



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 17, 2026 – 10:27 PM UTC

PDB ID : 8TOO / pdb_00008too
Title : Crystal structure of Epstein-Barr virus gp42 in complex with antibody 4C12
Authors : Bu, W.; Kumar, A.; Board, N.; Kim, J.; Dowdell, K.; Zhang, S.; Lei, Y.; Hostal, A.; Krogmann, T.; Wang, Y.; Pittaluga, S.; Marcotrigiano, J.; Cohen, J.I.
Deposited on : 2023-08-03
Resolution : 2.60 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	:	4-5-2 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

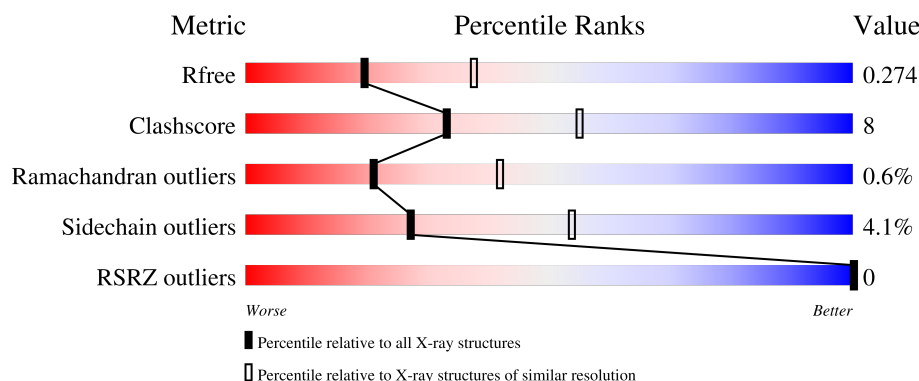
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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}	180053	4008 (2.60-2.60)
Clashscore	190562	4347 (2.60-2.60)
Ramachandran outliers	187476	4277 (2.60-2.60)
Sidechain outliers	187428	4277 (2.60-2.60)
RSRZ outliers	180081	4008 (2.60-2.60)

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	214	 85% 14% .
1	C	214	 80% 17% ..
1	E	214	 79% 18% ..
1	F	214	 83% 16% .
2	B	234	 70% 23% . 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	234	 75% 19% 6%
2	G	234	 73% 21% 6%
2	H	234	 69% 24% 6%
3	I	139	 77% 19% ..
3	J	139	 83% 14% .
3	K	139	 73% 23% ..
3	L	139	 84% 13% ..

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 17213 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 4C12 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	211	Total	C	N	O	S	0	0	0
			1589	987	266	331	5			
1	C	211	Total	C	N	O	S	0	0	0
			1582	985	266	326	5			
1	E	211	Total	C	N	O	S	0	0	0
			1597	999	272	321	5			
1	F	211	Total	C	N	O	S	0	0	0
			1579	984	268	322	5			

- Molecule 2 is a protein called 4C12 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	220	Total	C	N	O	S	0	0	0
			1605	1015	259	324	7			
2	D	220	Total	C	N	O	S	0	0	0
			1603	1014	261	321	7			
2	G	220	Total	C	N	O	S	0	0	0
			1622	1025	265	325	7			
2	H	220	Total	C	N	O	S	0	0	0
			1633	1032	266	328	7			

- Molecule 3 is a protein called Glycoprotein 42.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	136	Total	C	N	O	S	0	0	0
			1044	684	168	181	11			
3	J	136	Total	C	N	O	S	0	0	0
			1045	681	167	186	11			
3	K	136	Total	C	N	O	S	0	0	0
			1057	688	170	188	11			
3	L	136	Total	C	N	O	S	0	0	0
			1066	692	174	189	11			

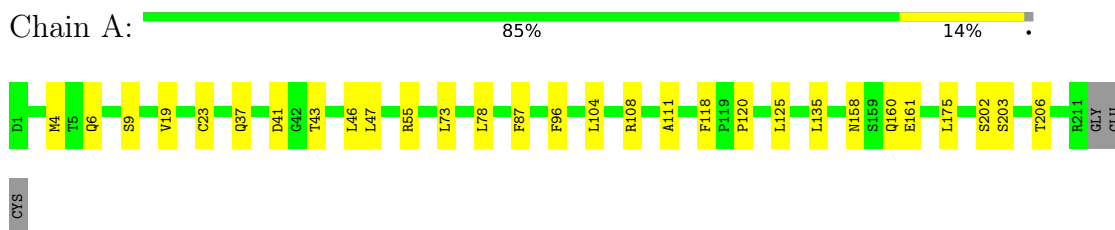
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	15	Total 15	O 15	0	0
4	B	20	Total 20	O 20	0	0
4	C	14	Total 14	O 14	0	0
4	D	23	Total 23	O 23	0	0
4	E	22	Total 22	O 22	0	0
4	F	21	Total 21	O 21	0	0
4	G	19	Total 19	O 19	0	0
4	H	20	Total 20	O 20	0	0
4	I	11	Total 11	O 11	0	0
4	J	8	Total 8	O 8	0	0
4	K	9	Total 9	O 9	0	0
4	L	9	Total 9	O 9	0	0

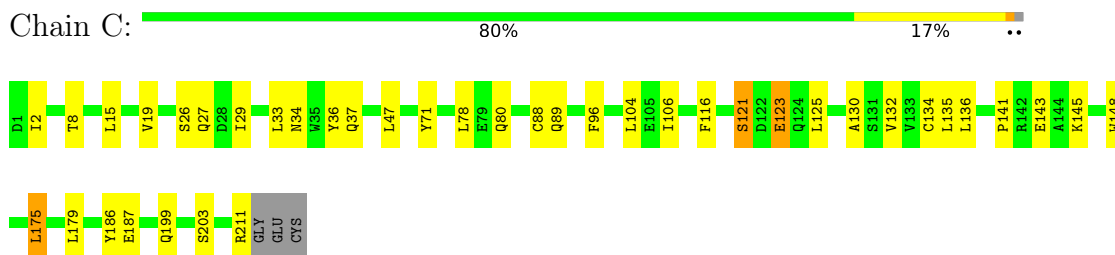
3 Residue-property plots

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.

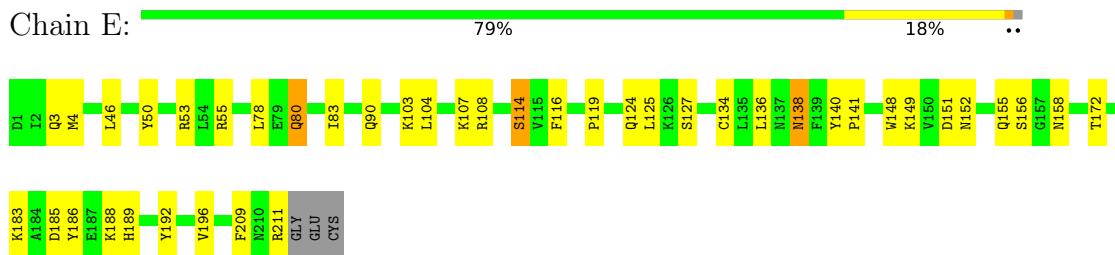
- Molecule 1: 4C12 light chain



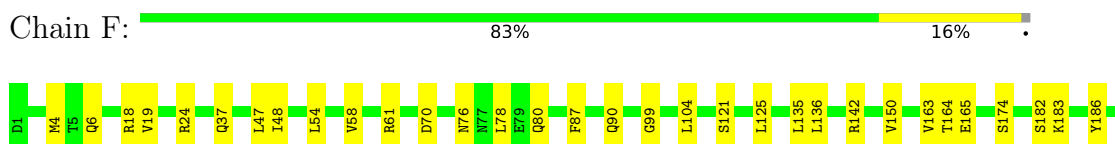
- Molecule 1: 4C12 light chain

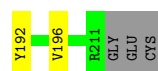


- Molecule 1: 4C12 light chain



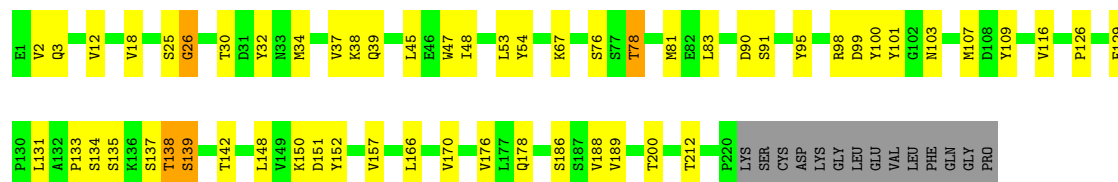
- Molecule 1: 4C12 light chain





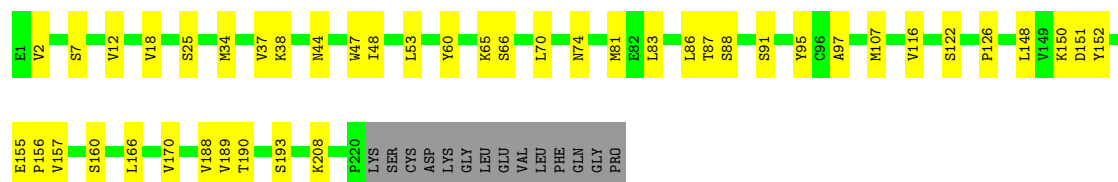
• Molecule 2: 4C12 heavy chain

Chain B: 70% 23% 6%



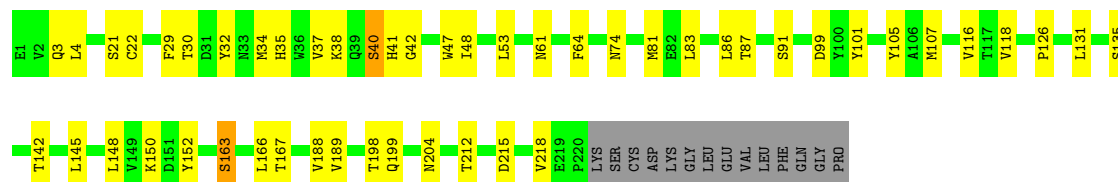
• Molecule 2: 4C12 heavy chain

Chain D: 75% 19% 6%



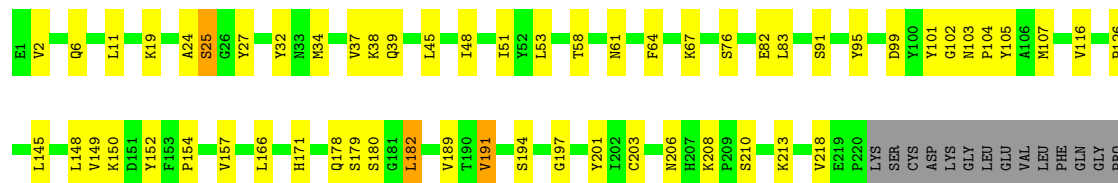
• Molecule 2: 4C12 heavy chain

Chain G: 73% 21% 6%



• Molecule 2: 4C12 heavy chain

Chain H: 69% 24% 6%

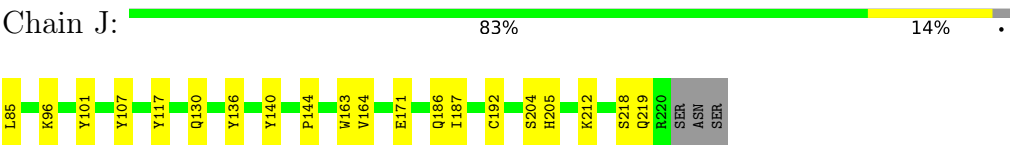


• Molecule 3: Glycoprotein 42

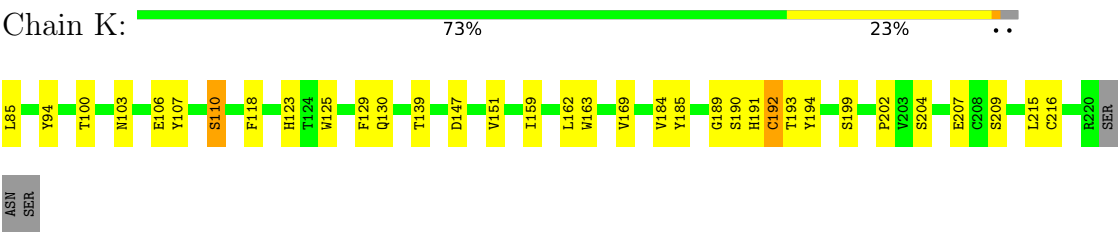
Chain I: 77% 19% ..



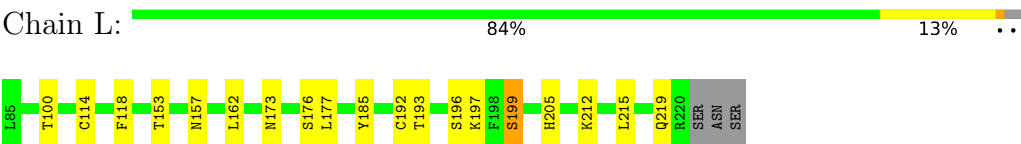
● Molecule 3: Glycoprotein 42



● Molecule 3: Glycoprotein 42



● Molecule 3: Glycoprotein 42



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	53.30Å 63.57Å 199.28Å 89.81° 89.75° 88.86°	Depositor
Resolution (Å)	53.51 – 2.60 53.51 – 2.60	Depositor EDS
% Data completeness (in resolution range)	91.5 (53.51-2.60) 90.5 (53.51-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.16	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.24 (at 2.61Å)	Xtriage
Refinement program	PHENIX (1.19.2_4158: ???)	Depositor
R, R_{free}	0.226 , 0.275 0.225 , 0.274	Depositor DCC
R_{free} test set	3755 reflections (4.66%)	wwPDB-VP
Wilson B-factor (Å ²)	42.5	Xtriage
Anisotropy	0.800	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 34.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.095 for h,-k,-l 0.094 for -h,k,-l 0.400 for -h,-k,l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	17213	wwPDB-VP
Average B, all atoms (Å ²)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.95% 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	0.11	0/1620	0.34	0/2210
1	C	0.11	0/1613	0.34	0/2201
1	E	0.14	0/1628	0.39	0/2214
1	F	0.12	0/1610	0.35	0/2194
2	B	0.21	0/1649	0.46	0/2261
2	D	0.24	0/1647	0.43	0/2257
2	G	0.24	0/1666	0.50	0/2281
2	H	0.22	0/1677	0.45	0/2292
3	I	0.16	0/1082	0.37	0/1482
3	J	0.30	0/1083	0.46	0/1485
3	K	0.53	0/1095	0.74	0/1500
3	L	0.21	0/1104	0.45	0/1511
All	All	0.23	0/17474	0.44	0/23888

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 [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	1589	0	1482	16	0
1	C	1582	0	1478	24	0
1	E	1597	0	1539	32	0
1	F	1579	0	1486	19	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1605	0	1473	35	0
2	D	1603	0	1469	22	0
2	G	1622	0	1508	30	0
2	H	1633	0	1530	33	0
3	I	1044	0	947	15	0
3	J	1045	0	934	12	0
3	K	1057	0	956	19	0
3	L	1066	0	971	12	0
4	A	15	0	0	0	0
4	B	20	0	0	0	0
4	C	14	0	0	1	0
4	D	23	0	0	1	0
4	E	22	0	0	1	0
4	F	21	0	0	0	0
4	G	19	0	0	0	0
4	H	20	0	0	3	0
4	I	11	0	0	0	0
4	J	8	0	0	0	0
4	K	9	0	0	0	0
4	L	9	0	0	1	0
All	All	17213	0	15773	254	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 8.

The worst 5 of 254 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:E:4:MET:HE2	1:E:90:GLN:HB3	1.54	0.89
2:B:48:ILE:HD13	2:B:81:MET:HE1	1.56	0.87
2:G:131:LEU:HD21	2:G:148:LEU:HB2	1.59	0.85
1:F:142:ARG:HD2	1:F:163:VAL:HG11	1.65	0.78
1:E:136:LEU:HD21	1:E:196:VAL:HG21	1.68	0.75

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	209/214 (98%)	198 (95%)	11 (5%)	0	100	100
1	C	209/214 (98%)	199 (95%)	10 (5%)	0	100	100
1	E	209/214 (98%)	200 (96%)	8 (4%)	1 (0%)	24	46
1	F	209/214 (98%)	200 (96%)	9 (4%)	0	100	100
2	B	218/234 (93%)	197 (90%)	15 (7%)	6 (3%)	4	6
2	D	218/234 (93%)	203 (93%)	14 (6%)	1 (0%)	24	46
2	G	218/234 (93%)	207 (95%)	10 (5%)	1 (0%)	24	46
2	H	218/234 (93%)	205 (94%)	12 (6%)	1 (0%)	24	46
3	I	134/139 (96%)	126 (94%)	7 (5%)	1 (1%)	18	38
3	J	134/139 (96%)	128 (96%)	5 (4%)	1 (1%)	18	38
3	K	134/139 (96%)	127 (95%)	6 (4%)	1 (1%)	18	38
3	L	134/139 (96%)	125 (93%)	9 (7%)	0	100	100
All	All	2244/2348 (96%)	2115 (94%)	116 (5%)	13 (1%)	21	42

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	2	VAL
2	B	26	GLY
2	B	133	PRO
3	I	191	HIS
3	J	171	GLU

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	177/191 (93%)	171 (97%)	6 (3%)	32	60
1	C	175/191 (92%)	167 (95%)	8 (5%)	24	49
1	E	179/191 (94%)	176 (98%)	3 (2%)	53	78
1	F	174/191 (91%)	169 (97%)	5 (3%)	37	65
2	B	172/200 (86%)	166 (96%)	6 (4%)	32	59
2	D	170/200 (85%)	160 (94%)	10 (6%)	18	38
2	G	176/200 (88%)	168 (96%)	8 (4%)	24	50
2	H	179/200 (90%)	172 (96%)	7 (4%)	28	55
3	I	109/126 (86%)	106 (97%)	3 (3%)	38	66
3	J	110/126 (87%)	107 (97%)	3 (3%)	39	67
3	K	113/126 (90%)	103 (91%)	10 (9%)	9	21
3	L	115/126 (91%)	109 (95%)	6 (5%)	21	44
All	All	1849/2068 (89%)	1774 (96%)	75 (4%)	27	54

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	J	192	CYS
3	L	173	ASN
3	K	85	LEU
3	K	159	ILE
2	D	122	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
3	K	130	GLN
3	K	191	HIS
1	F	6	GLN
1	E	38	GLN
3	L	123	HIS

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 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, 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	211/214 (98%)	-1.20	0 100 100	28, 47, 78, 102	0
1	C	211/214 (98%)	-1.24	0 100 100	26, 47, 77, 101	0
1	E	211/214 (98%)	-1.24	0 100 100	27, 46, 77, 92	0
1	F	211/214 (98%)	-1.22	0 100 100	28, 47, 71, 100	0
2	B	220/234 (94%)	-1.18	0 100 100	29, 47, 83, 98	0
2	D	220/234 (94%)	-1.23	0 100 100	27, 45, 83, 108	0
2	G	220/234 (94%)	-1.19	0 100 100	30, 46, 82, 105	0
2	H	220/234 (94%)	-1.20	0 100 100	28, 46, 78, 114	0
3	I	136/139 (97%)	-1.08	0 100 100	31, 51, 82, 107	0
3	J	136/139 (97%)	-1.09	0 100 100	32, 51, 82, 103	0
3	K	136/139 (97%)	-1.02	0 100 100	30, 55, 84, 98	0
3	L	136/139 (97%)	-1.01	0 100 100	31, 56, 85, 99	0
All	All	2268/2348 (96%)	-1.17	0 100 100	26, 48, 80, 114	0

There are no RSRZ outliers to report.

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 [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.