



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 23, 2024 – 03:09 AM EDT

PDB ID : 4WHY
Title : Structure of the Hepatitis C virus envelope glycoprotein E2 antigenic region
412-423 bound to the broadly neutralizing antibody 3/11, P21 crystal form
Authors : Krey, T.; Rey, F.A.
Deposited on : 2014-09-24
Resolution : 2.62 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
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)	:	2.37.1

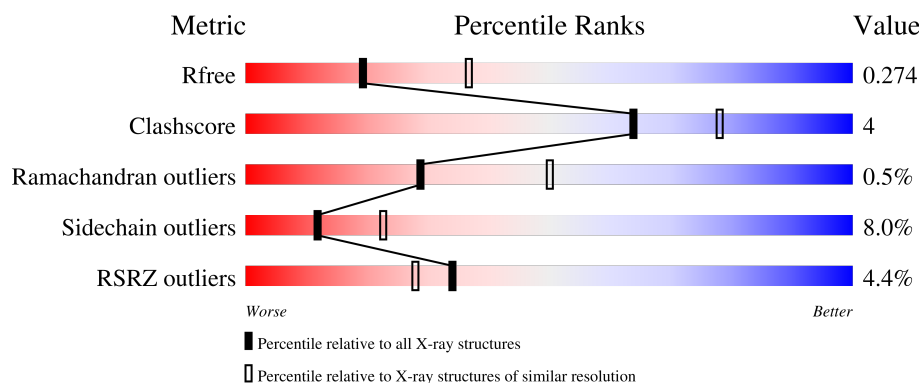
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.62 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	3797 (2.64-2.60)
Clashscore	141614	4168 (2.64-2.60)
Ramachandran outliers	138981	4093 (2.64-2.60)
Sidechain outliers	138945	4093 (2.64-2.60)
RSRZ outliers	127900	3731 (2.64-2.60)

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 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 $\leq 5\%$. 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	12	<div> <div>8%</div> <div> <div></div> <div>75%</div> <div>17%</div> <div>8%</div> </div> </div>
1	B	12	<div> <div>75%</div> <div>8%</div> <div>8%</div> <div>8%</div> </div>
1	C	12	<div> <div>58%</div> <div>33%</div> <div>8%</div> </div>
1	D	12	<div> <div>50%</div> <div>42%</div> <div>8%</div> </div>
2	G	252	<div> <div>6%</div> <div> <div>67%</div> <div>12%</div> <div>21%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	I	252	<div><div></div><div>3%</div><div>70%</div><div>9%</div><div>•</div><div>19%</div></div>
2	K	252	<div><div></div><div>3%</div><div>65%</div><div>13%</div><div>•</div><div>21%</div></div>
2	M	252	<div><div></div><div>2%</div><div>69%</div><div>10%</div><div>•</div><div>21%</div></div>
3	H	220	<div><div></div><div>4%</div><div>83%</div><div>15%</div><div>•</div></div>
3	J	220	<div><div></div><div>4%</div><div>85%</div><div>11%</div><div>• •</div></div>
3	L	220	<div><div></div><div>4%</div><div>76%</div><div>20%</div><div>• •</div></div>
3	N	220	<div><div></div><div>6%</div><div>82%</div><div>14%</div><div>• •</div></div>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 13178 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 epitope peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	11	Total	C	N	O	0	0	0
			89	56	17	16			
1	B	11	Total	C	N	O	0	0	0
			89	56	17	16			
1	C	11	Total	C	N	O	0	0	0
			89	56	17	16			
1	D	11	Total	C	N	O	0	0	0
			89	56	17	16			

- Molecule 2 is a protein called Heavy chain of Fab fragment derived from neutralizing antibody 3/11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	199	Total	C	N	O	S	0	0	0
			1516	959	255	295	7			
2	I	203	Total	C	N	O	S	0	0	0
			1542	976	259	300	7			
2	K	200	Total	C	N	O	S	0	0	0
			1525	967	256	295	7			
2	M	200	Total	C	N	O	S	0	0	0
			1523	964	256	296	7			

- Molecule 3 is a protein called Light chain of Fab fragment derived from neutralizing antibody 3/11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	215	Total	C	N	O	S	0	0	0
			1634	1014	273	340	7			
3	J	215	Total	C	N	O	S	0	0	0
			1634	1014	273	340	7			
3	L	218	Total	C	N	O	S	0	0	0
			1658	1026	277	347	8			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	N	215	Total	C	N	O	S	0	0	0
			1634	1014	273	340	7			

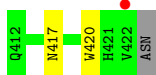
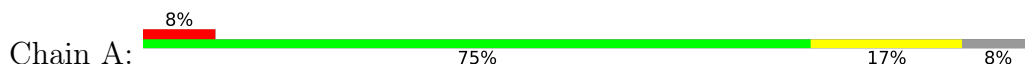
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	O	0	0
			1	1		
4	C	2	Total	O	0	0
			2	2		
4	G	15	Total	O	0	0
			15	15		
4	H	14	Total	O	0	0
			14	14		
4	I	26	Total	O	0	0
			26	26		
4	J	23	Total	O	0	0
			23	23		
4	K	14	Total	O	0	0
			14	14		
4	L	13	Total	O	0	0
			13	13		
4	M	23	Total	O	0	0
			23	23		
4	N	25	Total	O	0	0
			25	25		

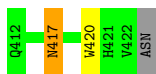
3 Residue-property plots [i](#)

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: epitope peptide



- Molecule 1: epitope peptide



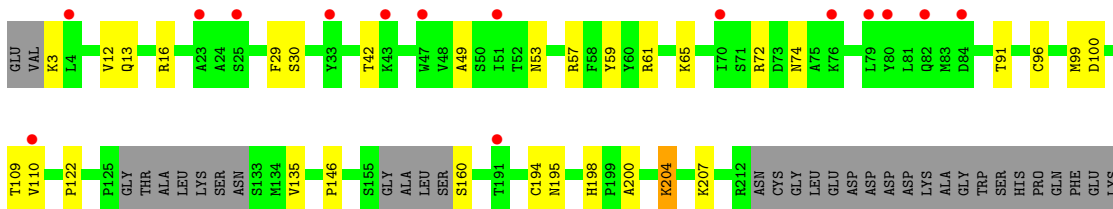
- Molecule 1: epitope peptide



- Molecule 1: epitope peptide



- Molecule 2: Heavy chain of Fab fragment derived from neutralizing antibody 3/11



GLY
GLY
GLY
SER
GLY
GLY
GLY
SER
GLY
GLY
GLY
SER
TRP
SER
HIS
PRO
PHE
GLU
LYS

- Molecule 2: Heavy chain of Fab fragment derived from neutralizing antibody 3/11

Chain I: 3% 70% 9% 19%

GLU V2 Q13 R16 A23 F27 A40 T42 K43 G44 A49 S50 I51 T52 Y59 Y60 R61 K65 L79 Y80 Y83 L86 T91 A92 T93 C96 T109 P122 P125 T125 THR ALA LEU LYS SER ASN S133 M134 V135 S155 A157

L158 S159 V192 N195 K204 K207 V210 P211 R212 ASN CYS GLY LEU GLU ASP ASP ASP LYS ALA GLY TRP SER HIS PRO GLN PHE GLU LYS

- Molecule 2: Heavy chain of Fab fragment derived from neutralizing antibody 3/11

Chain K: 3% 65% 13% 21%

GLU V2 Q13 R16 F27 T28 Y33 T42 K43 G44 A49 S50 T52 N53 S54 G55 G56 Y59 Y60 R61 K65 R72 S85 L86 R87 S88 E89 D90 T91 A92 T93 C96 T109 V110 S111 SER S114 E114 P122 P125 GLY THR ALA LYS

SER ASN S133 M134 V136 T137 F145 S146 E147 N154 S154 A157 L158 S159 Y174 S179 S187 V187 T191 C194 A200 K204 V205 D206 K207 K208 I209 R212 ASN CYS GLY LEU GLU ASP ASP ASP LYS ALA GLY TRP SER HIS PRO GLN PHE GLU LYS

GLY
GLY
SER
GLY
GLY
SER
GLY
GLY
TRP
SER
HIS
PRO
GLN
PHE
GLU
LYS

- Molecule 2: Heavy chain of Fab fragment derived from neutralizing antibody 3/11

Chain M: 2% 69% 10% 21%

GLU V2 K3 S7 V12 R16 S17 L18 T42 A49 N53 R57 F58 Y59 R61 K65 L79 T91 C96 T109 S119 P122 L123 A124 P125 GLY THR ALA LEU LYS SER ASN S133 M134 V135 S155 GLY ALA LEU SER S160 G161 V168

N195 K204 V205 D206 K207 K208 I209 R212 ASN CYS GLY LEU GLU ASP ASP ASP LYS ALA GLY TRP SER HIS PRO GLN PHE GLU LYS


- Molecule 3: Light chain of Fab fragment derived from neutralizing antibody 3/11

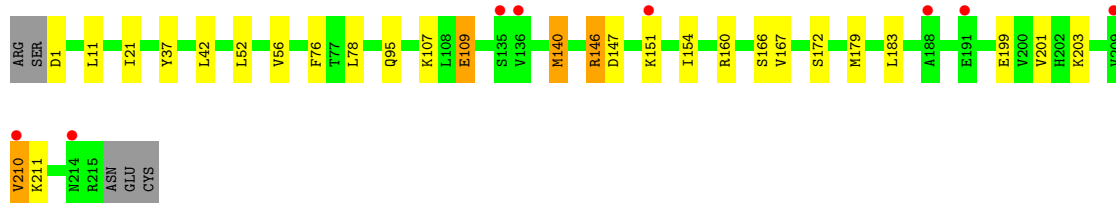
Chain H: 4% 83% 15%

ARG SER D1 L11 V19 S20 I21 Y37 L38 L52 S55 S72 F76 T77 L78 E84 Y90 Y91 M94 Q95 T96 F102 V119 S120 I121 F122 R146 D147 K151 I154 R159 S166 V167 S172 M179 L183 S184 L185 T186 K187


Y190 H193 K203 V210 K211 S212 R215 ASN GLU CYS

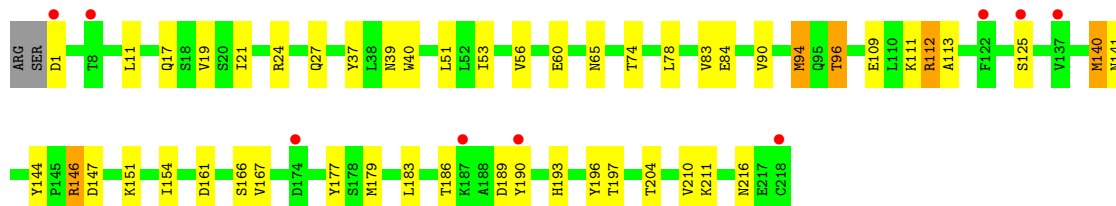
- Molecule 3: Light chain of Fab fragment derived from neutralizing antibody 3/11

Chain J: 




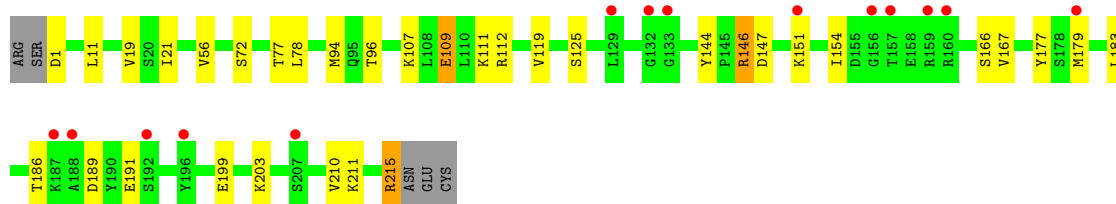
- Molecule 3: Light chain of Fab fragment derived from neutralizing antibody 3/11

Chain L: 



- Molecule 3: Light chain of Fab fragment derived from neutralizing antibody 3/11

Chain N: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	64.76Å 205.51Å 69.02Å 90.00° 103.18° 90.00°	Depositor
Resolution (Å)	47.97 – 2.62 47.97 – 2.62	Depositor EDS
% Data completeness (in resolution range)	96.6 (47.97-2.62) 97.0 (47.97-2.62)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.64 (at 2.61Å)	Xtriage
Refinement program	BUSTER 2.11.2	Depositor
R, R_{free}	0.204 , 0.258 0.219 , 0.274	Depositor DCC
R_{free} test set	2541 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	53.3	Xtriage
Anisotropy	0.729	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 67.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	13178	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 17.30% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0.47	0/91	0.68	0/124
1	B	0.50	0/91	0.70	0/124
1	C	0.52	0/91	0.68	0/124
1	D	0.44	0/91	0.67	0/124
2	G	0.46	0/1553	0.71	0/2118
2	I	0.51	0/1579	0.74	0/2154
2	K	0.47	0/1561	0.76	0/2128
2	M	0.51	0/1560	0.77	1/2128 (0.0%)
3	H	0.48	0/1664	0.76	0/2264
3	J	0.55	0/1664	0.77	0/2264
3	L	0.50	0/1688	0.78	0/2295
3	N	0.55	0/1664	0.78	0/2264
All	All	0.51	0/13297	0.76	1/18111 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	M	119	SER	N-CA-C	-5.42	96.38	111.00

There are no chirality outliers.

There are no planarity outliers.

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	89	0	82	0	0
1	B	89	0	82	1	0
1	C	89	0	82	3	0
1	D	89	0	82	4	0
2	G	1516	0	1486	12	0
2	I	1542	0	1516	12	0
2	K	1525	0	1500	21	0
2	M	1523	0	1495	10	0
3	H	1634	0	1593	14	0
3	J	1634	0	1593	11	0
3	L	1658	0	1610	25	0
3	N	1634	0	1593	12	0
4	A	1	0	0	1	0
4	C	2	0	0	0	0
4	G	15	0	0	0	0
4	H	14	0	0	0	0
4	I	26	0	0	1	0
4	J	23	0	0	1	0
4	K	14	0	0	0	0
4	L	13	0	0	1	0
4	M	23	0	0	0	0
4	N	25	0	0	1	0
All	All	13178	0	12714	108	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 4.

The worst 5 of 108 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:L:51:LEU:HD23	3:L:60:GLU:HG3	1.59	0.83
3:L:40:TRP:HB2	3:L:53:ILE:HG13	1.60	0.81
2:K:146:PRO:O	2:K:147:GLU:HB2	1.82	0.77
2:I:13:GLN:HG3	2:I:16:ARG:HD3	1.74	0.69
3:J:140:MET:HG3	3:J:179:MET:HB3	1.77	0.66

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	9/12 (75%)	9 (100%)	0	0	100	100
1	B	9/12 (75%)	9 (100%)	0	0	100	100
1	C	9/12 (75%)	9 (100%)	0	0	100	100
1	D	9/12 (75%)	9 (100%)	0	0	100	100
2	G	193/252 (77%)	185 (96%)	7 (4%)	1 (0%)	29	50
2	I	197/252 (78%)	190 (96%)	7 (4%)	0	100	100
2	K	192/252 (76%)	185 (96%)	4 (2%)	3 (2%)	9	18
2	M	194/252 (77%)	186 (96%)	6 (3%)	2 (1%)	15	30
3	H	213/220 (97%)	202 (95%)	10 (5%)	1 (0%)	29	50
3	J	213/220 (97%)	205 (96%)	8 (4%)	0	100	100
3	L	216/220 (98%)	207 (96%)	8 (4%)	1 (0%)	29	50
3	N	213/220 (97%)	204 (96%)	8 (4%)	1 (0%)	29	50
All	All	1667/1936 (86%)	1600 (96%)	58 (4%)	9 (0%)	29	50

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	K	55	GLY
2	M	53	ASN
2	G	53	ASN
2	K	147	GLU
3	L	204	THR

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	10/11 (91%)	8 (80%)	2 (20%)	1	2
1	B	10/11 (91%)	8 (80%)	2 (20%)	1	2
1	C	10/11 (91%)	8 (80%)	2 (20%)	1	2
1	D	10/11 (91%)	8 (80%)	2 (20%)	1	2
2	G	171/208 (82%)	159 (93%)	12 (7%)	15	29
2	I	174/208 (84%)	163 (94%)	11 (6%)	18	35
2	K	172/208 (83%)	159 (92%)	13 (8%)	13	25
2	M	172/208 (83%)	160 (93%)	12 (7%)	15	29
3	H	191/196 (97%)	177 (93%)	14 (7%)	14	27
3	J	191/196 (97%)	177 (93%)	14 (7%)	14	27
3	L	194/196 (99%)	173 (89%)	21 (11%)	6	11
3	N	191/196 (97%)	177 (93%)	14 (7%)	14	27
All	All	1496/1660 (90%)	1377 (92%)	119 (8%)	12	23

5 of 119 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	J	210	VAL
3	N	107	LYS
2	K	207	LYS
3	N	77	THR
3	N	215	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	412	GLN

5.3.3 RNA ⓘ

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](#)

There are no ligands in this entry.

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	11/12 (91%)	0.43	1 (9%) 9 6	40, 52, 69, 71	0
1	B	11/12 (91%)	0.59	0 100 100	34, 49, 70, 73	0
1	C	11/12 (91%)	-0.10	0 100 100	49, 58, 67, 67	0
1	D	11/12 (91%)	0.47	0 100 100	56, 61, 69, 75	0
2	G	199/252 (78%)	0.45	15 (7%) 14 10	41, 67, 85, 106	0
2	I	203/252 (80%)	0.23	7 (3%) 45 38	32, 56, 76, 90	0
2	K	200/252 (79%)	0.22	7 (3%) 44 37	38, 58, 76, 88	0
2	M	200/252 (79%)	0.06	5 (2%) 57 51	32, 50, 72, 94	0
3	H	215/220 (97%)	0.32	9 (4%) 36 30	40, 66, 87, 106	0
3	J	215/220 (97%)	0.20	8 (3%) 41 35	27, 55, 83, 97	0
3	L	218/220 (99%)	0.30	9 (4%) 37 31	41, 61, 82, 105	0
3	N	215/220 (97%)	0.30	14 (6%) 18 14	31, 51, 78, 102	0
All	All	1709/1936 (88%)	0.26	75 (4%) 34 28	27, 58, 82, 106	0

The worst 5 of 75 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	L	218	CYS	6.8
3	N	156	GLY	5.2
3	L	8	THR	4.8
3	L	1	ASP	4.7
3	J	188	ALA	4.2

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.