WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Oct 31, 2023 – 03:03 PM JST

|		
PDB ID	:	5GJ6
Title	:	Functional and structural characterization of P[19] rotavirus VP8* interaction with histo-blood group antigens
Authors	:	Sun, X.; Duan, Z.
Deposited on	:	2016-06-28
Resolution	:	2.39 Å(reported)

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 2.39 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.
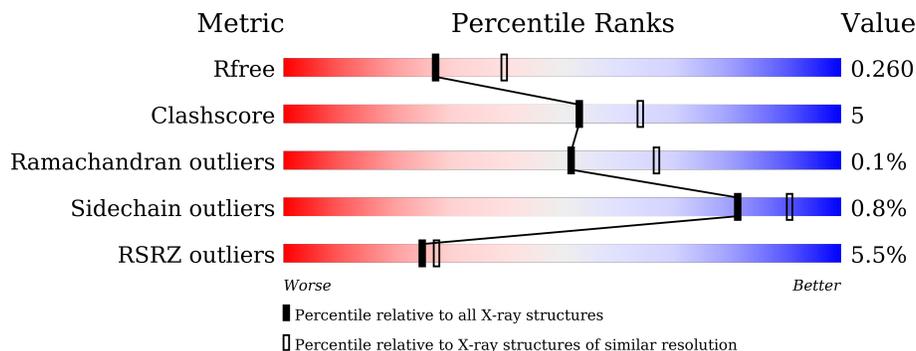
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5509 (2.40-2.36)
Clashscore	141614	6082 (2.40-2.36)
Ramachandran outliers	138981	5973 (2.40-2.36)
Sidechain outliers	138945	5975 (2.40-2.36)
RSRZ outliers	127900	5397 (2.40-2.36)

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 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	160	3% / 84% / 15% •
1	B	160	4% / 84% / 16%
1	C	160	2% / 91% / 8% •
1	D	160	4% / 85% / 14% ••
1	E	160	8% / 87% / 12% •
1	F	160	8% / 88% / 11% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	160	5% ▬▬ 90% ▬▬▬▬▬▬▬▬▬▬▬▬ 9% ·
1	H	160	9% ▬▬▬ 89% ▬▬▬▬▬▬▬▬▬▬▬ 11%

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	A	201	-	-	X	-

2 Entity composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 10614 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 Outer capsid protein VP4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	160	Total	C	N	O	S	0	0	0
			1279	808	216	252	3			
1	B	160	Total	C	N	O	S	0	0	0
			1279	808	216	252	3			
1	C	160	Total	C	N	O	S	0	0	0
			1279	808	216	252	3			
1	D	160	Total	C	N	O	S	0	0	0
			1279	808	216	252	3			
1	E	160	Total	C	N	O	S	0	0	0
			1279	808	216	252	3			
1	F	160	Total	C	N	O	S	0	0	0
			1279	808	216	252	3			
1	G	160	Total	C	N	O	S	0	0	0
			1279	808	216	252	3			
1	H	160	Total	C	N	O	S	0	0	0
			1279	808	216	252	3			

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

SO4

O1

O

O3 O⁻ — S — O⁻ O4

S

O

O2
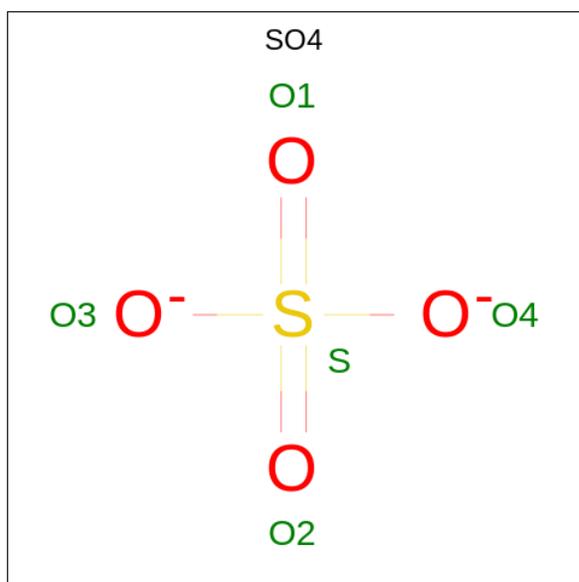
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0
2	E	1	Total 5	O 4	S 1	0	0
2	F	1	Total 5	O 4	S 1	0	0

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	54	Total 54	O 54	0	0
3	B	66	Total 66	O 66	0	0
3	C	47	Total 47	O 47	0	0
3	D	50	Total 50	O 50	0	0
3	E	48	Total 48	O 48	0	0
3	F	46	Total 46	O 46	0	0
3	G	26	Total 26	O 26	0	0
3	H	25	Total 25	O 25	0	0

3   Residue-property plots ⓘ

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: Outer capsid protein VP4
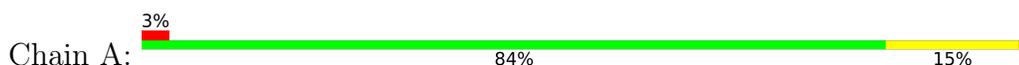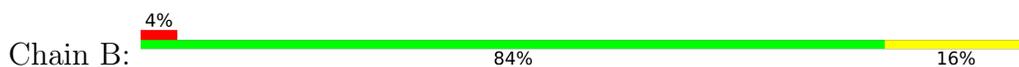
Chain A:

- Molecule 1: Outer capsid protein VP4

Chain B:

- Molecule 1: Outer capsid protein VP4

Chain C:

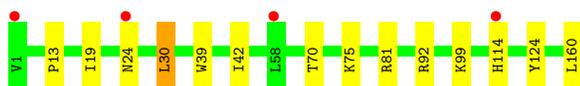- Molecule 1: Outer capsid protein VP4

Chain D:

- Molecule 1: Outer capsid protein VP4

Chain E:

- Molecule 1: Outer capsid protein VP4

Chain F:

8%

88%

11%

V1 L2 Y6 K12 P13 D16 N24 G27 L30 E31 D37 I43 Q55 L58 N62 I65 N72 K73 N85 R92 K99 K105 H106 G107 G108 Y124 E132 S150 K151 E154 Y155 L160

• Molecule 1: Outer capsid protein VP4

Chain G:

5%

90%

9%

V1 L2 D3 G4 P13 P14 I19 N24 S25 L30 T36 D37 V38 W39 Q55 K63 T70 R81 E87 R92 Y124 S128 Q148 E149 T153 L160

• Molecule 1: Outer capsid protein VP4

Chain H:

9%

89%

11%

V1 L2 P13 D16 Y17 W18 I19 S25 L30 E31 T36 D37 I45 Q55 Q64 F76 A86 R92 T93 L94 G102 A120 Y124 S125 T126 I143 I144 Q148 K151 C152 T153 E154 L160
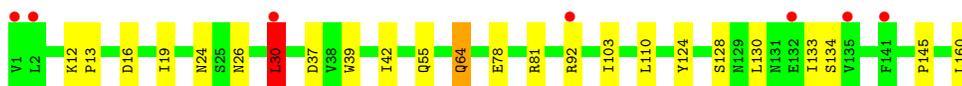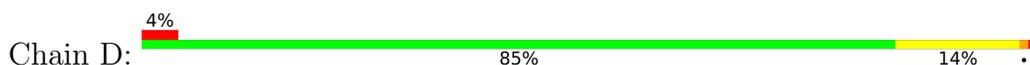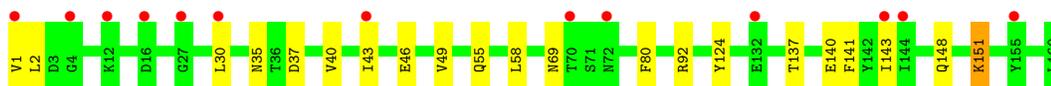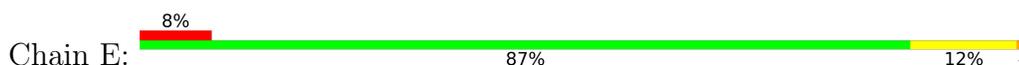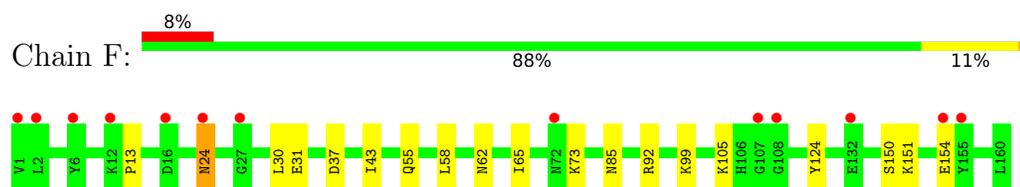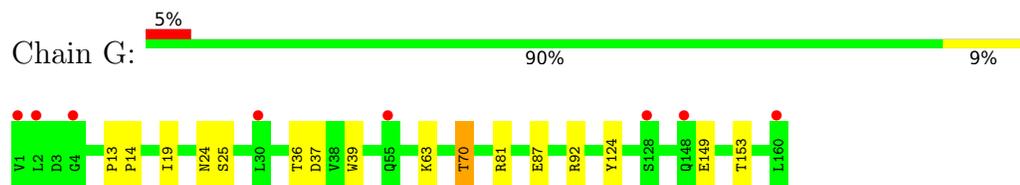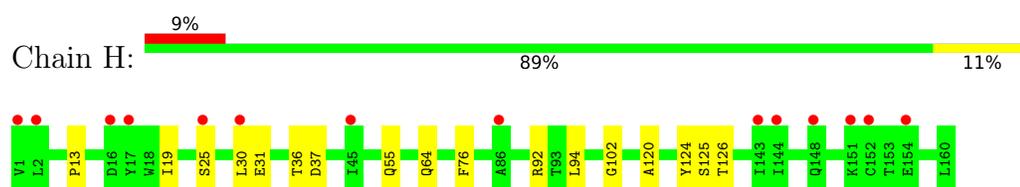
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	180.90Å 129.78Å 86.82Å 90.00° 115.90° 90.00°	Depositor
Resolution (Å)	33.82 – 2.39 34.84 – 2.39	Depositor EDS
% Data completeness (in resolution range)	98.3 (33.82-2.39) 98.4 (34.84-2.39)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	0.15	Depositor
$< I/\sigma(I) >$ [1]	2.79 (at 2.39Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.215 , 0.259 0.219 , 0.260	Depositor DCC
R_{free} test set	3420 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å^2)	34.8	Xtriage
Anisotropy	0.249	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 41.4	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.058 for -h-2*l,-k,l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	10614	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.02% 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	0.26	0/1310	0.44	0/1786
1	B	0.25	0/1310	0.44	0/1786
1	C	0.26	0/1310	0.45	0/1786
1	D	0.25	0/1310	0.46	1/1786 (0.1%)
1	E	0.23	0/1310	0.44	0/1786
1	F	0.26	0/1310	0.45	0/1786
1	G	0.24	0/1310	0.44	0/1786
1	H	0.24	0/1310	0.42	0/1786
All	All	0.25	0/10480	0.44	1/14288 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	30	LEU	CA-CB-CG	5.09	127.01	115.30

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	1279	0	1237	19	0
1	B	1279	0	1237	17	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1279	0	1237	11	0
1	D	1279	0	1237	19	1
1	E	1279	0	1237	14	0
1	F	1279	0	1237	17	0
1	G	1279	0	1237	9	0
1	H	1279	0	1237	11	1
2	A	5	0	0	2	0
2	B	5	0	0	0	0
2	E	5	0	0	0	0
2	F	5	0	0	0	0
3	A	54	0	0	4	2
3	B	66	0	0	4	0
3	C	47	0	0	2	1
3	D	50	0	0	7	1
3	E	48	0	0	3	1
3	F	46	0	0	3	0
3	G	26	0	0	0	1
3	H	25	0	0	4	0
All	All	10614	0	9896	108	4

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 5.

The worst 5 of 108 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:F:151:LYS:NZ	3:F:301:HOH:O	1.82	1.08
1:D:64:GLN:HE21	1:D:64:GLN:HA	1.22	1.01
1:H:120:ALA:O	3:H:201:HOH:O	1.88	0.93
1:C:92:ARG:HD3	1:C:114:HIS:HD2	1.35	0.91
1:C:160:LEU:OXT	3:C:201:HOH:O	1.93	0.85

All (4) 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)
3:A:331:HOH:O	3:E:331:HOH:O[1_554]	2.07	0.13
3:C:210:HOH:O	3:G:225:HOH:O[1_554]	2.11	0.09
3:A:313:HOH:O	3:D:229:HOH:O[4_748]	2.12	0.08
1:D:24:ASN:OD1	1:H:25:SER:OG[2_758]	2.16	0.04

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	158/160 (99%)	152 (96%)	6 (4%)	0	100	100
1	B	158/160 (99%)	153 (97%)	5 (3%)	0	100	100
1	C	158/160 (99%)	151 (96%)	7 (4%)	0	100	100
1	D	158/160 (99%)	153 (97%)	5 (3%)	0	100	100
1	E	158/160 (99%)	152 (96%)	5 (3%)	1 (1%)	25	34
1	F	158/160 (99%)	153 (97%)	5 (3%)	0	100	100
1	G	158/160 (99%)	152 (96%)	6 (4%)	0	100	100
1	H	158/160 (99%)	151 (96%)	7 (4%)	0	100	100
All	All	1264/1280 (99%)	1217 (96%)	46 (4%)	1 (0%)	51	67

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	2	LEU

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	146/146 (100%)	145 (99%)	1 (1%)	84	92
1	B	146/146 (100%)	145 (99%)	1 (1%)	84	92
1	C	146/146 (100%)	145 (99%)	1 (1%)	84	92
1	D	146/146 (100%)	144 (99%)	2 (1%)	67	81

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	146/146 (100%)	144 (99%)	2 (1%)	67	81
1	F	146/146 (100%)	145 (99%)	1 (1%)	84	92
1	G	146/146 (100%)	145 (99%)	1 (1%)	84	92
1	H	146/146 (100%)	146 (100%)	0	100	100
All	All	1168/1168 (100%)	1159 (99%)	9 (1%)	81	91

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	24	ASN
1	G	70	THR
1	D	30	LEU
1	D	64	GLN
1	E	1	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	106	HIS
1	F	62	ASN
1	E	64	GLN
1	E	55	GLN
1	F	55	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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

4 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	B	201	-	4,4,4	0.92	0	6,6,6	0.25	0
2	SO4	E	201	-	4,4,4	0.81	0	6,6,6	0.38	0
2	SO4	A	201	-	4,4,4	1.16	0	6,6,6	0.59	0
2	SO4	F	201	-	4,4,4	0.81	0	6,6,6	0.28	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	A	201	SO4	2	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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, 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	160/160 (100%)	0.40	5 (3%)	49	51	22, 34, 50, 86	0
1	B	160/160 (100%)	0.48	7 (4%)	34	37	21, 33, 48, 70	0
1	C	160/160 (100%)	0.43	4 (2%)	57	59	21, 35, 50, 66	0
1	D	160/160 (100%)	0.53	7 (4%)	34	37	22, 37, 52, 63	0
1	E	160/160 (100%)	0.75	13 (8%)	12	13	23, 39, 61, 84	0
1	F	160/160 (100%)	0.71	13 (8%)	12	13	23, 39, 61, 98	0
1	G	160/160 (100%)	0.58	8 (5%)	28	31	27, 42, 62, 94	0
1	H	160/160 (100%)	0.79	14 (8%)	10	11	26, 47, 67, 95	0
All	All	1280/1280 (100%)	0.58	71 (5%)	25	27	21, 38, 60, 98	0

The worst 5 of 71 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	1	VAL	7.3
1	F	1	VAL	6.3
1	H	2	LEU	6.0
1	C	1	VAL	4.7
1	G	1	VAL	4.3

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 monosaccharides 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, 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(Å^2)	Q<0.9
2	SO4	F	201	5/5	0.93	0.17	29,29,29,29	0
2	SO4	B	201	5/5	0.96	0.12	29,29,29,29	0
2	SO4	E	201	5/5	0.97	0.11	29,29,29,29	0
2	SO4	A	201	5/5	0.98	0.08	29,29,29,29	0

6.5 Other polymers ⓘ

There are no such residues in this entry.