Full wwPDB X-ray Structure Validation Report ⓘ

Oct 15, 2023 – 11:05 AM EDT

PDB ID	:	8CYH
Title	:	Novel Anti-Mesothelin Antibodies Enable Crystallography of the Intact Mesothelin Ectodo- main and Engineering of Potent, T cell-engaging Bispecific Therapeutics
Authors	:	Bandaranayake, A.D.; Rupert, P.B.; Lin, I.; Pilat, K.; Ruff, R.O.; Friend, D.J.; Chan, M.K.; Clarke, M.; Carter, J.; Meshinchi, S.; Mehlin, C.; Olson, J.M.; Strong, R.K.; Correnti, C.E.
Deposited on	:	2022-05-23
Resolution	:	3.38 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:

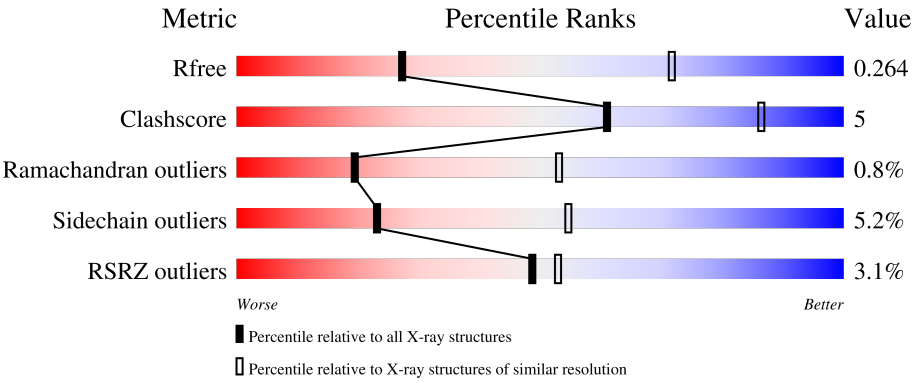
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA,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

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 3.38 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	1691 (3.46-3.30)
Clashscore	141614	1762 (3.46-3.30)
Ramachandran outliers	138981	1732 (3.46-3.30)
Sidechain outliers	138945	1731 (3.46-3.30)
RSRZ outliers	127900	1635 (3.46-3.30)

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	L	214	2% 80% 18%
2	H	222	% 86% 13%
3	M	327	5% 77% 10% 13%

Validation Pipeline (wwPDB-VP) : 2.36

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5244 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 A12 antibody light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	L	212	Total	C	N	O	S	0	0	0
			1555	974	253	323	5			

- Molecule 2 is a protein called A12 antibody heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	220	Total	C	N	O	S	0	0	0
			1560	986	263	305	6			

- Molecule 3 is a protein called Mesothelin, cleaved form.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	M	285	Total	C	N	O	S	0	0	0
			2101	1350	337	405	9			

There are 17 discrepancies between the modelled and reference sequences:

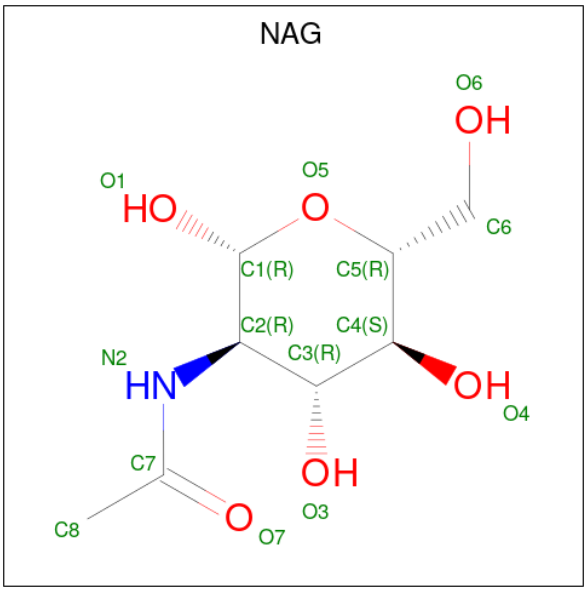
Chain	Residue	Modelled	Actual	Comment	Reference
M	606	GLY	-	expression tag	UNP Q13421
M	607	SER	-	expression tag	UNP Q13421
M	608	GLY	-	expression tag	UNP Q13421
M	609	LEU	-	expression tag	UNP Q13421
M	610	ASN	-	expression tag	UNP Q13421
M	611	ASP	-	expression tag	UNP Q13421
M	612	ILE	-	expression tag	UNP Q13421
M	613	PHE	-	expression tag	UNP Q13421
M	614	GLU	-	expression tag	UNP Q13421
M	615	ALA	-	expression tag	UNP Q13421
M	616	GLN	-	expression tag	UNP Q13421
M	617	LYS	-	expression tag	UNP Q13421
M	618	ILE	-	expression tag	UNP Q13421
M	619	GLU	-	expression tag	UNP Q13421

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
M	620	TRP	-	expression tag	UNP Q13421
M	621	HIS	-	expression tag	UNP Q13421
M	622	GLU	-	expression tag	UNP Q13421

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



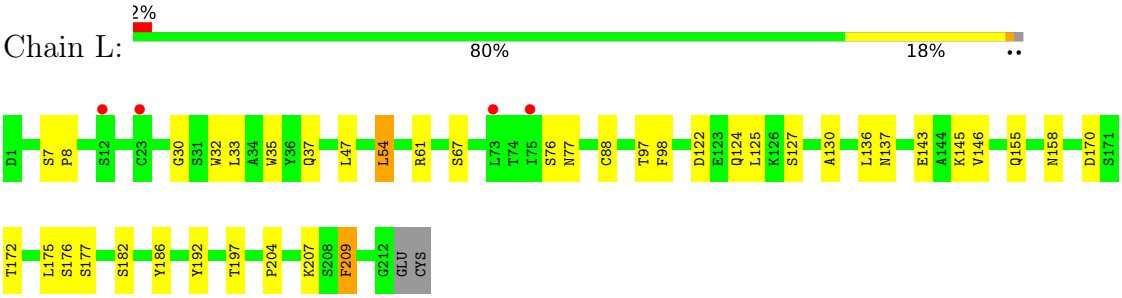
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	M	1	Total	C	N	O	0	0
			14	8	1	5		
4	M	1	Total	C	N	O	0	0
			14	8	1	5		

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.

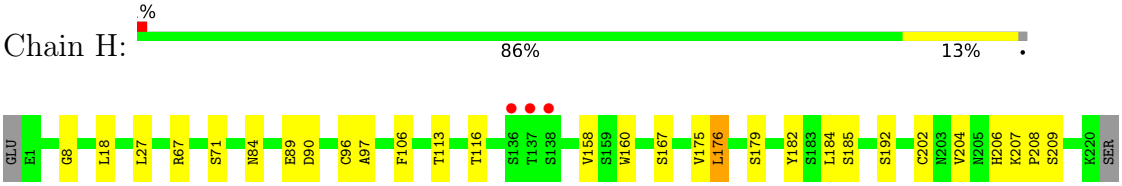
• Molecule 1: A12 antibody light chain

Chain L:



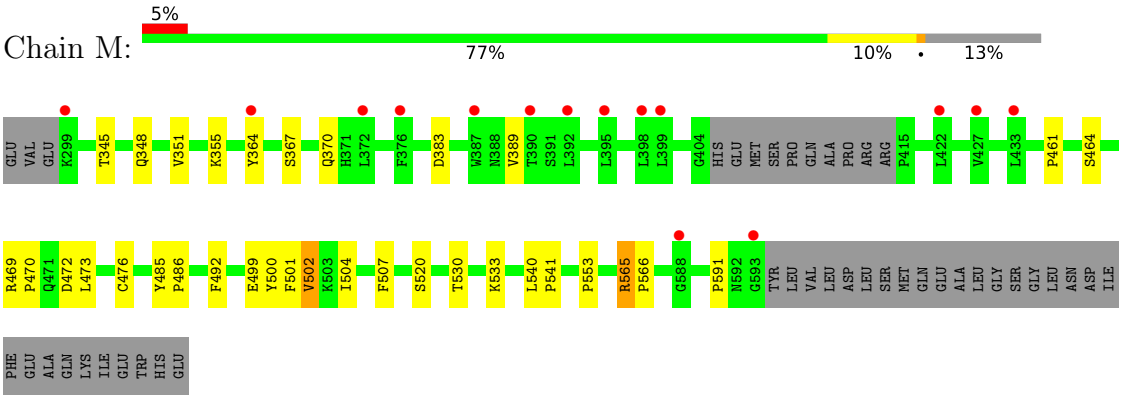
• Molecule 2: A12 antibody heavy chain

Chain H:



• Molecule 3: Mesothelin, cleaved form

Chain M:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	65.11Å 124.86Å 302.75Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.38 – 3.38 44.34 – 3.38	Depositor EDS
% Data completeness (in resolution range)	97.9 (44.38-3.38) 98.0 (44.34-3.38)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.58 (at 3.40Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.230 , 0.268 0.227 , 0.264	Depositor DCC
R_{free} test set	851 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å^2)	105.1	Xtriage
Anisotropy	0.503	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 89.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	5244	wwPDB-VP
Average B, all atoms (Å^2)	133.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.06% 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: NAG

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	L	0.67	0/1591	0.70	0/2177
2	H	0.71	0/1596	0.72	0/2183
3	M	0.70	0/2146	0.67	0/2940
All	All	0.70	0/5333	0.69	0/7300

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	L	1555	0	1419	17	0
2	H	1560	0	1471	12	0
3	M	2101	0	1978	18	0
4	M	28	0	26	0	0
All	All	5244	0	4894	46	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 5.

All (46) 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)
3:M:565:ARG:CB	3:M:566:PRO:HD3	2.05	0.86
3:M:565:ARG:CB	3:M:566:PRO:CD	2.76	0.64
1:L:155:GLN:HB3	1:L:158:ASN:HD21	1.65	0.60
2:H:176:LEU:HD12	2:H:182:TYR:CD1	2.41	0.55
3:M:351:VAL:O	3:M:355:LYS:HG2	2.08	0.54
2:H:97:ALA:HB1	2:H:106:PHE:HB3	1.92	0.52
1:L:170:ASP:CB	1:L:172:THR:HG22	2.40	0.52
1:L:145:LYS:HB3	1:L:197:THR:OG1	2.10	0.52
3:M:540:LEU:HB2	3:M:541:PRO:HD3	1.92	0.51
2:H:67:ARG:NH2	2:H:90:ASP:OD2	2.45	0.50
1:L:32:TRP:CE3	3:M:553:PRO:HB3	2.47	0.49
1:L:136:LEU:HB2	1:L:175:LEU:HB3	1.95	0.49
1:L:125:LEU:HD23	1:L:130:ALA:HB2	1.97	0.47
3:M:485:TYR:HB3	3:M:486:PRO:HD3	1.96	0.47
3:M:499:GLU:O	3:M:502:VAL:HG22	2.13	0.47
3:M:499:GLU:O	3:M:502:VAL:CG2	2.63	0.47
2:H:207:LYS:N	2:H:208:PRO:CD	2.78	0.47
2:H:184:LEU:HD23	2:H:185:SER:N	2.30	0.47
2:H:8:GLY:O	2:H:18:LEU:HD11	2.15	0.46
3:M:461:PRO:HG2	3:M:464:SER:HB2	1.97	0.46
1:L:37:GLN:HB2	1:L:47:LEU:HD11	1.98	0.45
3:M:530:THR:HA	3:M:533:LYS:HD3	1.99	0.45
1:L:124:GLN:O	1:L:127:SER:OG	2.29	0.45
2:H:176:LEU:HD12	2:H:182:TYR:CG	2.52	0.45
1:L:192:TYR:CD1	1:L:192:TYR:N	2.84	0.45
1:L:54:LEU:HD23	1:L:54:LEU:HA	1.82	0.44
3:M:473:LEU:HD11	3:M:507:PHE:CE1	2.53	0.44
1:L:192:TYR:HB2	1:L:209:PHE:CE1	2.53	0.43
1:L:61:ARG:HB2	1:L:76:SER:O	2.18	0.42
3:M:345:THR:OG1	3:M:348:GLN:HG3	2.19	0.42
3:M:367:SER:O	3:M:370:GLN:HB2	2.20	0.42
1:L:97:THR:HG22	1:L:98:PHE:O	2.20	0.41
2:H:160:TRP:CZ3	2:H:202:CYS:HB3	2.55	0.41
3:M:500:TYR:CE1	3:M:504:ILE:HG21	2.56	0.41
1:L:7:SER:HA	1:L:8:PRO:HA	1.89	0.41
2:H:176:LEU:HD12	2:H:182:TYR:CE1	2.55	0.41
2:H:206:HIS:ND1	2:H:209:SER:OG	2.32	0.41
1:L:35:TRP:CZ3	1:L:88:CYS:HB3	2.56	0.41
3:M:501:PHE:O	3:M:502:VAL:C	2.59	0.41
1:L:186:TYR:HA	1:L:192:TYR:OH	2.21	0.41
1:L:146:VAL:HG11	1:L:177:SER:CB	2.51	0.41
3:M:469:ARG:HB3	3:M:470:PRO:HD2	2.02	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:207:LYS:N	2:H:208:PRO:HD2	2.35	0.41
3:M:501:PHE:HA	3:M:504:ILE:HG22	2.03	0.41
2:H:158:VAL:HG22	2:H:204:VAL:HG22	2.04	0.40
3:M:492:PHE:O	3:M:500:TYR:CD2	2.75	0.40

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	L	210/214 (98%)	184 (88%)	22 (10%)	4 (2%)	8 34
2	H	218/222 (98%)	191 (88%)	26 (12%)	1 (0%)	29 63
3	M	281/327 (86%)	246 (88%)	34 (12%)	1 (0%)	34 68
All	All	709/763 (93%)	621 (88%)	82 (12%)	6 (1%)	19 53

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	M	565	ARG
1	L	143	GLU
1	L	30	GLY
1	L	77	ASN
1	L	204	PRO
2	H	175	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	L	167/187 (89%)	158 (95%)	9 (5%)	22	53
2	H	160/182 (88%)	149 (93%)	11 (7%)	15	46
3	M	213/286 (74%)	205 (96%)	8 (4%)	33	62
All	All	540/655 (82%)	512 (95%)	28 (5%)	23	54

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	33	LEU
1	L	54	LEU
1	L	67	SER
1	L	122	ASP
1	L	137	ASN
1	L	176	SER
1	L	182	SER
1	L	207	LYS
1	L	209	PHE
2	H	27	LEU
2	H	71	SER
2	H	84	ASN
2	H	89	GLU
2	H	96	CYS
2	H	113	THR
2	H	116	THR
2	H	167	SER
2	H	176	LEU
2	H	179	SER
2	H	192	SER
3	M	364	TYR
3	M	383	ASP
3	M	389	VAL
3	M	472	ASP
3	M	476	CYS
3	M	502	VAL
3	M	520	SER
3	M	591	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	L	55	GLN
1	L	92	ASN
2	H	82	GLN
2	H	84	ASN
3	M	522	GLN

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	$\#\|Z\| > 2$
4	NAG	M	701	3	14,14,15	0.50	0	17,19,21	0.90	0
4	NAG	M	702	3	14,14,15	0.75	1 (7%)	17,19,21	0.65	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
4	NAG	M	701	3	-	1/6/23/26	0/1/1/1
4	NAG	M	702	3	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	M	702	NAG	C1-C2	2.40	1.55	1.52

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	M	701	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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	L	212/214 (99%)	0.23	4 (1%)	66	71	85, 113, 138, 159	0
2	H	220/222 (99%)	0.18	3 (1%)	75	79	81, 116, 158, 252	0
3	M	285/327 (87%)	0.24	15 (5%)	26	30	115, 159, 183, 202	0
All	All	717/763 (93%)	0.22	22 (3%)	49	53	81, 129, 178, 252	0

All (22) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	137	THR	4.0
3	M	433	LEU	3.3
3	M	299	LYS	2.9
3	M	387	TRP	2.9
1	L	75	ILE	2.7
3	M	593	GLY	2.7
3	M	364	TYR	2.6
1	L	12	SER	2.6
3	M	422	LEU	2.5
2	H	136	SER	2.4
3	M	395	LEU	2.4
3	M	376	PHE	2.3
3	M	399	LEU	2.3
3	M	372	LEU	2.2
3	M	588	GLY	2.2
3	M	398	LEU	2.2
1	L	23	CYS	2.1
3	M	390	THR	2.1
3	M	427	VAL	2.0
2	H	138	SER	2.0
3	M	392	LEU	2.0
1	L	73	LEU	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
4	NAG	M	702	14/15	0.59	0.33	158,205,218,221	0
4	NAG	M	701	14/15	0.80	0.27	163,202,214,223	0

6.5 Other polymers (i)

There are no such residues in this entry.