



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 6, 2026 – 11:22 PM UTC

PDB ID : 4PLK / pdb\_00004plk  
Title : Hepatitis E Virus E2s domain (Genotype I) in complex with a neutralizing antibody 8G12  
Authors : Tang, X.H.; Li, S.W.; Sivaraman, J.  
Deposited on : 2014-05-18  
Resolution : 4.00 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](mailto: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 : 4-5-2 with Phenix2.0  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

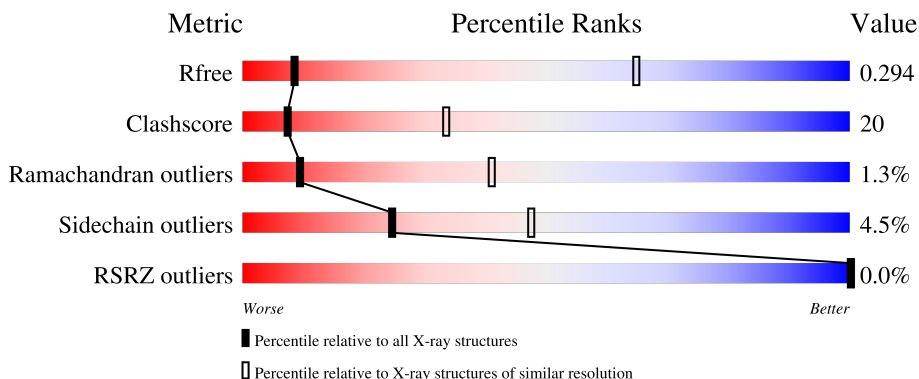
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.00 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))
$R_{free}$	180053	1082 (4.20-3.80)
Clashscore	190562	1129 (4.20-3.80)
Ramachandran outliers	187476	1064 (4.20-3.80)
Sidechain outliers	187428	1055 (4.20-3.80)
RSRZ outliers	180081	1082 (4.20-3.80)

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  $\geq 3$ , 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	147	55% (green), 42% (yellow), 3% (orange), 0% (red), 0% (grey)
1	B	147	65% (green), 33% (yellow), 2% (orange), 0% (red), 0% (grey)
1	E	147	59% (green), 39% (yellow), 2% (orange), 0% (red), 0% (grey)
1	F	147	55% (green), 40% (yellow), 5% (orange), 0% (red), 0% (grey)
2	C	212	44% (green), 38% (yellow), 9% (orange), 7% (red), 2% (grey)

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	G	212	
2	K	212	
2	L	212	
3	D	229	
3	H	229	
3	I	229	
3	J	229	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 17362 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 Capsid protein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	A	147	1104	702	186	216	0	0	0
1	B	147	1104	702	186	216	0	0	0
1	E	147	1104	702	186	216	0	0	0
1	F	147	1104	702	186	216	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	604	PRO	-	expression tag	UNP L0L7P5
A	605	PRO	-	expression tag	UNP L0L7P5
B	604	PRO	-	expression tag	UNP L0L7P5
B	605	PRO	-	expression tag	UNP L0L7P5
E	604	PRO	-	expression tag	UNP L0L7P5
E	605	PRO	-	expression tag	UNP L0L7P5
F	604	PRO	-	expression tag	UNP L0L7P5
F	605	PRO	-	expression tag	UNP L0L7P5

- Molecule 2 is a protein called 8G12 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	L	208	1618	1010	271	331	6	0	0	0
2	C	198	1533	959	257	312	5	0	0	0
2	K	208	1618	1010	271	331	6	0	0	0
2	G	198	1531	957	257	312	5	0	0	0

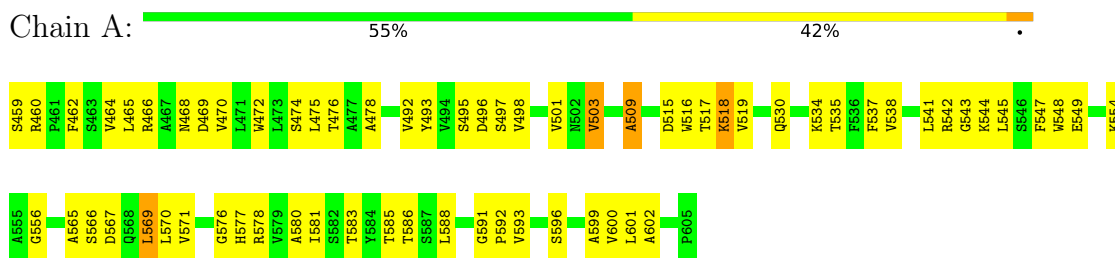
- Molecule 3 is a protein called 8G12 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	H	216	Total 1654	C 1056	N 271	O 320	S 7	0	0	0
3	D	218	Total 1669	C 1064	N 274	O 324	S 7	0	0	0
3	J	216	Total 1654	C 1056	N 271	O 320	S 7	0	0	0
3	I	218	Total 1669	C 1064	N 274	O 324	S 7	0	0	0

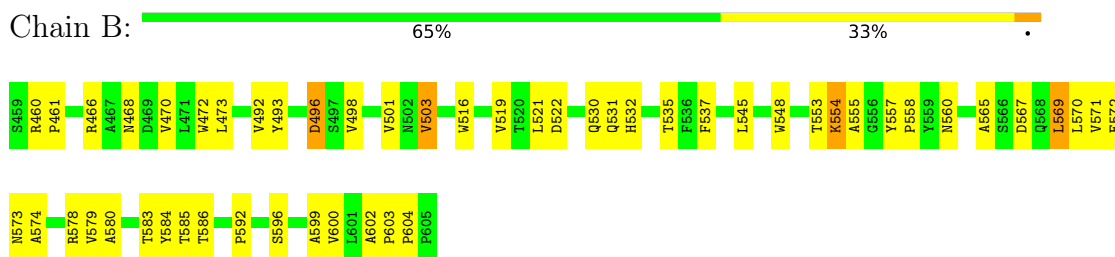
### 3 Residue-property plots [i](#)

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.

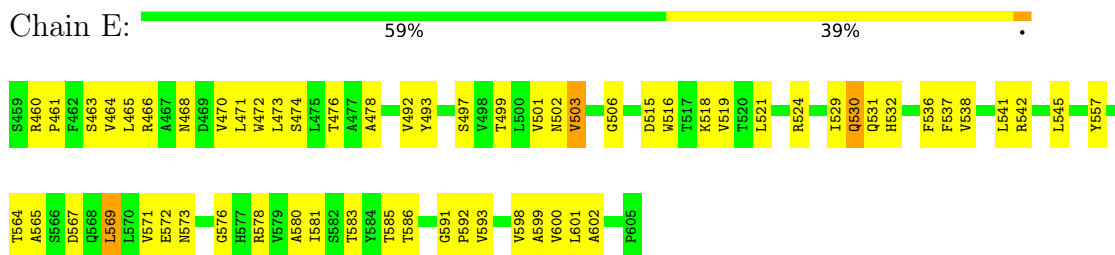
- Molecule 1: Capsid protein



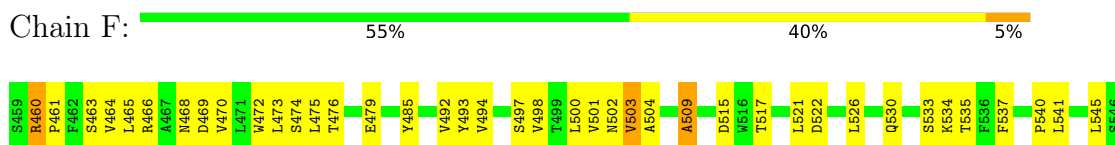
- Molecule 1: Capsid protein



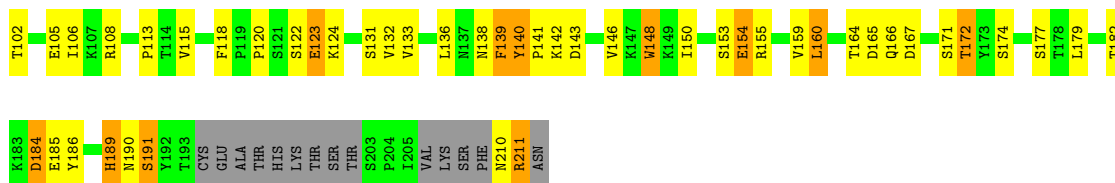
- Molecule 1: Capsid protein



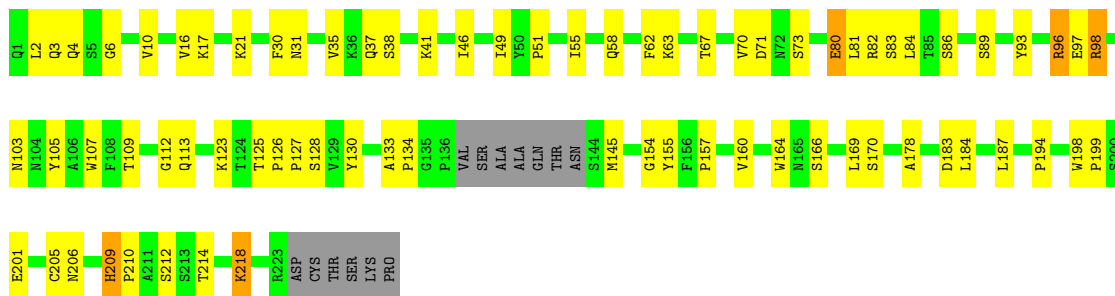
- Molecule 1: Capsid protein



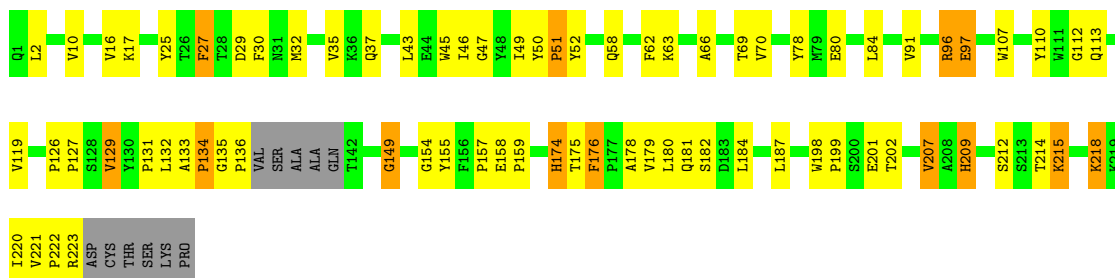




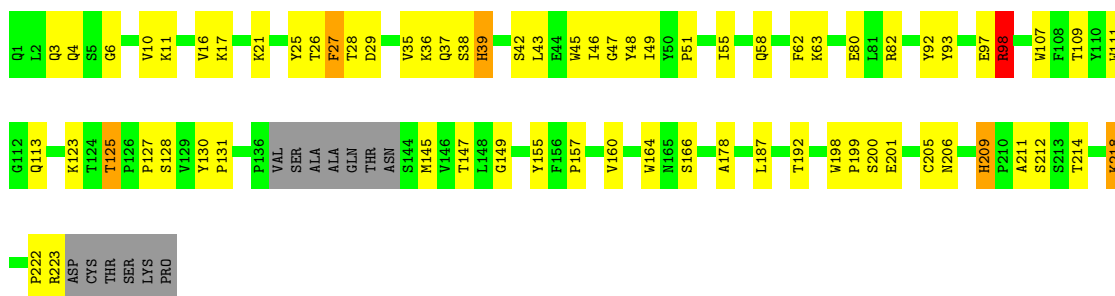
• Molecule 3: 8G12 heavy chain



• Molecule 3: 8G12 heavy chain



• Molecule 3: 8G12 heavy chain



• Molecule 3: 8G12 heavy chain





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	54.20Å 191.99Å 192.58Å 90.00° 90.03° 90.00°	Depositor
Resolution (Å)	33.99 – 4.00 33.99 – 4.00	Depositor EDS
% Data completeness (in resolution range)	91.2 (33.99-4.00) 82.2 (33.99-4.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.79 (at 3.87Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.264 , 0.296 0.273 , 0.294	Depositor DCC
$R_{free}$ test set	1971 reflections (6.20%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	78.6	Xtrriage
Anisotropy	0.580	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 71.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	0.378 for -h,l,k 0.379 for -h,-l,-k 0.399 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	17362	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	141.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.31% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical 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	0.41	0/1131	1.01	3/1552 (0.2%)
1	B	0.44	0/1131	1.00	5/1552 (0.3%)
1	E	0.45	0/1131	1.08	9/1552 (0.6%)
1	F	0.43	0/1131	1.11	6/1552 (0.4%)
2	C	0.48	0/1567	1.29	20/2126 (0.9%)
2	G	0.48	0/1565	1.28	18/2122 (0.8%)
2	K	0.47	0/1655	1.23	13/2246 (0.6%)
2	L	0.43	0/1655	1.15	14/2246 (0.6%)
3	D	0.45	0/1717	1.14	15/2346 (0.6%)
3	H	0.37	0/1702	1.06	8/2325 (0.3%)
3	I	0.47	0/1717	1.40	27/2346 (1.2%)
3	J	0.39	0/1702	1.07	12/2325 (0.5%)
All	All	0.44	0/17804	1.17	150/24290 (0.6%)

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
2	C	0	1
2	G	0	1
2	K	0	2
3	I	0	2
All	All	0	6

There are no bond length outliers.

The worst 5 of 150 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	130	TYR	CA-C-N	14.97	134.97	119.85
3	I	130	TYR	C-N-CA	14.97	134.97	119.85

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	220	ILE	N-CA-C	13.46	128.82	108.23
2	C	152	GLY	N-CA-C	12.61	130.08	115.08
3	I	131	PRO	N-CA-C	12.25	130.85	111.38

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	172	THR	Peptide
2	G	141	PRO	Peptide
3	I	207	VAL	Peptide
2	K	159	VAL	Peptide
2	K	209	PHE	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	1104	0	1092	67	0
1	B	1104	0	1090	53	0
1	E	1104	0	1092	38	0
1	F	1104	0	1092	54	0
2	C	1533	0	1464	104	0
2	G	1531	0	1456	67	0
2	K	1618	0	1544	77	0
2	L	1618	0	1542	97	0
3	D	1669	0	1624	67	0
3	H	1654	0	1611	53	0
3	I	1669	0	1624	66	0
3	J	1654	0	1611	53	0
All	All	17362	0	16842	696	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 20.

The worst 5 of 696 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)
1:B:554:LYS:CD	2:C:92:TRP:CZ2	2.15	1.27
1:B:554:LYS:HD2	2:C:92:TRP:CE2	1.70	1.24
1:B:554:LYS:CD	2:C:92:TRP:CE2	2.19	1.23
1:B:554:LYS:HD2	2:C:92:TRP:CZ2	1.79	1.16
1:A:554:LYS:HE2	2:L:93:GLU:HG3	1.23	1.15

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	145/147 (99%)	139 (96%)	4 (3%)	2 (1%)	9	39
1	B	145/147 (99%)	140 (97%)	4 (3%)	1 (1%)	18	54
1	E	145/147 (99%)	138 (95%)	6 (4%)	1 (1%)	18	54
1	F	145/147 (99%)	136 (94%)	6 (4%)	3 (2%)	5	32
2	C	192/212 (91%)	176 (92%)	11 (6%)	5 (3%)	4	29
2	G	192/212 (91%)	176 (92%)	13 (7%)	3 (2%)	7	37
2	K	204/212 (96%)	193 (95%)	8 (4%)	3 (2%)	8	38
2	L	204/212 (96%)	192 (94%)	9 (4%)	3 (2%)	8	38
3	D	214/229 (93%)	202 (94%)	10 (5%)	2 (1%)	14	48
3	H	212/229 (93%)	208 (98%)	4 (2%)	0	100	100
3	I	214/229 (93%)	205 (96%)	5 (2%)	4 (2%)	6	34
3	J	212/229 (93%)	207 (98%)	3 (1%)	2 (1%)	14	48
All	All	2224/2352 (95%)	2112 (95%)	83 (4%)	29 (1%)	9	41

5 of 29 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	577	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	503	VAL
2	L	169	LYS
2	C	154	GLU
2	C	157	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	120/120 (100%)	116 (97%)	4 (3%)	33	56
1	B	120/120 (100%)	116 (97%)	4 (3%)	33	56
1	E	120/120 (100%)	114 (95%)	6 (5%)	22	45
1	F	120/120 (100%)	116 (97%)	4 (3%)	33	56
2	C	175/190 (92%)	158 (90%)	17 (10%)	8	27
2	G	175/190 (92%)	161 (92%)	14 (8%)	11	35
2	K	186/190 (98%)	182 (98%)	4 (2%)	45	65
2	L	186/190 (98%)	178 (96%)	8 (4%)	26	48
3	D	190/199 (96%)	183 (96%)	7 (4%)	30	52
3	H	188/199 (94%)	182 (97%)	6 (3%)	34	56
3	I	190/199 (96%)	179 (94%)	11 (6%)	18	42
3	J	188/199 (94%)	184 (98%)	4 (2%)	47	65
All	All	1958/2036 (96%)	1869 (96%)	89 (4%)	24	47

5 of 89 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	K	162	SER
2	G	123	GLU
2	K	205	ILE
2	G	34	VAL
2	G	160	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 39 such sidechains are listed below:

Mol	Chain	Res	Type
3	J	31	ASN
3	I	31	ASN
3	J	174	HIS
2	G	37	GLN
3	I	58	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 oligosaccharides 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, 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	A	147/147 (100%)	-0.92	0 100 100	73, 103, 146, 165	0
1	B	147/147 (100%)	-0.95	0 100 100	79, 104, 153, 193	0
1	E	147/147 (100%)	-0.87	0 100 100	81, 108, 159, 180	0
1	F	147/147 (100%)	-0.91	0 100 100	68, 107, 158, 175	0
2	C	198/212 (93%)	-0.83	0 100 100	55, 169, 218, 283	0
2	G	198/212 (93%)	-0.82	0 100 100	51, 167, 208, 237	0
2	K	208/212 (98%)	-0.84	0 100 100	63, 167, 209, 234	0
2	L	208/212 (98%)	-0.83	0 100 100	49, 163, 234, 272	0
3	D	218/229 (95%)	-0.87	0 100 100	96, 135, 210, 233	0
3	H	216/229 (94%)	-0.92	0 100 100	85, 128, 202, 253	0
3	I	218/229 (95%)	-0.86	1 (0%) 87 73	98, 135, 211, 235	0
3	J	216/229 (94%)	-0.86	0 100 100	95, 134, 192, 219	0
All	All	2268/2352 (96%)	-0.87	1 (0%) 100 100	49, 135, 210, 283	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	I	189	SER	2.6

### 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 oligosaccharides 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.