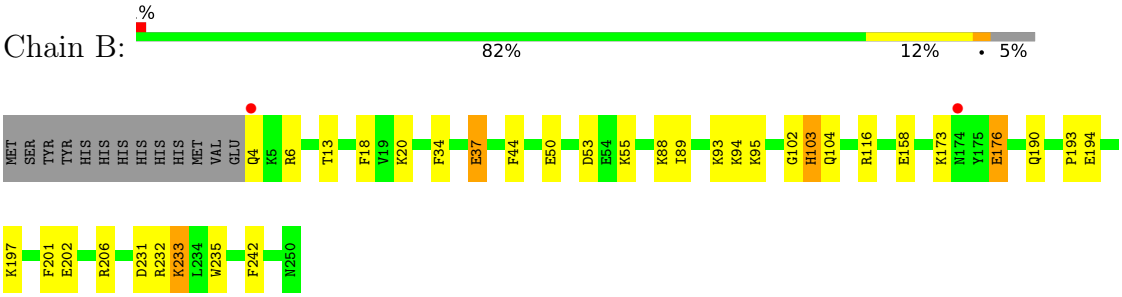
3 Residue-property plots (i)

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: Gamma-glutamyl peptidase 1



• Molecule 1: Gamma-glutamyl peptidase 1



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	72.54Å 75.59Å 103.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.72 – 1.90 46.72 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.9 (46.72-1.90) 100.0 (46.72-1.90)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.83 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.170 , 0.219 0.180 , 0.222	Depositor DCC
R_{free} test set	2358 reflections (5.20%)	wwPDB-VP
Wilson B-factor (Å^2)	22.0	Xtriage
Anisotropy	0.016	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 38.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.020 for k,h,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4379	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 4.35% 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: GOL, ACT

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	0.93	1/2062 (0.0%)	1.38	13/2781 (0.5%)
1	B	0.99	2/2043 (0.1%)	1.40	21/2752 (0.8%)
All	All	0.96	3/4105 (0.1%)	1.39	34/5533 (0.6%)

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
All	All	0	3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	116	ARG	NE-CZ	-6.79	1.25	1.33
1	B	231	ASP	C-O	-5.47	1.17	1.23
1	A	116	ARG	NE-CZ	-5.02	1.27	1.33

All (34) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	140	TYR	CB-CA-C	10.12	130.09	110.67
1	A	103	HIS	CA-CB-CG	8.48	122.28	113.80
1	B	13	THR	OG1-CB-CG2	-7.00	95.31	109.30
1	A	140	TYR	CA-CB-CG	6.68	125.93	113.90
1	B	103	HIS	CA-CB-CG	6.53	120.33	113.80
1	A	129	THR	CA-CB-OG1	-6.45	99.93	109.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	158	GLU	CG-CD-OE2	-6.42	103.64	118.40
1	A	20	LYS	CB-CA-C	-6.31	100.98	110.88
1	B	176[A]	GLU	CB-CG-CD	6.29	123.29	112.60
1	B	176[B]	GLU	CB-CG-CD	6.29	123.29	112.60
1	B	50	GLU	CB-CG-CD	6.25	123.23	112.60
1	A	92	MET	CG-SD-CE	-6.11	87.45	100.90
1	A	176	GLU	CB-CG-CD	6.10	122.97	112.60
1	B	158	GLU	CG-CD-OE1	6.08	132.38	118.40
1	B	88	LYS	N-CA-CB	-6.08	101.19	110.12
1	A	176	GLU	CB-CA-C	6.07	122.67	109.99
1	B	194	GLU	CB-CG-CD	5.98	122.77	112.60
1	A	90	ASP	CA-CB-CG	5.96	118.56	112.60
1	B	6	ARG	NE-CZ-NH2	5.90	124.51	119.20
1	A	140	TYR	N-CA-C	-5.88	105.00	111.82
1	B	158	GLU	CB-CA-C	5.69	123.35	110.67
1	B	34	PHE	CA-C-N	-5.55	115.83	122.16
1	B	34	PHE	C-N-CA	-5.55	115.83	122.16
1	B	53	ASP	CA-CB-CG	5.54	118.14	112.60
1	B	116	ARG	NE-CZ-NH1	-5.53	115.97	121.50
1	B	44	PHE	CA-CB-CG	-5.39	108.41	113.80
1	A	206	ARG	NE-CZ-NH2	5.37	124.04	119.20
1	A	55	LYS	CB-CG-CD	5.32	123.53	111.30
1	B	173	LYS	O-C-N	5.30	127.74	122.12
1	B	93	LYS	CB-CA-C	-5.24	104.28	111.73
1	B	37	GLU	CB-CG-CD	5.24	121.51	112.60
1	A	206	ARG	CA-CB-CG	5.21	124.51	114.10
1	B	102	GLY	CA-C-N	5.14	127.12	120.44
1	B	102	GLY	C-N-CA	5.14	127.12	120.44

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	116	ARG	Sidechain
1	A	206	ARG	Sidechain
1	B	232	ARG	Sidechain

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	2017	0	1989	21	0
1	B	2001	0	1970	12	0
2	A	4	0	3	2	0
2	B	4	0	3	0	0
3	B	6	0	8	0	0
4	A	158	0	0	2	0
4	B	189	0	0	1	0
All	All	4379	0	3973	32	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 4.

All (32) 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:202:GLU:OE2	1:B:206:ARG:NH2	2.10	0.84
1:A:202:GLU:OE2	1:A:206:ARG:NH2	2.11	0.83
1:A:139:SER:OG	1:A:140:TYR:N	2.12	0.76
1:A:18:PHE:CG	1:A:202:GLU:HG3	2.29	0.68
1:A:55:LYS:NZ	4:A:401:HOH:O	2.31	0.61
1:A:128:ILE:O	1:A:148:SER:HB2	2.02	0.59
1:A:158:GLU:OE1	1:A:175[B]:TYR:CD1	2.55	0.59
1:A:127:ASP:H	2:A:301:ACT:CH3	2.20	0.54
1:B:206:ARG:HG2	1:B:206:ARG:HH21	1.73	0.53
1:A:139:SER:O	1:A:142:GLY:N	2.40	0.53
1:A:18:PHE:CD1	1:A:202:GLU:HG3	2.44	0.52
1:A:175[B]:TYR:HE2	1:A:190:GLN:HE21	1.58	0.52
1:B:95:LYS:HE3	1:B:242:PHE:O	2.11	0.51
1:B:89:ILE:HG23	1:B:94:LYS:HB2	1.95	0.49
1:B:104:GLN:HE21	1:B:190:GLN:HE22	1.60	0.48
1:A:127:ASP:H	2:A:301:ACT:H1	1.76	0.48
1:A:135[B]:ILE:CG2	1:A:136[B]:THR:N	2.77	0.48
1:B:37:GLU:O	1:B:37:GLU:HG2	2.13	0.48
1:A:233:LYS:HE3	4:A:514:HOH:O	2.13	0.47
1:A:175[B]:TYR:CE2	1:A:177:VAL:O	2.68	0.46
1:B:95:LYS:CE	4:B:443:HOH:O	2.65	0.45
1:A:40:HIS:NE2	1:B:4:GLN:NE2	2.64	0.44
1:B:18:PHE:CD1	1:B:202:GLU:HG3	2.52	0.44
1:B:197:LYS:HE3	1:B:201:PHE:CZ	2.52	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:214:LYS:HE3	1:A:216:GLU:OE2	2.18	0.44
1:B:193:PRO:HB3	1:B:235:TRP:HB3	2.00	0.42
1:A:122:GLU:OE2	1:A:124:LYS:HE2	2.19	0.42
1:A:89:ILE:CG2	1:A:96:ILE:HD11	2.50	0.41
1:B:233:LYS:HB3	1:B:233:LYS:HE2	1.89	0.40
1:A:158:GLU:HB2	1:A:175[B]:TYR:CZ	2.56	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	250/260 (96%)	242 (97%)	8 (3%)	0	100	100
1	B	247/260 (95%)	239 (97%)	8 (3%)	0	100	100
All	All	497/520 (96%)	481 (97%)	16 (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	216/225 (96%)	210 (97%)	6 (3%)	38	33
1	B	215/225 (96%)	209 (97%)	6 (3%)	38	33

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	431/450 (96%)	419 (97%)	12 (3%)	44 33

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

Mol	Chain	Res	Type
1	A	85	ILE
1	A	132	LYS
1	A	135[A]	ILE
1	A	135[B]	ILE
1	A	140	TYR
1	A	164	GLU
1	B	20	LYS
1	B	55	LYS
1	B	103	HIS
1	B	176[A]	GLU
1	B	176[B]	GLU
1	B	233	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	103	HIS
1	A	104	GLN
1	A	190	GLN
1	B	4	GLN
1	B	104	GLN
1	B	190	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

3 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	ACT	A	301	-	3,3,3	2.06	2 (66%)	3,3,3	0.38	0
3	GOL	B	302	-	5,5,5	0.35	0	5,5,5	0.57	0
2	ACT	B	301	-	3,3,3	1.34	1 (33%)	3,3,3	1.35	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	GOL	B	302	-	-	0/4/4/4	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	301	ACT	OXT-C	-2.59	1.18	1.30
2	A	301	ACT	O-C	2.32	1.33	1.22
2	B	301	ACT	CH3-C	2.19	1.58	1.49

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	A	301	ACT	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

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	245/260 (94%)	-0.07	12 (4%) 36 37	11, 22, 48, 91	7 (2%)
1	B	247/260 (95%)	-0.46	2 (0%) 82 84	10, 18, 37, 53	3 (1%)
All	All	492/520 (94%)	-0.27	14 (2%) 55 57	10, 20, 44, 91	10 (2%)

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	247	VAL	6.4
1	A	248	PRO	5.7
1	A	135[A]	ILE	4.2
1	A	142	GLY	3.4
1	A	143	ASN	3.3
1	A	175[A]	TYR	2.9
1	A	134[A]	ALA	2.8
1	A	140	TYR	2.8
1	A	144	GLU	2.7
1	B	174	ASN	2.7
1	A	139	SER	2.6
1	A	137[A]	PRO	2.5
1	A	141	PHE	2.4
1	B	4	GLN	2.1

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no oligosaccharides in this entry.

6.4 Ligands

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	ACT	A	301	4/4	0.90	0.10	19,27,30,38	0
2	ACT	B	301	4/4	0.91	0.09	23,28,28,29	0
3	GOL	B	302	6/6	0.94	0.13	29,32,35,37	0

6.5 Other polymers

There are no such residues in this entry.