



# Full wwPDB X-ray Structure Validation Report

Oct 13, 2014 – 10:39 AM EDT

PDB ID : 4N8C  
Title : Three-dimensional structure of the extracellular domain of Matrix protein 2 of influenza A virus  
Authors : Cho, K.J.; Seok, J.H.; Kim, S.; Roose, K.; Schepens, B.; Fiers, W.; Saelens, X.; Kim, K.H.  
Deposited on : 2013-10-17  
Resolution : 1.60 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

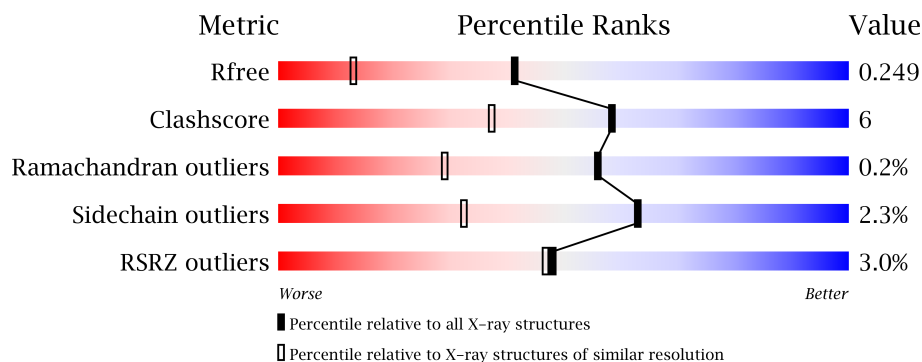
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1439  
EDS : stable23828  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable23828

# 1 Overall quality at a glance

The reported resolution of this entry is 1.60 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))
$R_{free}$	66092	1872 (1.60-1.60)
Clashscore	79885	2199 (1.60-1.60)
Ramachandran outliers	78287	2126 (1.60-1.60)
Sidechain outliers	78261	2125 (1.60-1.60)
RSRZ outliers	66119	1872 (1.60-1.60)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	H	221	
1	I	221	
2	L	220	
2	M	220	
3	X	23	
3	Y	23	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7783 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Heavy chain of monoclonal antibody.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	220	Total	C	N	O	S	0	1	0
			1675	1051	284	331	9			
1	I	221	Total	C	N	O	S	0	6	0
			1712	1071	290	342	9			

- Molecule 2 is a protein called Light chain of monoclonal antibody.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	219	Total	C	N	O	S	0	3	0
			1728	1085	282	351	10			
2	M	220	Total	C	N	O	S	0	4	0
			1739	1091	284	354	10			

- Molecule 3 is a protein called Extracellular domain of influenza Matrix protein 2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	X	15	Total	C	N	O	0	0	0
			122	77	20	25			
3	Y	14	Total	C	N	O	0	0	0
			116	74	19	23			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	H	138	Total	O	0	0
			138	138		
4	I	192	Total	O	0	0
			192	192		
4	L	142	Total	O	0	0
			142	142		
4	M	197	Total	O	0	0
			197	197		

*Continued on next page...*

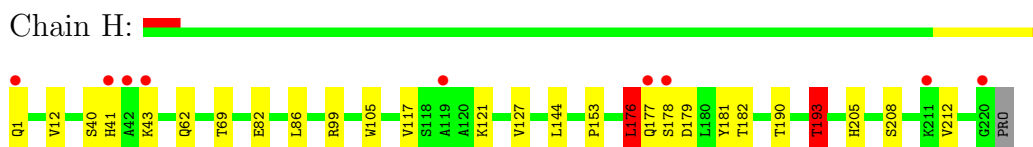
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	X	11	Total 11	O 11	0	0
4	Y	11	Total 11	O 11	0	0

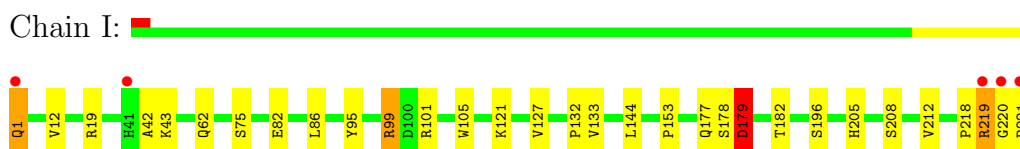
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

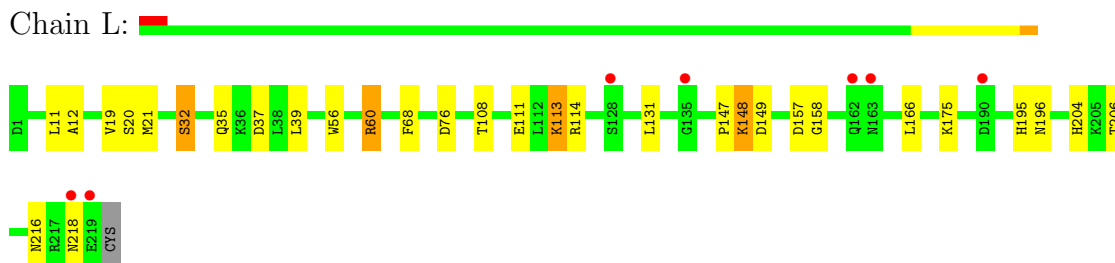
- Molecule 1: Heavy chain of monoclonal antibody



- Molecule 1: Heavy chain of monoclonal antibody



- Molecule 2: Light chain of monoclonal antibody



- Molecule 2: Light chain of monoclonal antibody

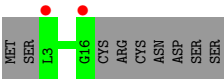


- Molecule 3: Extracellular domain of influenza Matrix protein 2



- Molecule 3: Extracellular domain of influenza Matrix protein 2





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	40.35Å 72.28Å 78.45Å 86.88° 77.53° 84.63°	Depositor
Resolution (Å)	38.43 – 1.60 35.75 – 1.60	Depositor EDS
% Data completeness (in resolution range)	96.6 (38.43-1.60) 96.6 (35.75-1.60)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.82 (at 1.60Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.193 , 0.242 0.204 , 0.249	Depositor DCC
$R_{free}$ test set	5494 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.0	Xtriage
Anisotropy	0.071	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 36.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 109542 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7783	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.40% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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	H	0.89	0/1716	0.94	2/2342 (0.1%)
1	I	1.02	0/1754	1.03	3/2394 (0.1%)
2	L	0.90	1/1768 (0.1%)	1.02	3/2398 (0.1%)
2	M	1.06	0/1779	1.06	1/2414 (0.0%)
3	X	0.84	0/124	0.95	0/169
3	Y	1.15	0/118	1.00	0/161
All	All	0.97	1/7259 (0.0%)	1.01	9/9878 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	L	20	SER	CB-OG	5.34	1.49	1.42

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	L	60	ARG	NE-CZ-NH2	-11.46	114.57	120.30
2	L	60	ARG	NE-CZ-NH1	9.98	125.29	120.30
2	M	76	ASP	CB-CG-OD1	9.35	126.72	118.30
1	I	179	ASP	CB-CG-OD1	6.83	124.45	118.30
1	I	95	TYR	CB-CG-CD1	5.95	124.57	121.00
1	I	99	ARG	NE-CZ-NH2	-5.81	117.39	120.30
1	H	176	LEU	CA-CB-CG	5.75	128.53	115.30
2	L	76	ASP	CB-CG-OD2	-5.57	113.29	118.30
1	H	193	THR	N-CA-CB	-5.52	99.81	110.30

There are no chirality outliers.

There are no planarity outliers.



## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	H	1675	0	1627	21	0
1	I	1712	0	1655	35	0
2	L	1728	0	1658	16	0
2	M	1739	0	1667	18	0
3	X	122	0	117	1	0
3	Y	116	0	112	0	0
4	H	138	0	0	4	0
4	I	192	0	0	4	0
4	L	142	0	0	3	0
4	M	197	0	0	5	0
4	X	11	0	0	0	0
4	Y	11	0	0	0	0
All	All	7783	0	6836	86	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 6.

All (86) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:L:35:GLN:HA	4:L:354:HOH:O	1.13	1.26
1:I:75:SER:HB2	4:I:461:HOH:O	1.32	1.22
1:H:12:VAL:HB	4:H:399:HOH:O	1.45	1.13
1:I:221:PRO:HB2	2:M:125:PRO:HB2	1.40	1.02
1:H:40:SER:HB2	1:H:43:LYS:HD3	1.50	0.93
1:I:219:ARG:HH11	1:I:219:ARG:HG3	1.41	0.85
1:I:221:PRO:HB3	2:M:126:PRO:O	1.76	0.83
1:I:177:GLN:HE21	1:I:182:THR:HG21	1.43	0.81
1:H:40:SER:CB	1:H:43:LYS:HD3	2.17	0.75
1:I:19:ARG:HB2	1:I:82:GLU:OE2	1.89	0.72
1:I:219:ARG:HG3	1:I:219:ARG:NH1	2.02	0.71
1:I:221:PRO:HB2	2:M:125:PRO:CB	2.19	0.71
1:I:132:PRO:HD3	1:I:144:LEU:CD2	2.20	0.71
1:I:43:LYS:HD3	2:M:91:VAL:HG21	1.76	0.67
1:H:41:HIS:HB3	4:H:395:HOH:O	1.96	0.66

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:L:204:HIS:HD2	2:L:206:THR:OG1	1.81	0.64
1:H:177:GLN:HE21	1:H:182:THR:HG21	1.63	0.64
2:M:51:LYS:HE2	4:M:324:HOH:O	1.99	0.63
1:I:153:PRO:O	1:I:205:HIS:HE1	1.82	0.62
1:H:153:PRO:O	1:H:205:HIS:HE1	1.82	0.61
1:H:178:SER:OG	1:H:179:ASP:N	2.34	0.61
1:I:218:PRO:O	1:I:219:ARG:CG	2.49	0.60
1:I:218:PRO:O	1:I:219:ARG:CB	2.51	0.59
1:I:178:SER:OG	1:I:179:ASP:N	2.35	0.58
1:H:205:HIS:HD2	1:H:208:SER:OG	1.87	0.57
1:I:220:GLY:O	1:I:221:PRO:C	2.42	0.57
2:M:60:ARG:HD2	4:M:468:HOH:O	2.04	0.57
1:H:40:SER:O	1:H:43:LYS:HB2	2.04	0.57
1:H:177:GLN:NE2	1:H:182:THR:HG21	2.21	0.56
2:M:204:HIS:HD2	2:M:206:THR:OG1	1.87	0.56
1:I:177:GLN:NE2	1:I:182:THR:HG21	2.18	0.56
2:L:12:ALA:HB2	2:L:111:GLU:OE2	2.07	0.55
1:I:132:PRO:HD3	1:I:144:LEU:HD22	1.88	0.54
1:H:117:VAL:HG22	4:H:399:HOH:O	2.07	0.53
1:I:205:HIS:HD2	1:I:208:SER:OG	1.91	0.53
1:I:132:PRO:HD3	1:I:144:LEU:HD23	1.92	0.52
1:I:12:VAL:HG11	1:I:86:LEU:HD12	1.92	0.52
2:M:218:ASN:OD1	2:M:219:GLU:HG3	2.10	0.52
2:M:142:LEU:HD21	2:M:152[B]:VAL:HG22	1.91	0.51
1:H:176:LEU:HD13	1:H:181:TYR:CE1	2.47	0.50
1:I:218:PRO:O	1:I:219:ARG:HB2	2.12	0.50
2:M:60:ARG:NE	2:M:68:PHE:O	2.45	0.50
2:L:157:ASP:OD2	2:L:195:HIS:ND1	2.42	0.50
1:H:12:VAL:HG11	1:H:86:LEU:HD12	1.94	0.49
2:L:148:LYS:HG3	2:L:149:ASP:N	2.28	0.49
1:I:101:ARG:CG	4:I:431:HOH:O	2.61	0.48
1:H:99:ARG:HA	1:H:105:TRP:O	2.14	0.48
2:L:60:ARG:HD3	2:L:68:PHE:O	2.14	0.47
2:M:51:LYS:HD3	4:M:377:HOH:O	2.13	0.47
2:L:147:PRO:O	2:L:204:HIS:HE1	1.98	0.47
1:I:1:GLN:HB2	1:I:1:GLN:HE21	1.61	0.47
2:L:196:ASN:OD1	2:L:218:ASN:ND2	2.46	0.47
2:M:189:LYS:O	2:M:193:GLU:HG2	2.15	0.46
1:I:62:GLN:HG2	4:I:432:HOH:O	2.16	0.46
1:H:127:VAL:HG21	1:H:212:VAL:CG2	2.46	0.45
1:I:99:ARG:HA	1:I:105:TRP:O	2.17	0.45
2:M:147:PRO:O	2:M:204:HIS:HE1	2.00	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:I:42:ALA:O	1:I:43:LYS:CB	2.65	0.45
2:L:196:ASN:O	2:L:216:ASN:HA	2.17	0.44
1:I:127:VAL:HG21	1:I:212:VAL:CG2	2.47	0.44
1:I:42:ALA:O	1:I:43:LYS:HB3	2.18	0.44
1:H:69:THR:HB	1:H:82:GLU:HB2	2.00	0.44
1:H:190:THR:O	1:H:193:THR:HB	2.17	0.43
1:I:205:HIS:CD2	1:I:208:SER:OG	2.71	0.43
1:I:133:VAL:HG12	1:I:220:GLY:HA3	2.00	0.43
2:M:51:LYS:CE	4:M:324:HOH:O	2.62	0.43
1:I:177:GLN:HG2	2:M:166:LEU:HD12	2.01	0.43
2:M:205:LYS:HB2	2:M:205:LYS:HE2	1.71	0.42
1:I:101:ARG:HG3	4:I:431:HOH:O	2.17	0.42
2:L:21[B]:MET:HG2	2:L:108:THR:HG21	2.01	0.42
2:L:32:SER:HA	4:L:354:HOH:O	2.18	0.42
1:I:12:VAL:HG11	1:I:86:LEU:CD1	2.50	0.42
3:X:3:LEU:CD1	3:X:16:GLY:HA3	2.49	0.42
2:L:37:ASP:O	2:L:56:TRP:HA	2.19	0.41
2:L:113:LYS:HG3	2:L:114:ARG:N	2.35	0.41
2:L:11:LEU:HD23	2:L:19:VAL:HG13	2.03	0.41
2:M:166:LEU:HD23	2:M:166:LEU:HA	1.84	0.41
1:I:218:PRO:O	1:I:219:ARG:HG2	2.20	0.41
1:H:144:LEU:HD12	1:H:144:LEU:N	2.36	0.41
1:H:82:GLU:HB3	4:H:366:HOH:O	2.21	0.41
2:L:113:LYS:NZ	4:L:316:HOH:O	2.48	0.41
1:I:219:ARG:HH11	1:I:219:ARG:CG	2.17	0.41
1:H:205:HIS:CD2	1:H:208:SER:OG	2.70	0.41
2:L:12:ALA:HA	2:L:111:GLU:HG3	2.02	0.41
1:H:12:VAL:HG11	1:H:86:LEU:CD1	2.51	0.41
2:M:51:LYS:NZ	4:M:324:HOH:O	2.50	0.41

There are no symmetry-related clashes.

## 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	H	219/221 (99%)	213 (97%)	6 (3%)	0	100	100
1	I	225/221 (102%)	218 (97%)	6 (3%)	1 (0%)	43	18
2	L	220/220 (100%)	214 (97%)	5 (2%)	1 (0%)	38	13
2	M	222/220 (101%)	216 (97%)	6 (3%)	0	100	100
3	X	13/23 (56%)	13 (100%)	0	0	100	100
3	Y	12/23 (52%)	12 (100%)	0	0	100	100
All	All	911/928 (98%)	886 (97%)	23 (2%)	2 (0%)	56	29

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	I	219	ARG
2	L	158	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	H	189/189 (100%)	184 (97%)	5 (3%)	59	28
1	I	195/189 (103%)	191 (98%)	4 (2%)	66	37
2	L	199/197 (101%)	192 (96%)	7 (4%)	48	18
2	M	201/197 (102%)	200 (100%)	1 (0%)	94	87
3	X	14/22 (64%)	13 (93%)	1 (7%)	21	4
3	Y	13/22 (59%)	13 (100%)	0	100	100
All	All	811/816 (99%)	793 (98%)	18 (2%)	63	34

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	1	GLN
1	H	62	GLN
1	H	121	LYS
1	H	176	LEU
1	H	193	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	I	1	GLN
1	I	121	LYS
1	I	179	ASP
1	I	196	SER
2	L	32	SER
2	L	39	LEU
2	L	113	LYS
2	L	131	LEU
2	L	148	LYS
2	L	166	LEU
2	L	175	LYS
2	M	66	ASP
3	X	3	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	H	177	GLN
1	H	205	HIS
1	I	1	GLN
1	I	177	GLN
1	I	205	HIS
2	L	35	GLN
2	L	144	ASN
2	L	163	ASN
2	L	204	HIS
2	M	144	ASN
2	M	204	HIS
2	M	216	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

There are no ligands in this entry.

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	H	220/221 (99%)	-0.01	9 (4%) 35 34	17, 28, 47, 73	0
1	I	221/221 (100%)	-0.15	5 (2%) 57 57	13, 21, 39, 67	0
2	L	219/220 (99%)	0.02	7 (3%) 45 44	16, 27, 58, 78	0
2	M	220/220 (100%)	-0.24	2 (0%) 81 83	12, 21, 44, 71	0
3	X	15/23 (65%)	0.95	2 (13%) 4 3	21, 25, 65, 65	0
3	Y	14/23 (60%)	0.11	2 (14%) 3 2	15, 20, 39, 52	0
All	All	909/928 (97%)	-0.07	27 (2%) 48 47	12, 24, 49, 78	0

All (27) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	X	3	LEU	7.4
2	M	220	CYS	6.6
1	I	221	PRO	6.5
1	H	220	GLY	5.4
3	X	2	SER	5.4
1	H	178	SER	4.5
1	I	219	ARG	4.3
3	Y	3	LEU	4.2
1	H	42	ALA	4.0
1	H	177	GLN	3.6
2	L	219	GLU	3.4
2	L	218	ASN	3.2
1	I	220	GLY	3.1
1	H	41	HIS	2.9
1	H	119	ALA	2.6
1	I	1	GLN	2.6
3	Y	16	GLY	2.5
1	H	211	LYS	2.3
2	L	135	GLY	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	M	218	ASN	2.3
2	L	163	ASN	2.2
1	H	1	GLN	2.2
2	L	128	SER	2.2
2	L	190	ASP	2.1
1	H	43	LYS	2.0
2	L	162	GLN	2.0
1	I	41	HIS	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 carbohydrates in this entry.

## 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.