Full wwPDB X-ray Structure Validation Report ⓘ

Sep 16, 2023 – 06:32 PM EDT

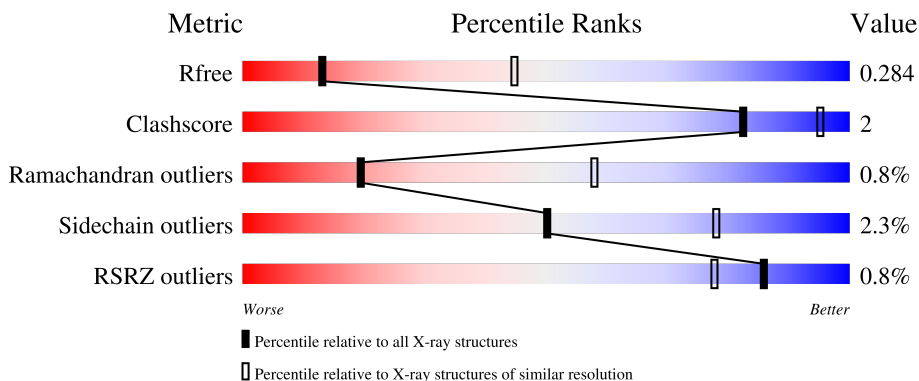|		
PDB ID	:	4W2O
Title	:	Anti-Marburgvirus Nucleoprotein Single Domain Antibody B Complexed with Nucleoprotein C-terminal domain
Authors	:	Taylor, A.B.; Garza, J.A.
Deposited on	:	2017-08-17
Resolution	:	3.20 Å(reported)

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 3.20 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	Percentile Ranks	Value
Rfree		0.284
Clashscore		2
Ramachandran outliers		0.8%
Sidechain outliers		2.3%
RSRZ outliers		0.8%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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	126	2% · 87% · 7% · 6%
1	C	126	2% · 86% · 8% · 6%
1	E	126	86% · 8% · 6%
1	G	126	.% · 85% · 9% · 6%
2	B	106	53% · 8% · 40%

Continued on next page...

W O R L D W I D E

wwPDB

PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	106	58% • 40%
2	F	106	60% 40%
2	H	106	56% • • 40%

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
3	SO4	A	201	-	-	X	-
3	SO4	C	201	-	-	X	-
3	SO4	E	201	-	-	X	-
3	SO4	G	201	-	-	X	-

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 5752 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 Anti-Marburgvirus Nucleoprotein Single Domain Antibody B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	118	Total	C	N	O	S	0	0	0
			900	560	163	173	4			
1	C	118	Total	C	N	O	S	0	0	0
			900	560	163	173	4			
1	E	118	Total	C	N	O	S	0	0	0
			900	560	163	173	4			
1	G	118	Total	C	N	O	S	0	0	0
			900	560	163	173	4			

- Molecule 2 is a protein called Nucleoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	64	Total	C	N	O	S	0	0	0
			528	335	91	100	2			
2	D	64	Total	C	N	O	S	0	0	0
			528	335	91	100	2			
2	F	64	Total	C	N	O	S	0	0	0
			528	335	91	100	2			
2	H	64	Total	C	N	O	S	0	0	0
			528	335	91	100	2			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	590	MET	-	initiating methionine	UNP P27588
B	591	GLY	-	expression tag	UNP P27588
B	592	HIS	-	expression tag	UNP P27588
B	593	HIS	-	expression tag	UNP P27588
B	594	HIS	-	expression tag	UNP P27588
B	595	HIS	-	expression tag	UNP P27588
B	596	HIS	-	expression tag	UNP P27588
B	597	HIS	-	expression tag	UNP P27588

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	598	GLY	-	expression tag	UNP P27588
B	599	GLY	-	expression tag	UNP P27588
B	600	GLY	-	expression tag	UNP P27588
D	590	MET	-	initiating methionine	UNP P27588
D	591	GLY	-	expression tag	UNP P27588
D	592	HIS	-	expression tag	UNP P27588
D	593	HIS	-	expression tag	UNP P27588
D	594	HIS	-	expression tag	UNP P27588
D	595	HIS	-	expression tag	UNP P27588
D	596	HIS	-	expression tag	UNP P27588
D	597	HIS	-	expression tag	UNP P27588
D	598	GLY	-	expression tag	UNP P27588
D	599	GLY	-	expression tag	UNP P27588
D	600	GLY	-	expression tag	UNP P27588
F	590	MET	-	initiating methionine	UNP P27588
F	591	GLY	-	expression tag	UNP P27588
F	592	HIS	-	expression tag	UNP P27588
F	593	HIS	-	expression tag	UNP P27588
F	594	HIS	-	expression tag	UNP P27588
F	595	HIS	-	expression tag	UNP P27588
F	596	HIS	-	expression tag	UNP P27588
F	597	HIS	-	expression tag	UNP P27588
F	598	GLY	-	expression tag	UNP P27588
F	599	GLY	-	expression tag	UNP P27588
F	600	GLY	-	expression tag	UNP P27588
H	590	MET	-	initiating methionine	UNP P27588
H	591	GLY	-	expression tag	UNP P27588
H	592	HIS	-	expression tag	UNP P27588
H	593	HIS	-	expression tag	UNP P27588
H	594	HIS	-	expression tag	UNP P27588
H	595	HIS	-	expression tag	UNP P27588
H	596	HIS	-	expression tag	UNP P27588
H	597	HIS	-	expression tag	UNP P27588
H	598	GLY	-	expression tag	UNP P27588
H	599	GLY	-	expression tag	UNP P27588
H	600	GLY	-	expression tag	UNP P27588

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

SO4

O1

O3 O⁻ — S — O⁻ O4

S

O2

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 5	O 4	S 1	0	0
3	A	1	Total 5	O 4	S 1	0	0
3	C	1	Total 5	O 4	S 1	0	0
3	C	1	Total 5	O 4	S 1	0	0
3	E	1	Total 5	O 4	S 1	0	0
3	E	1	Total 5	O 4	S 1	0	0
3	G	1	Total 5	O 4	S 1	0	0
3	G	1	Total 5	O 4	S 1	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: Anti-Marburgvirus Nucleoprotein Single Domain Antibody B

Chain A:
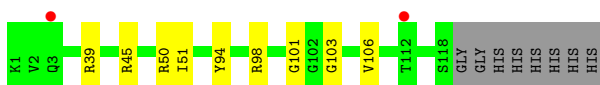
- Molecule 1: Anti-Marburgvirus Nucleoprotein Single Domain Antibody B

Chain C:

- Molecule 1: Anti-Marburgvirus Nucleoprotein Single Domain Antibody B
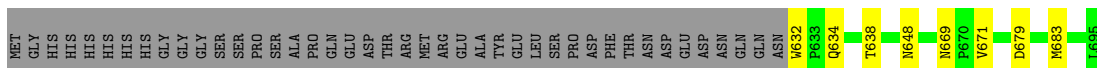
Chain E:

- Molecule 1: Anti-Marburgvirus Nucleoprotein Single Domain Antibody B

Chain G:

- Molecule 2: Nucleoprotein

Chain B:

- Molecule 2: Nucleoprotein

Chain D:

58% • 40%

MET GLY HIS HIS HIS HIS HIS HIS GLY GLY GLY SER SER PRO SER ALA PRO GLN GLU ASP THR ARG MET ARG GLU ALA TYR GLU LEU SER PRO ASP PHE THR ASN ASP GLU ASP ASN GLN GLN ASN W632 V671 D679 E687 L695

• Molecule 2: Nucleoprotein

Chain F:

60% 40%

MET GLY HIS HIS HIS HIS HIS HIS GLY GLY GLY SER SER PRO SER ALA PRO GLN GLU ASP THR ARG MET ARG GLU ALA TYR GLU LEU SER PRO ASP PHE THR ASN ASP GLU ASP ASN GLN GLN ASN W632 L695

• Molecule 2: Nucleoprotein

Chain H:

56% • • 40%

MET GLY HIS HIS HIS HIS HIS HIS GLY GLY GLY SER SER PRO SER ALA PRO GLN GLU ASP THR ARG MET ARG GLU ALA TYR GLU LEU SER PRO ASP PHE THR ASN ASP GLU ASP ASN GLN GLN ASN W632 K639 T643 S658 E675 E687 L695

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	58.00Å 108.66Å 141.27Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	54.33 – 3.20 54.33 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.7 (54.33-3.20) 100.0 (54.33-3.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.19	Depositor
$< I/\sigma(I) >$ [1]	2.62 (at 3.19Å)	Xtriage
Refinement program	PHENIX (1.10_2155)	Depositor
R, R_{free}	0.227 , 0.285 0.227 , 0.284	Depositor DCC
R_{free} test set	1534 reflections (10.01%)	wwPDB-VP
Wilson B-factor (Å^2)	55.1	Xtriage
Anisotropy	0.827	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 26.6	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	5752	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.06% 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	0.28	0/917	0.47	0/1236
1	C	0.28	0/917	0.48	0/1236
1	E	0.27	0/917	0.46	0/1236
1	G	0.27	0/917	0.47	0/1236
2	B	0.25	0/542	0.39	0/738
2	D	0.25	0/542	0.40	0/738
2	F	0.24	0/542	0.39	0/738
2	H	0.25	0/542	0.41	0/738
All	All	0.26	0/5836	0.44	0/7896

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	900	0	876	6	0
1	C	900	0	876	7	0
1	E	900	0	876	5	0
1	G	900	0	876	7	0
2	B	528	0	514	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	528	0	514	2	0
2	F	528	0	514	0	0
2	H	528	0	514	3	0
3	A	10	0	0	3	0
3	C	10	0	0	3	0
3	E	10	0	0	2	0
3	G	10	0	0	3	0
All	All	5752	0	5560	28	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 (28) 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:A:50:ARG:NH1	3:A:201:SO4:O2	2.16	0.78
1:G:50:ARG:NH2	3:G:201:SO4:O3	2.20	0.74
1:A:98:ARG:NH1	3:A:201:SO4:O4	2.25	0.68
1:C:50:ARG:NH1	3:C:201:SO4:O4	2.35	0.59
1:G:53:SER:OG	2:H:687:GLU:OE1	2.19	0.52
1:A:98:ARG:NH2	2:B:679:ASP:OD2	2.43	0.51
1:E:100:TRP:O	1:E:102:GLY:N	2.44	0.51
1:A:39:ARG:NH2	1:A:94:TYR:OH	2.44	0.51
1:A:50:ARG:NH1	3:A:201:SO4:S	2.84	0.51
1:C:53:SER:OG	2:D:687:GLU:OE1	2.20	0.49
1:G:50:ARG:NH2	3:G:201:SO4:S	2.86	0.49
1:C:98:ARG:NH1	3:C:201:SO4:O4	2.47	0.47
1:C:19:ARG:NE	1:C:81:ASP:OD1	2.39	0.47
1:C:98:ARG:NH2	2:D:679:ASP:OD2	2.48	0.46
1:G:30:SER:N	3:G:202:SO4:O1	2.48	0.46
1:E:98:ARG:NH2	3:E:201:SO4:O1	2.43	0.46
1:C:12:VAL:O	1:C:116:VAL:HA	2.16	0.46
2:H:643:THR:O	2:H:658:SER:OG	2.18	0.44
2:B:669:ASN:OD1	2:B:671:VAL:HG22	2.18	0.43
1:E:98:ARG:NH1	3:E:201:SO4:O2	2.52	0.43
1:G:98:ARG:NH1	2:H:675:GLU:OE1	2.52	0.42
2:B:683:MET:HB3	2:B:683:MET:HE2	1.97	0.42
1:E:29:PHE:O	1:E:71:ARG:NH2	2.53	0.41
2:B:648:ASN:ND2	1:G:54:GLY:O	2.52	0.41
1:C:50:ARG:NE	3:C:201:SO4:O1	2.53	0.41
1:A:50:ARG:HG2	1:A:51:ILE:N	2.34	0.41

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:72:ASP:OD2	1:E:75:LYS:HE2	2.21	0.41
1:G:100:TRP:O	1:G:102:GLY:N	2.54	0.41

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	116/126 (92%)	109 (94%)	5 (4%)	2 (2%)	9	42
1	C	116/126 (92%)	108 (93%)	7 (6%)	1 (1%)	17	56
1	E	116/126 (92%)	106 (91%)	9 (8%)	1 (1%)	17	56
1	G	116/126 (92%)	106 (91%)	8 (7%)	2 (2%)	9	42
2	B	62/106 (58%)	61 (98%)	1 (2%)	0	100	100
2	D	62/106 (58%)	62 (100%)	0	0	100	100
2	F	62/106 (58%)	61 (98%)	1 (2%)	0	100	100
2	H	62/106 (58%)	60 (97%)	2 (3%)	0	100	100
All	All	712/928 (77%)	673 (94%)	33 (5%)	6 (1%)	19	58

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	101	GLY
1	C	101	GLY
1	E	101	GLY
1	G	101	GLY
1	G	103	GLY
1	A	103	GLY

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	93/99 (94%)	91 (98%)	2 (2%)	52	79
1	C	93/99 (94%)	91 (98%)	2 (2%)	52	79
1	E	93/99 (94%)	91 (98%)	2 (2%)	52	79
1	G	93/99 (94%)	91 (98%)	2 (2%)	52	79
2	B	59/95 (62%)	56 (95%)	3 (5%)	24	60
2	D	59/95 (62%)	58 (98%)	1 (2%)	60	83
2	F	59/95 (62%)	59 (100%)	0	100	100
2	H	59/95 (62%)	57 (97%)	2 (3%)	37	70
All	All	608/776 (78%)	594 (98%)	14 (2%)	50	78

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	45	ARG
1	A	106	VAL
2	B	632	TRP
2	B	634	GLN
2	B	638	THR
1	C	45	ARG
1	C	106	VAL
2	D	671	VAL
1	E	45	ARG
1	E	110	GLN
1	G	1	LYS
1	G	45	ARG
2	H	639	LYS
2	H	675	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 ⓘ

8 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	$\#|Z| > 2$
3	SO4	A	202	-	4,4,4	0.14	0	6,6,6	0.05	0
3	SO4	E	201	-	4,4,4	0.13	0	6,6,6	0.09	0
3	SO4	C	201	-	4,4,4	0.15	0	6,6,6	0.06	0
3	SO4	G	201	-	4,4,4	0.16	0	6,6,6	0.04	0
3	SO4	A	201	-	4,4,4	0.15	0	6,6,6	0.05	0
3	SO4	C	202	-	4,4,4	0.13	0	6,6,6	0.06	0
3	SO4	G	202	-	4,4,4	0.14	0	6,6,6	0.08	0
3	SO4	E	202	-	4,4,4	0.14	0	6,6,6	0.07	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.

5 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	201	SO4	2	0
3	C	201	SO4	3	0
3	G	201	SO4	2	0
3	A	201	SO4	3	0
3	G	202	SO4	1	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	118/126 (93%)	0.30	2 (1%)	70	57	38, 52, 70, 80	0
1	C	118/126 (93%)	0.26	3 (2%)	57	43	36, 47, 65, 76	0
1	E	118/126 (93%)	0.08	0	100	100	33, 49, 69, 75	0
1	G	118/126 (93%)	0.35	1 (0%)	86	78	40, 56, 75, 86	0
2	B	64/106 (60%)	0.09	0	100	100	32, 43, 58, 74	0
2	D	64/106 (60%)	0.02	0	100	100	32, 48, 62, 73	0
2	F	64/106 (60%)	0.17	0	100	100	33, 47, 63, 78	0
2	H	64/106 (60%)	0.18	0	100	100	34, 54, 76, 85	0
All	All	728/928 (78%)	0.20	6 (0%)	86	78	32, 50, 71, 86	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	117	SER	2.7
1	A	112	THR	2.5
1	C	117	SER	2.2
1	C	116	VAL	2.1
1	C	13	GLN	2.1
1	A	3	GLN	2.0

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
3	SO4	G	201	5/5	0.74	0.37	72,74,77,86	0
3	SO4	E	201	5/5	0.78	0.26	60,62,73,86	0
3	SO4	A	201	5/5	0.82	0.27	61,64,68,73	0
3	SO4	A	202	5/5	0.86	0.26	61,62,70,80	0
3	SO4	C	201	5/5	0.86	0.40	65,71,80,92	0
3	SO4	C	202	5/5	0.93	0.22	55,65,83,85	0
3	SO4	G	202	5/5	0.94	0.16	40,52,60,61	0
3	SO4	E	202	5/5	0.95	0.21	58,63,73,78	0

6.5 Other polymers ⓘ

There are no such residues in this entry.