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

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 17, 2022 – 03:42 pm BST

PDB ID	:	7NME
Title	:	Human MHC Class I, A24 Allele presenting QLPRLFPLL, Complex with 4C6 TCR
Authors	:	Rizkallah, P.J.; Sewell, A.K.; Cole, D.K.; Wall, A.
Deposited on	:	2021-02-23
Resolution	:	2.20 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

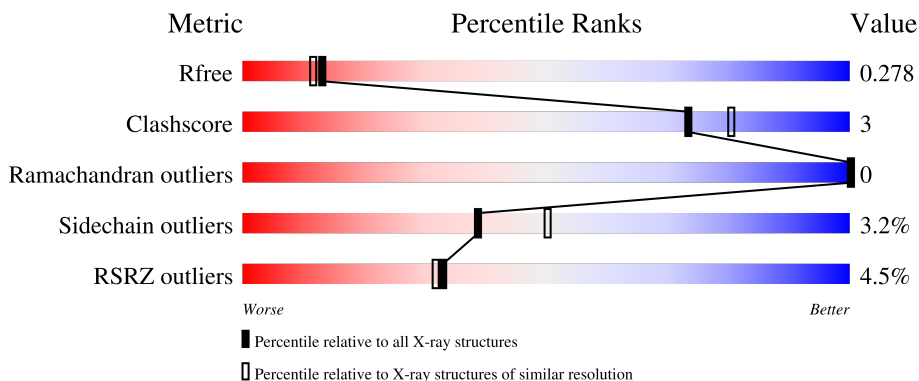|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 2.20 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		0.278
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		3.2%
RSRZ outliers		4.5%

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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	276	3% · 88% · 11% ·
2	B	100	2% · 93% · 7%
3	C	9	67% · 33%
4	D	192	8% · 85% · 14% ··
5	E	240	4% · 90% · 9%

WORLDWIDE wwPDB PROTEIN DATA BANK

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 6814 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 MHC class I antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	276	Total	C	N	O	S	0	0	0
			2238	1392	405	431	10			

- Molecule 2 is a protein called Human MHC Class I, beta 2 microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	100	Total	C	N	O	S	0	0	0
			837	533	141	159	4			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called GLN-LEU-PRO-ARG-LEU-PHE-PRO-LEU-LEU.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	0	0	0
			78	54	13	11			

- Molecule 4 is a protein called 4C6 Human T-cell Receptor, alpha Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	191	Total	C	N	O	S	0	0	0
			1495	932	246	308	9			

- Molecule 5 is a protein called 4C6 Human T-cell Receptor, beta Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	240	Total	C	N	O	S	0	0	0
			1943	1218	347	372	6			

- Molecule 6 is 2-[BIS-(2-HYDROXY-ETHYL)-AMINO]-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: BTB) (formula: $C_8H_{19}NO_5$).
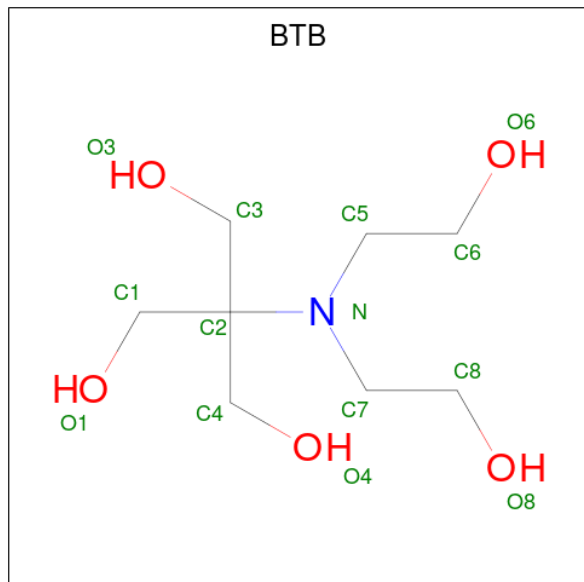
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
6	A	1	14	8	1	5	0	0

- Molecule 7 is SULFATE ION (three-letter code: SO4) (formula: O_4S).
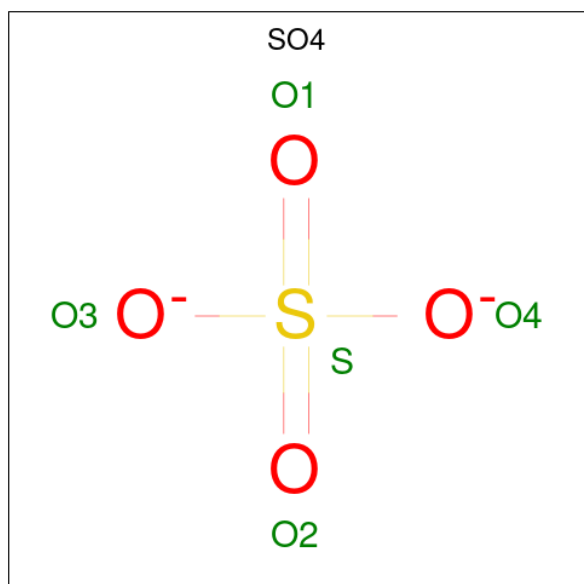
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
7	A	1	5	4	1	0	0

- Molecule 8 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).

EDO

O2

OH

C1

C2

HO

O1

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

- Molecule 9 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	78	Total 78	O 78	0	0
9	B	40	Total 40	O 40	0	0
9	C	4	Total 4	O 4	0	0
9	D	36	Total 36	O 36	0	0
9	E	42	Total 42	O 42	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.
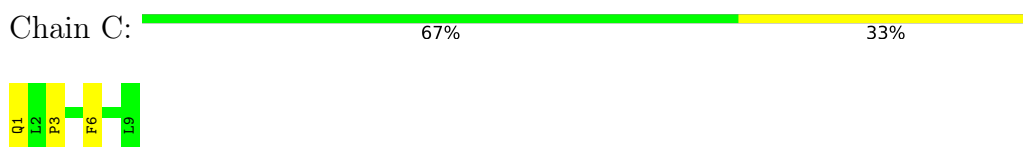
• Molecule 1: MHC class I antigen
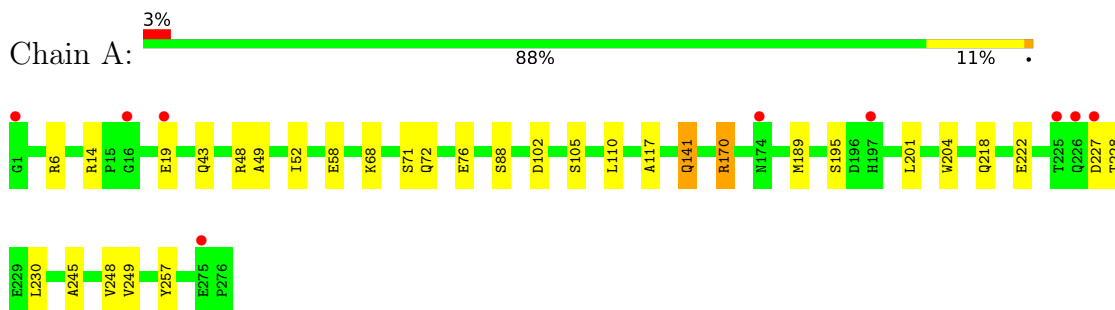
Chain A:

• Molecule 2: Human MHC Class I, beta 2 microglobulin
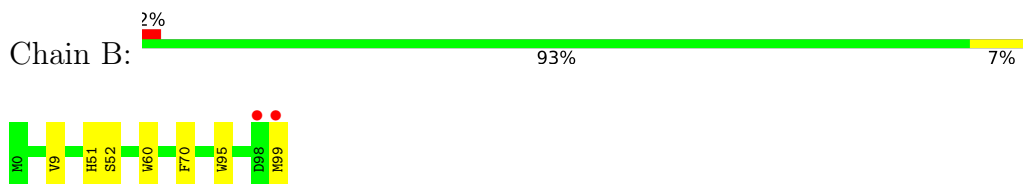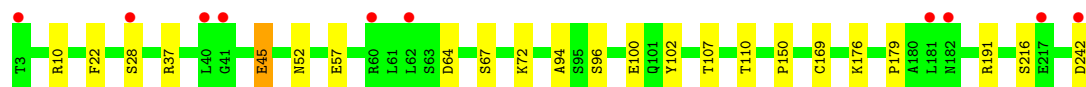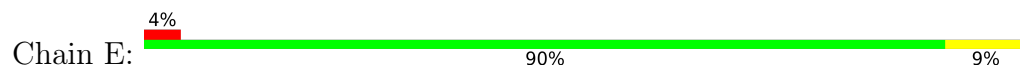
Chain B:

• Molecule 3: GLN-LEU-PRO-ARG-LEU-PHE-PRO-LEU-LEU

Chain C:

• Molecule 4: 4C6 Human T-cell Receptor, alpha Chain
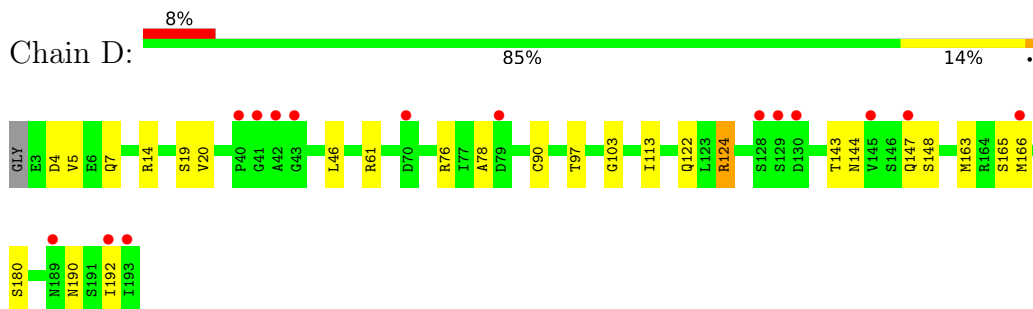
Chain D:

• Molecule 5: 4C6 Human T-cell Receptor, beta Chain

Chain E:

4%

90%

9%

T3 R10 F22 S28 R37 L40 G41 E45 N52 E57 R60 L61 L62 S63 D64 S67 K72 A94 S95 S96 E100 Q101 Y102 T107 T110 P150 C169 K176 P179 A180 L181 N182 R191 S216 E217 D242

4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	55.16Å 71.35Å 220.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	110.09 – 2.20 110.09 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.5 (110.09-2.20) 99.5 (110.09-2.20)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.48 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.215 , 0.274 0.220 , 0.278	Depositor DCC
R_{free} test set	2169 reflections (4.83%)	wwPDB-VP
Wilson B-factor (Å^2)	36.3	Xtriage
Anisotropy	0.060	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< |L| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6814	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 4.25% 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, EDO, BTB

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	0.67	0/2299	0.71	0/3116
2	B	0.64	0/860	0.70	0/1162
3	C	0.59	0/80	0.62	0/107
4	D	0.66	0/1521	0.70	0/2058
5	E	0.65	0/1996	0.69	0/2712
All	All	0.66	0/6756	0.70	0/9155

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
4	D	0	1
5	E	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	D	165	SER	Peptide
5	E	72	LYS	Peptide

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	2238	0	2095	17	0
2	B	837	0	803	4	0
3	C	78	0	90	2	0
4	D	1495	0	1436	13	0
5	E	1943	0	1838	9	0
6	A	14	0	19	3	0
7	A	5	0	0	0	0
8	A	4	0	6	0	0
9	A	78	0	0	1	0
9	B	40	0	0	0	0
9	C	4	0	0	0	0
9	D	36	0	0	0	0
9	E	42	0	0	1	0
All	All	6814	0	6287	42	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.

The worst 5 of 42 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)
4:D:163:MET:O	4:D:166:MET:O	1.93	0.86
4:D:14:ARG:HB3	4:D:113:ILE:HD11	1.70	0.71
4:D:7:GLN:HE21	4:D:103:GLY:HA3	1.56	0.71
5:E:22:PHE:CD1	5:E:107:THR:HG21	2.29	0.67
1:A:6:ARG:NH2	1:A:102:ASP:OD1	2.30	0.64

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	274/276 (99%)	263 (96%)	11 (4%)	0	100	100
2	B	98/100 (98%)	96 (98%)	2 (2%)	0	100	100
3	C	7/9 (78%)	7 (100%)	0	0	100	100
4	D	189/192 (98%)	182 (96%)	7 (4%)	0	100	100
5	E	238/240 (99%)	226 (95%)	12 (5%)	0	100	100
All	All	806/817 (99%)	774 (96%)	32 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	232/232 (100%)	224 (97%)	8 (3%)	37	47
2	B	95/95 (100%)	94 (99%)	1 (1%)	73	85
3	C	9/9 (100%)	9 (100%)	0	100	100
4	D	173/173 (100%)	167 (96%)	6 (4%)	36	46
5	E	212/212 (100%)	204 (96%)	8 (4%)	33	42
All	All	721/721 (100%)	698 (97%)	23 (3%)	39	50

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	180	SER
5	E	64	ASP
5	E	45	GLU
5	E	67	SER
1	A	195	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
4	D	178	ASN
4	D	147	GLN
4	D	38	GLN
4	D	142	GLN
4	D	22	ASN

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 ⓘ

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	$\#|Z| > 2$
8	EDO	A	303	-	3,3,3	0.25	0	2,2,2	0.23	0
6	BTB	A	301	-	13,13,13	1.16	3 (23%)	7,16,16	0.29	0
7	SO4	A	302	-	4,4,4	0.37	0	6,6,6	0.05	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
8	EDO	A	303	-	-	1/1/1/1	-
6	BTB	A	301	-	-	4/21/21/21	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	301	BTB	C2-N	2.49	1.53	1.48
6	A	301	BTB	C7-N	2.30	1.51	1.48
6	A	301	BTB	C5-N	2.07	1.51	1.48

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	301	BTB	C1-C2-C3-O3
6	A	301	BTB	C4-C2-C3-O3
6	A	301	BTB	N-C2-C3-O3
8	A	303	EDO	O1-C1-C2-O2
6	A	301	BTB	N-C7-C8-O8

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	301	BTB	3	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(Å²)	Q<0.9
1	A	276/276 (100%)	0.08	9 (3%)	46	44	21, 35, 61, 80	0
2	B	100/100 (100%)	0.04	2 (2%)	65	63	21, 34, 65, 82	0
3	C	9/9 (100%)	-0.10	0	100	100	23, 27, 40, 42	0
4	D	191/192 (99%)	0.34	16 (8%)	11	9	24, 43, 83, 99	0
5	E	240/240 (100%)	0.13	10 (4%)	36	34	22, 38, 65, 86	0
All	All	816/817 (99%)	0.15	37 (4%)	33	32	21, 38, 70, 99	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	128	SER	9.5
5	E	3	THR	6.2
5	E	41	GLY	5.5
4	D	41	GLY	5.1
4	D	192	ILE	5.0

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
8	EDO	A	303	4/4	0.82	0.21	41,47,50,51	0
7	SO4	A	302	5/5	0.85	0.27	88,98,102,108	0
6	BTB	A	301	14/14	0.86	0.19	37,44,52,54	0

6.5 Other polymers ⓘ

There are no such residues in this entry.