Full wwPDB X-ray Structure Validation Report (i)

May 15, 2020 – 10:16 pm BST

PDB ID	:	6SB1
Title	:	Crystal structure of murine perforin-2 P2 domain crystal form 1
Authors	:	Ni, T.; Ginger, L.; Gilbert, R.J.C.
Deposited on	:	2019-07-18
Resolution	:	2.05 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 (1)) were used in the production of this report:

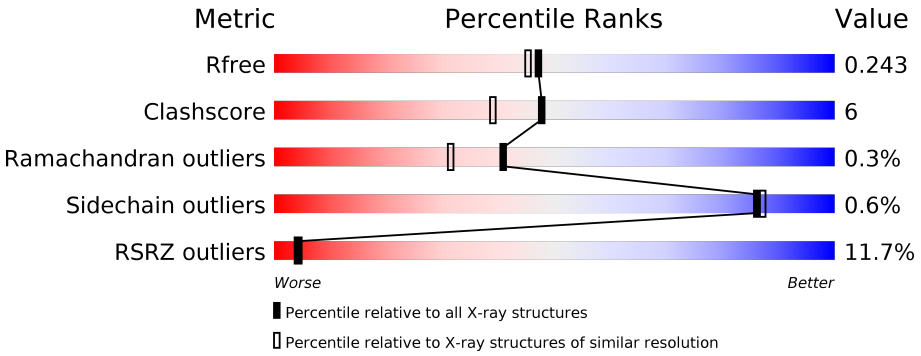
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.05 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	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)
RSRZ outliers	127900	1672 (2.04-2.04)

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	295	9% 57% 9% 34%
1	B	295	6% 59% 7% 34%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3098 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.

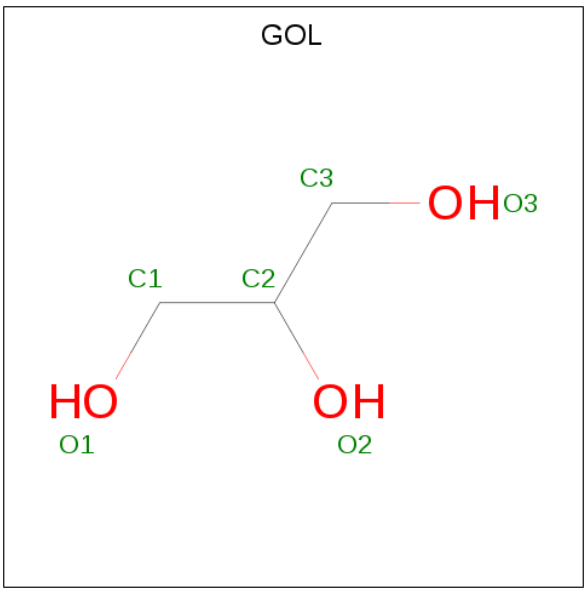
- Molecule 1 is a protein called Macrophage-expressed gene 1 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	196	Total	C	N	O	S	0	0	0
			1493	959	241	277	16			
1	B	196	Total	C	N	O	S	0	0	0
			1495	960	241	278	16			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	346	GLU	-	expression tag	UNP A1L314
A	347	THR	-	expression tag	UNP A1L314
A	348	GLY	-	expression tag	UNP A1L314
A	632	GLY	-	expression tag	UNP A1L314
A	633	THR	-	expression tag	UNP A1L314
A	634	LYS	-	expression tag	UNP A1L314
A	635	HIS	-	expression tag	UNP A1L314
A	636	HIS	-	expression tag	UNP A1L314
A	637	HIS	-	expression tag	UNP A1L314
A	638	HIS	-	expression tag	UNP A1L314
A	639	HIS	-	expression tag	UNP A1L314
A	640	HIS	-	expression tag	UNP A1L314
B	346	GLU	-	expression tag	UNP A1L314
B	347	THR	-	expression tag	UNP A1L314
B	348	GLY	-	expression tag	UNP A1L314
B	632	GLY	-	expression tag	UNP A1L314
B	633	THR	-	expression tag	UNP A1L314
B	634	LYS	-	expression tag	UNP A1L314
B	635	HIS	-	expression tag	UNP A1L314
B	636	HIS	-	expression tag	UNP A1L314
B	637	HIS	-	expression tag	UNP A1L314
B	638	HIS	-	expression tag	UNP A1L314
B	639	HIS	-	expression tag	UNP A1L314
B	640	HIS	-	expression tag	UNP A1L314

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Cl 1 1	0	0

- Molecule 4 is water.

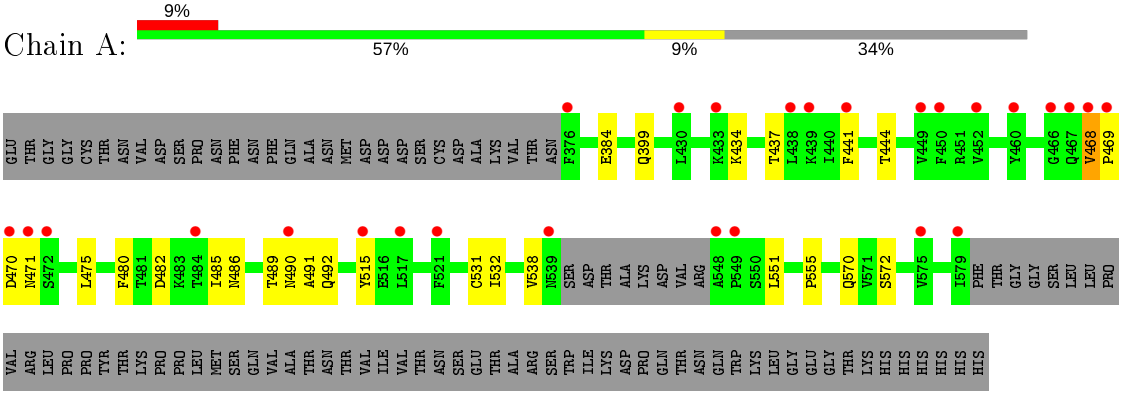
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	47	Total O 47 47	0	0
4	B	44	Total O 44 44	0	0

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.

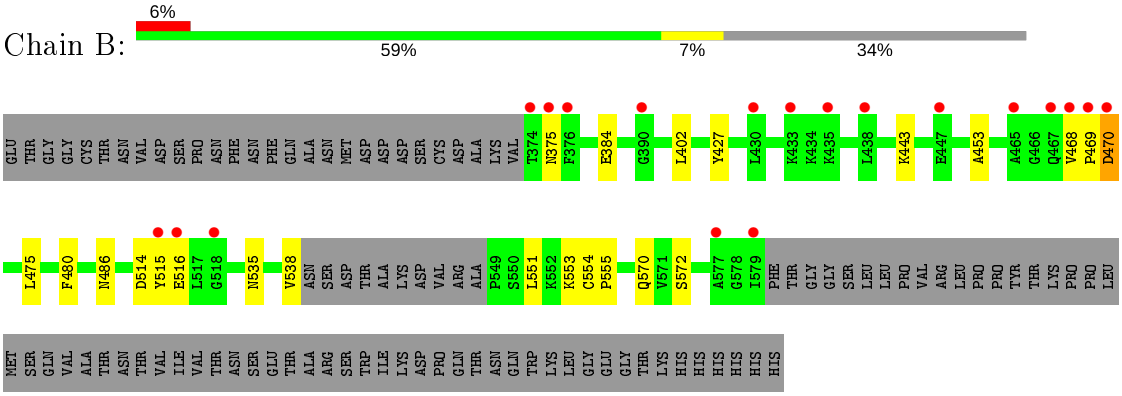
- Molecule 1: Macrophage-expressed gene 1 protein

Chain A:



- Molecule 1: Macrophage-expressed gene 1 protein

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	70.53Å 31.27Å 103.65Å 90.00° 91.42° 90.00°	Depositor
Resolution (Å)	57.63 – 2.05 57.63 – 2.05	Depositor EDS
% Data completeness (in resolution range)	99.3 (57.63-2.05) 99.4 (57.63-2.05)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.82 (at 2.05Å)	Xtriage
Refinement program	PHENIX dev_2645	Depositor
R, R_{free}	0.212 , 0.243 0.213 , 0.243	Depositor DCC
R_{free} test set	1432 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å^2)	43.3	Xtriage
Anisotropy	0.364	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 64.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.020 for h,-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3098	wwPDB-VP
Average B, all atoms (Å^2)	77.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 52.88 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 4.5687e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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, 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	0.47	0/1529	0.64	0/2068
1	B	0.46	0/1531	0.65	0/2070
All	All	0.46	0/3060	0.65	0/4138

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	1493	0	1450	19	0
1	B	1495	0	1453	17	0
2	A	12	0	16	2	0
2	B	6	0	8	1	0
3	A	1	0	0	0	0
4	A	47	0	0	0	2
4	B	44	0	0	2	1
All	All	3098	0	2927	36	3

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 6.

All (36) 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:570:GLN:NE2	1:B:572:SER:OG	2.14	0.80
1:B:443:LYS:NZ	4:B:801:HOH:O	2.13	0.80
1:A:570:GLN:NE2	1:A:572:SER:OG	2.16	0.79
1:A:384:GLU:HB3	2:A:701:GOL:H32	1.73	0.71
1:A:470:ASP:HB3	1:A:471:ASN:HA	1.74	0.69
1:A:434:LYS:HG2	1:A:444:THR:HG22	1.78	0.66
1:B:470:ASP:OD1	1:B:470:ASP:N	2.30	0.64
1:A:468:VAL:HB	1:A:469:PRO:HD2	1.80	0.62
1:B:538:VAL:HG21	1:B:551:LEU:HB2	1.83	0.59
1:A:538:VAL:HA	1:A:555:PRO:HB3	1.86	0.58
1:A:475:LEU:HD11	1:A:515:TYR:HA	1.87	0.56
1:A:399:GLN:HE22	2:A:702:GOL:H2	1.70	0.56
1:B:553:LYS:HD2	1:B:554:CYS:H	1.70	0.56
1:A:468:VAL:HB	1:A:469:PRO:CD	2.36	0.55
1:B:480:PHE:HB3	1:B:486:ASN:HB2	1.88	0.55
1:A:489:THR:HG23	1:A:491:ALA:H	1.71	0.55
1:A:437:THR:OG1	1:A:441:PHE:HB2	2.06	0.55
1:B:538:VAL:HA	1:B:555:PRO:HB3	1.90	0.54
1:A:538:VAL:HG21	1:A:551:LEU:HB2	1.89	0.54
1:B:468:VAL:N	1:B:469:PRO:CD	2.72	0.53
1:B:553:LYS:CD	1:B:554:CYS:H	2.24	0.50
1:B:384:GLU:HB3	2:B:701:GOL:H2	1.93	0.50
1:B:375:ASN:ND2	4:B:806:HOH:O	2.48	0.46
1:A:485:ILE:HD11	1:A:491:ALA:HA	1.98	0.46
1:B:468:VAL:N	1:B:469:PRO:HD3	2.31	0.45
1:B:535:ASN:O	1:B:538:VAL:HG22	2.16	0.45
1:B:514:ASP:OD1	1:B:516:GLU:HG2	2.17	0.45
1:B:475:LEU:HD11	1:B:515:TYR:HA	1.99	0.44
1:A:480:PHE:HB3	1:A:486:ASN:HB2	2.00	0.43
1:A:531:CYS:SG	1:A:532:ILE:HG23	2.59	0.42
1:B:468:VAL:H	1:B:469:PRO:HD3	1.85	0.42
1:B:427:TYR:CE1	1:B:453:ALA:HB2	2.55	0.42
1:A:486:ASN:O	1:A:490:ASN:N	2.48	0.42
1:A:470:ASP:CB	1:A:471:ASN:HA	2.42	0.42
1:A:482:ASP:O	1:A:492:GLN:NE2	2.51	0.40
1:A:486:ASN:HB3	1:A:489:THR:HG22	2.04	0.40

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:B:811:HOH:O	4:B:837:HOH:O[2_647]	2.08	0.12
4:A:803:HOH:O	4:A:835:HOH:O[1_545]	2.09	0.11
4:A:812:HOH:O	4:A:828:HOH:O[2_556]	2.09	0.11

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	192/295 (65%)	185 (96%)	6 (3%)	1 (0%)	29 18
1	B	192/295 (65%)	183 (95%)	9 (5%)	0	100 100
All	All	384/590 (65%)	368 (96%)	15 (4%)	1 (0%)	41 31

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	468	VAL

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	167/255 (66%)	167 (100%)	0	100 100
1	B	168/255 (66%)	166 (99%)	2 (1%)	71 70
All	All	335/510 (66%)	333 (99%)	2 (1%)	86 87

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	402	LEU
1	B	470	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	467	GLN
1	A	570	GLN
1	B	570	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 carbohydrates in this entry.

5.6 Ligand geometry

Of 4 ligands modelled in this entry, 1 is monoatomic - leaving 3 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	GOL	A	701	-	5,5,5	0.99	0	5,5,5	0.74	0
2	GOL	A	702	-	5,5,5	0.99	0	5,5,5	0.97	0
2	GOL	B	701	-	5,5,5	0.86	0	5,5,5	0.98	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
2	GOL	A	701	-	-	4/4/4/4	-
2	GOL	A	702	-	-	0/4/4/4	-
2	GOL	B	701	-	-	1/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	701	GOL	O2-C2-C3-O3
2	A	701	GOL	O1-C1-C2-C3
2	A	701	GOL	C1-C2-C3-O3
2	A	701	GOL	O1-C1-C2-O2
2	B	701	GOL	O1-C1-C2-C3

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	701	GOL	1	0
2	A	702	GOL	1	0
2	B	701	GOL	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	196/295 (66%)	0.99	27 (13%) 2 2	36, 67, 138, 165	0
1	B	196/295 (66%)	0.87	19 (9%) 7 8	34, 71, 138, 188	0
All	All	392/590 (66%)	0.93	46 (11%) 4 4	34, 70, 138, 188	0

All (46) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	579	ILE	8.9
1	A	468	VAL	8.5
1	A	548	ALA	6.4
1	B	467	GLN	6.3
1	B	376	PHE	5.8
1	A	469	PRO	5.1
1	A	466	GLY	4.9
1	B	374	THR	4.9
1	A	376	PHE	4.5
1	B	469	PRO	4.2
1	B	375	ASN	4.1
1	B	515	TYR	4.1
1	A	549	PRO	4.0
1	A	438	LEU	4.0
1	A	515	TYR	4.0
1	B	468	VAL	3.6
1	B	465	ALA	3.6
1	B	438	LEU	3.3
1	B	435	LYS	3.3
1	A	470	ASP	3.3
1	B	516	GLU	3.2
1	B	470	ASP	3.1
1	A	430	LEU	3.1
1	A	439	LYS	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	452	VAL	3.0
1	B	433	LYS	3.0
1	A	539	ASN	3.0
1	A	575	VAL	2.9
1	A	471	ASN	2.9
1	B	518	GLY	2.8
1	A	449	VAL	2.6
1	A	467	GLN	2.6
1	B	577	ALA	2.6
1	A	433	LYS	2.6
1	B	579	ILE	2.6
1	A	441	PHE	2.6
1	A	450	PHE	2.5
1	A	472	SER	2.4
1	A	517	LEU	2.4
1	A	490	ASN	2.4
1	B	430	LEU	2.4
1	A	521	PHE	2.3
1	B	447	GLU	2.1
1	B	390	GLY	2.1
1	A	484	THR	2.1
1	A	460	TYR	2.1

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, 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	GOL	A	702	6/6	0.83	0.26	76,89,90,97	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	GOL	B	701	6/6	0.87	0.24	52,76,85,87	0
3	CL	A	703	1/1	0.90	0.29	91,91,91,91	0
2	GOL	A	701	6/6	0.90	0.21	59,73,79,83	0

6.5 Other polymers (i)

There are no such residues in this entry.