Full wwPDB X-ray Structure Validation Report ⓘ

Nov 21, 2023 – 01:18 PM JST

PDB ID	:	7FI9
Title	:	Crystal structure of human MICA mutants in complex with natural killer cell receptor NKG2D
Authors	:	Cai, W.; Peng, S.; Xu, T.; Tian, Y.; Li, Y.; Liu, J.
Deposited on	:	2021-07-30
Resolution	:	2.16 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 ⓘ 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 ⓘ) were used in the production of this report:

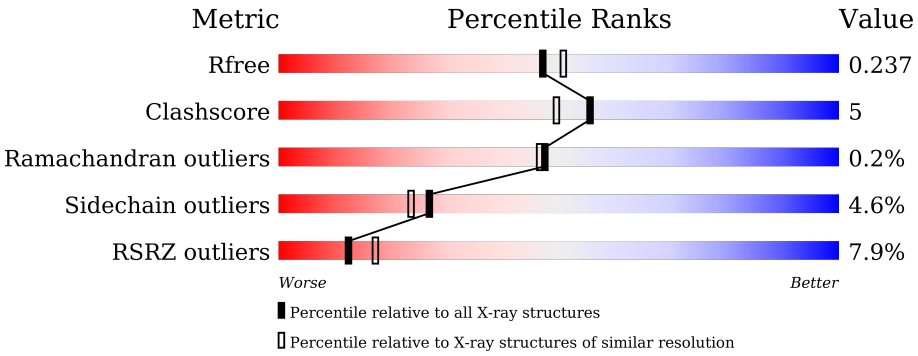
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.16 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	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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	139	6% 81% 7% • 11%
1	B	139	11% 87% 9% ••
2	C	275	7% 79% 14% • 5%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4398 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 NKG2-D type II integral membrane protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	124	Total	C	N	O	S	0	0	0
			1004	638	161	193	12			
1	B	136	Total	C	N	O	S	0	0	0
			1101	701	176	212	12			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	78	MET	-	initiating methionine	UNP P26718
A	79	GLU	-	expression tag	UNP P26718
B	78	MET	-	initiating methionine	UNP P26718
B	79	GLU	-	expression tag	UNP P26718

- Molecule 2 is a protein called MHC class I polypeptide-related sequence A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	260	Total	C	N	O	S	0	2	0
			2133	1324	392	401	16			

There are 11 discrepancies between the modelled and reference sequences:

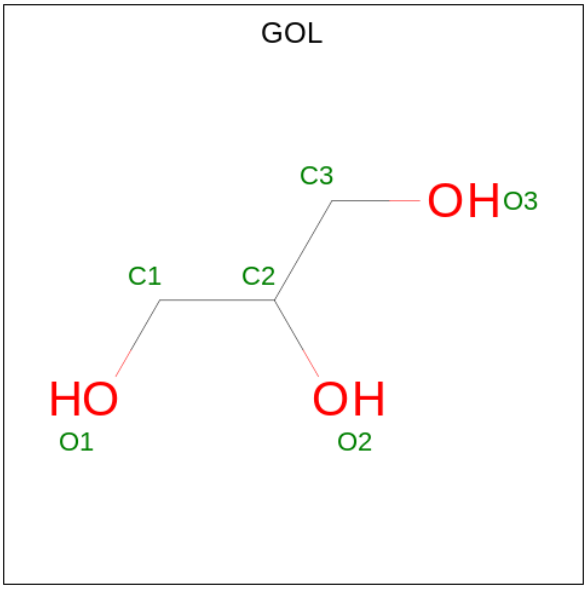
Chain	Residue	Modelled	Actual	Comment	Reference
C	0	MET	-	initiating methionine	UNP Q29983
C	8	ASP	ASN	engineered mutation	UNP Q29983
C	9	PHE	LEU	engineered mutation	UNP Q29983
C	34	ILE	LEU	engineered mutation	UNP Q29983
C	108	HIS	GLN	engineered mutation	UNP Q29983
C	120	TRP	GLN	engineered mutation	UNP Q29983
C	127	PHE	TRP	engineered mutation	UNP Q29983
C	146	TRP	LEU	engineered mutation	UNP Q29983
C	157	PHE	TYR	engineered mutation	UNP Q29983
C	161	ARG	HIS	engineered mutation	UNP Q29983

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	177	ILE	VAL	engineered mutation	UNP Q29983

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

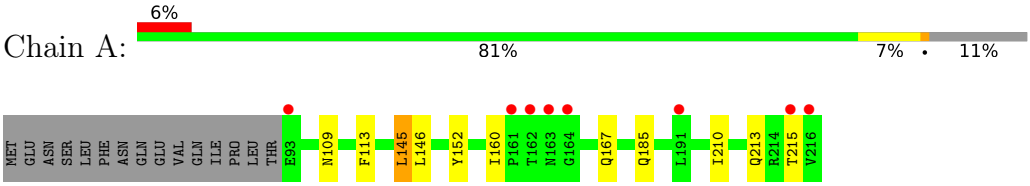
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	37	Total O 37 37	0	0
4	B	35	Total O 35 35	0	0
4	C	76	Total O 76 76	0	0

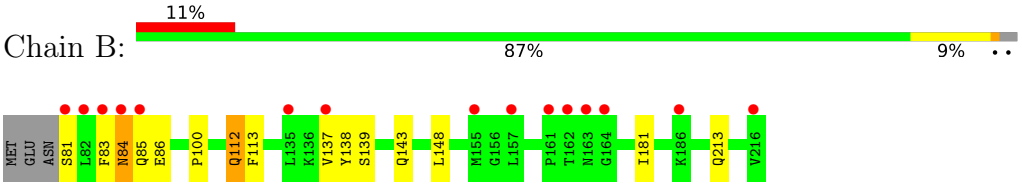
3 Residue-property plots (i)

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.

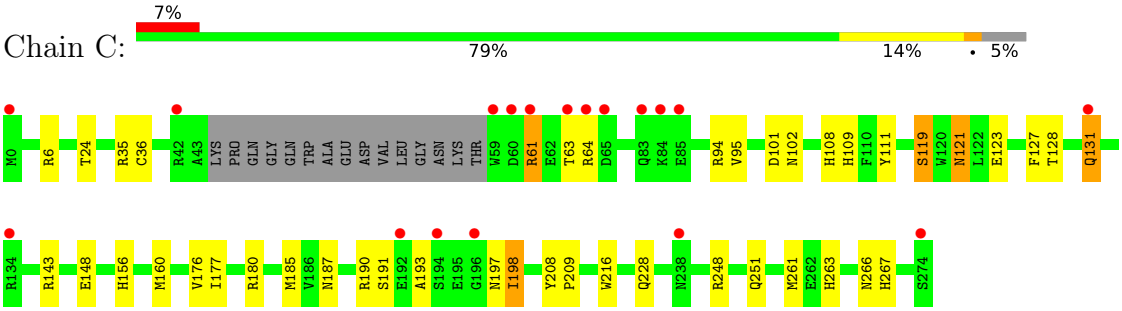
• Molecule 1: NKG2-D type II integral membrane protein



• Molecule 1: NKG2-D type II integral membrane protein



• Molecule 2: MHC class I polypeptide-related sequence A



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, α, β, γ	123.81Å 123.81Å 181.32Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.71 – 2.16 19.71 – 2.16	Depositor EDS
% Data completeness (in resolution range)	99.0 (19.71-2.16) 99.2 (19.71-2.16)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.48 (at 2.17Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.188 , 0.236 0.195 , 0.237	Depositor DCC
R_{free} test set	1891 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	44.7	Xtriage
Anisotropy	0.178	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 36.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.025 for -1/2*h-1/2*k-1/2*l,-1/2*h-1/2*k+1/2*l,-h+k 0.009 for -1/2*h+1/2*k-1/2*l,1/2*h-1/2*k-1/2*l,-h-k	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4398	wwPDB-VP
Average B, all atoms (Å^2)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.43% 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

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	0.71	0/1032	0.86	0/1398
1	B	0.68	0/1131	0.82	0/1534
2	C	0.69	0/2193	0.95	2/2965 (0.1%)
All	All	0.69	0/4356	0.90	2/5897 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	6	ARG	NE-CZ-NH2	-6.75	116.92	120.30
2	C	6	ARG	NE-CZ-NH1	6.72	123.66	120.30

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	1004	0	942	7	0
1	B	1101	0	1040	11	0
2	C	2133	0	2024	25	1
3	B	6	0	8	1	0
3	C	6	0	8	0	0
4	A	37	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	35	0	0	0	0
4	C	76	0	0	2	0
All	All	4398	0	4022	42	1

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.

All (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)
2:C:156:HIS:CD2	2:C:160[B]:MET:CE	2.68	0.76
2:C:156:HIS:CD2	2:C:160[B]:MET:HE2	2.22	0.74
1:B:138:TYR:H	1:B:143:GLN:HE22	1.35	0.71
1:B:138:TYR:H	1:B:143:GLN:NE2	1.91	0.69
1:B:112:GLN:HE21	1:B:113:PHE:H	1.43	0.66
1:A:213:GLN:O	4:A:301:HOH:O	2.15	0.64
1:A:185:GLN:O	4:A:302:HOH:O	2.15	0.63
2:C:193:ALA:HA	2:C:197:ASN:O	1.99	0.62
2:C:94:ARG:HD3	2:C:108:HIS:CE1	2.38	0.58
2:C:216:TRP:H	2:C:228:GLN:HE22	1.52	0.58
2:C:209:PRO:O	2:C:263:HIS:HE1	1.88	0.57
1:A:145:LEU:HD13	1:A:146:LEU:N	2.19	0.57
1:B:112:GLN:HE21	1:B:112:GLN:HA	1.72	0.55
2:C:191:SER:O	2:C:198:ILE:HG22	2.07	0.55
1:B:181:ILE:HD13	3:B:301:GOL:H2	1.90	0.54
2:C:156:HIS:NE2	2:C:160[B]:MET:SD	2.81	0.54
2:C:131:GLN:HE21	2:C:131:GLN:HA	1.74	0.53
1:A:113:PHE:CD2	1:B:148:LEU:HG	2.43	0.52
2:C:216:TRP:H	2:C:228:GLN:NE2	2.07	0.52
1:B:112:GLN:HA	1:B:112:GLN:NE2	2.25	0.51
2:C:127:PHE:HD2	2:C:143:ARG:HH11	1.57	0.51
2:C:156:HIS:CD2	2:C:160[B]:MET:HE1	2.47	0.50
1:B:139:SER:H	1:B:143:GLN:HE21	1.59	0.49
1:B:137:VAL:HA	1:B:143:GLN:HE22	1.78	0.48
1:B:100:PRO:HD3	1:B:213:GLN:NE2	2.28	0.48
2:C:101:ASP:O	2:C:102:ASN:HB2	2.14	0.47
1:A:160:ILE:HD12	1:A:167:GLN:NE2	2.30	0.47
2:C:127:PHE:HD2	2:C:143:ARG:NH1	2.13	0.46
2:C:261:MET:O	2:C:267:HIS:HA	2.16	0.45
2:C:121:ASN:C	2:C:121:ASN:HD22	2.19	0.45
1:A:113:PHE:CD2	1:A:210:ILE:HG12	2.52	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:109:ASN:OD1	1:A:215:THR:HG22	2.17	0.45
2:C:143:ARG:NH1	4:C:406:HOH:O	2.49	0.44
2:C:109:HIS:CD2	2:C:119:SER:HB3	2.53	0.43
2:C:193:ALA:HB1	2:C:251:GLN:OE1	2.19	0.42
2:C:94:ARG:HD2	4:C:461:HOH:O	2.19	0.42
2:C:121:ASN:HD21	2:C:123:GLU:HB2	1.84	0.42
2:C:185:MET:CE	2:C:187:ASN:HB2	2.49	0.42
2:C:61:ARG:H	2:C:61:ARG:HG2	1.53	0.42
2:C:208:TYR:CG	2:C:209:PRO:HA	2.54	0.41
2:C:24:THR:HB	2:C:36:CYS:HB2	2.03	0.41
1:B:81:SER:HB2	1:B:84:ASN:HA	2.03	0.41

All (1) 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)
2:C:111:TYR:OH	2:C:266:ASN:ND2[3_655]	2.14	0.06

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	122/139 (88%)	119 (98%)	3 (2%)	0	100 100
1	B	134/139 (96%)	129 (96%)	4 (3%)	1 (1%)	22 15
2	C	258/275 (94%)	246 (95%)	12 (5%)	0	100 100
All	All	514/553 (93%)	494 (96%)	19 (4%)	1 (0%)	47 46

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	84	ASN

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	114/129 (88%)	112 (98%)	2 (2%)	59	63
1	B	126/129 (98%)	122 (97%)	4 (3%)	39	38
2	C	236/246 (96%)	220 (93%)	16 (7%)	16	10
All	All	476/504 (94%)	454 (95%)	22 (5%)	27	23

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	145	LEU
1	A	152	TYR
1	B	83	PHE
1	B	85	GLN
1	B	86	GLU
1	B	112	GLN
2	C	35	ARG
2	C	61	ARG
2	C	63	THR
2	C	64	ARG
2	C	95	VAL
2	C	119	SER
2	C	121	ASN
2	C	128	THR
2	C	131	GLN
2	C	148	GLU
2	C	176	VAL
2	C	177	ILE
2	C	180	ARG
2	C	190	ARG
2	C	198	ILE
2	C	248	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	167	GLN
1	B	85	GLN
1	B	112	GLN
1	B	130	GLN
1	B	143	GLN
1	B	213	GLN
2	C	102	ASN
2	C	109	HIS
2	C	121	ASN
2	C	131	GLN
2	C	156	HIS
2	C	166	GLN
2	C	197	ASN
2	C	228	GLN
2	C	263	HIS
2	C	267	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

2 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	GOL	C	301	-	5,5,5	0.15	0	5,5,5	0.47	0
3	GOL	B	301	-	5,5,5	0.19	0	5,5,5	0.34	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	C	301	-	-	3/4/4/4	-
3	GOL	B	301	-	-	2/4/4/4	-

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
3	B	301	GOL	C1-C2-C3-O3
3	C	301	GOL	O1-C1-C2-C3
3	C	301	GOL	C1-C2-C3-O3
3	B	301	GOL	O2-C2-C3-O3
3	C	301	GOL	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	301	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	124/139 (89%)	0.21	8 (6%) 18 25	34, 45, 87, 159	0
1	B	136/139 (97%)	0.37	15 (11%) 5 8	38, 52, 116, 136	0
2	C	260/275 (94%)	0.16	18 (6%) 16 23	35, 49, 92, 146	0
All	All	520/553 (94%)	0.23	41 (7%) 12 17	34, 49, 97, 159	0

All (41) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	216	VAL	7.7
2	C	59	TRP	6.6
1	A	162	THR	6.1
1	B	81	SER	6.0
2	C	64	ARG	5.5
1	B	82	LEU	5.5
2	C	60	ASP	5.1
1	A	163	ASN	5.1
1	B	216	VAL	4.6
2	C	274	SER	4.5
2	C	83	GLN	4.3
1	B	163	ASN	4.1
1	B	161	PRO	3.9
1	B	84	ASN	3.8
2	C	192	GLU	3.8
2	C	42	ARG	3.7
1	B	164	GLY	3.6
1	A	215	THR	3.6
1	A	161	PRO	3.6
2	C	61	ARG	3.5
1	B	85	GLN	3.3
1	B	83	PHE	3.2
1	A	164	GLY	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	C	196	GLY	3.1
2	C	84	LYS	3.0
2	C	238	ASN	3.0
1	B	186	LYS	3.0
2	C	131	GLN	2.9
1	A	191	LEU	2.9
1	B	162	THR	2.8
2	C	194	SER	2.8
1	B	135	LEU	2.6
2	C	65	ASP	2.6
2	C	0	MET	2.5
1	B	157	LEU	2.4
1	A	93	GLU	2.3
1	B	155	MET	2.2
2	C	134	ARG	2.2
2	C	85	GLU	2.2
2	C	63	THR	2.1
1	B	137	VAL	2.0

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 monosaccharides 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
3	GOL	B	301	6/6	0.55	0.35	76,81,85,87	0
3	GOL	C	301	6/6	0.80	0.26	74,82,85,87	0

6.5 Other polymers (i)

There are no such residues in this entry.