



wwPDB X-ray Structure Validation Summary Report

Feb 1, 2022 – 03:14 pm GMT

PDB ID : 7PI2
Title : PfCyRPA bound to monoclonal antibody Cy.003 Fab fragment
Authors : Ragotte, R.J.; Higgins, M.K.
Deposited on : 2021-08-19
Resolution : 3.14 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 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 : 2.26
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.26

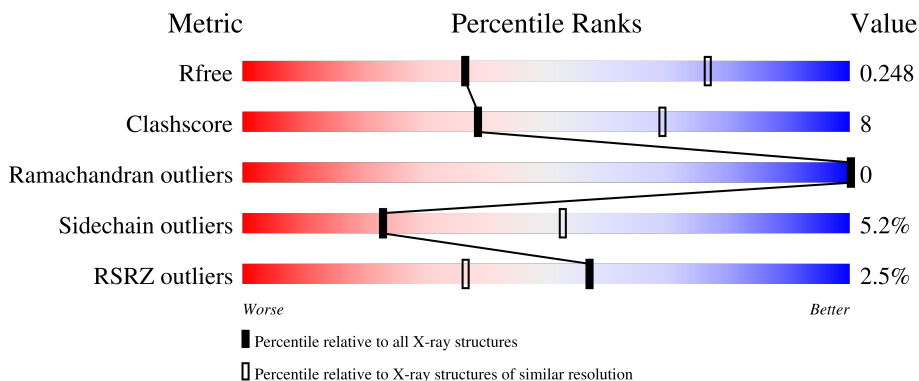
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.14 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	1626 (3.18-3.10)
Clashscore	141614	1735 (3.18-3.10)
Ramachandran outliers	138981	1677 (3.18-3.10)
Sidechain outliers	138945	1677 (3.18-3.10)
RSRZ outliers	127900	1588 (3.18-3.10)

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	343	 2% 75% 18% • 6%
1	D	343	 3% 74% 18% • 5%
1	G	343	 5% 68% 20% • 10%
1	J	343	 6% 73% 19% • 7%
2	B	231	 86% 10% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	231	 <p>81% 15% ..</p>
2	H	231	 <p>84% 13% ..</p>
2	K	231	 <p>87% 10% ..</p>
3	C	210	 <p>81% 16% ..</p>
3	F	210	 <p>82% 15% ..</p>
3	I	210	 <p>87% 11% ..</p>
3	L	210	 <p>82% 15% ..</p>

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 23353 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 Cysteine-rich protective antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	321	2674	1719	431	511	13	0	0	0
1	D	327	2725	1752	440	520	13	0	0	0
1	G	310	2580	1663	413	491	13	0	0	0
1	J	319	2646	1702	426	505	13	0	0	0

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	147	ALA	SER	conflict	UNP Q8IFM8
A	324	ALA	THR	conflict	UNP Q8IFM8
A	340	ALA	THR	conflict	UNP Q8IFM8
A	363	GLY	-	expression tag	UNP Q8IFM8
A	364	GLY	-	expression tag	UNP Q8IFM8
A	365	GLY	-	expression tag	UNP Q8IFM8
A	366	GLY	-	expression tag	UNP Q8IFM8
A	367	SER	-	expression tag	UNP Q8IFM8
A	368	GLU	-	expression tag	UNP Q8IFM8
A	369	PRO	-	expression tag	UNP Q8IFM8
A	370	GLU	-	expression tag	UNP Q8IFM8
A	371	ALA	-	expression tag	UNP Q8IFM8
D	147	ALA	SER	conflict	UNP Q8IFM8
D	324	ALA	THR	conflict	UNP Q8IFM8
D	340	ALA	THR	conflict	UNP Q8IFM8
D	363	GLY	-	expression tag	UNP Q8IFM8
D	364	GLY	-	expression tag	UNP Q8IFM8
D	365	GLY	-	expression tag	UNP Q8IFM8
D	366	GLY	-	expression tag	UNP Q8IFM8
D	367	SER	-	expression tag	UNP Q8IFM8
D	368	GLU	-	expression tag	UNP Q8IFM8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	369	PRO	-	expression tag	UNP Q8IFM8
D	370	GLU	-	expression tag	UNP Q8IFM8
D	371	ALA	-	expression tag	UNP Q8IFM8
G	147	ALA	SER	conflict	UNP Q8IFM8
G	324	ALA	THR	conflict	UNP Q8IFM8
G	340	ALA	THR	conflict	UNP Q8IFM8
G	363	GLY	-	expression tag	UNP Q8IFM8
G	364	GLY	-	expression tag	UNP Q8IFM8
G	365	GLY	-	expression tag	UNP Q8IFM8
G	366	GLY	-	expression tag	UNP Q8IFM8
G	367	SER	-	expression tag	UNP Q8IFM8
G	368	GLU	-	expression tag	UNP Q8IFM8
G	369	PRO	-	expression tag	UNP Q8IFM8
G	370	GLU	-	expression tag	UNP Q8IFM8
G	371	ALA	-	expression tag	UNP Q8IFM8
J	147	ALA	SER	conflict	UNP Q8IFM8
J	324	ALA	THR	conflict	UNP Q8IFM8
J	340	ALA	THR	conflict	UNP Q8IFM8
J	363	GLY	-	expression tag	UNP Q8IFM8
J	364	GLY	-	expression tag	UNP Q8IFM8
J	365	GLY	-	expression tag	UNP Q8IFM8
J	366	GLY	-	expression tag	UNP Q8IFM8
J	367	SER	-	expression tag	UNP Q8IFM8
J	368	GLU	-	expression tag	UNP Q8IFM8
J	369	PRO	-	expression tag	UNP Q8IFM8
J	370	GLU	-	expression tag	UNP Q8IFM8
J	371	ALA	-	expression tag	UNP Q8IFM8

- Molecule 2 is a protein called Monoclonal antibody Cy.003 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	222	Total	C	N	O	S	0	0	0
			1616	1012	277	320	7			
2	E	223	Total	C	N	O	S	0	0	0
			1622	1015	278	322	7			
2	H	225	Total	C	N	O	S	0	0	0
			1637	1024	281	325	7			
2	K	225	Total	C	N	O	S	0	0	0
			1637	1024	281	325	7			

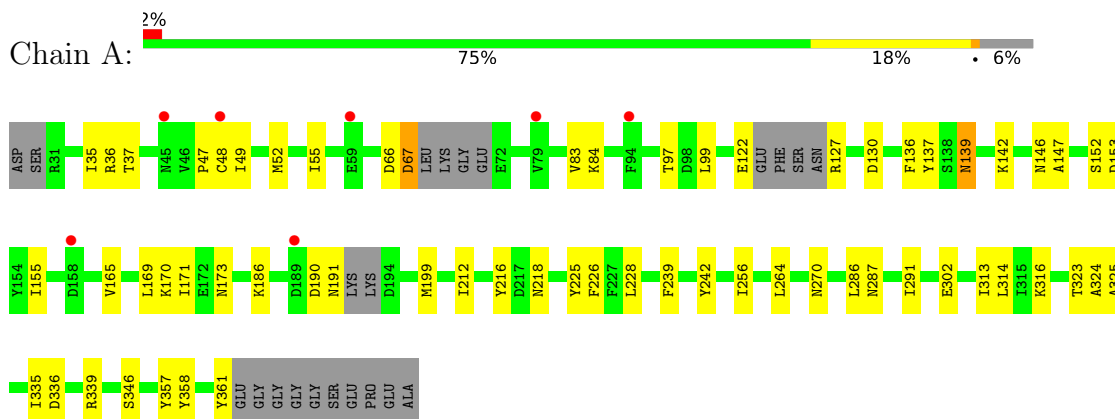
- Molecule 3 is a protein called Monoclonal antibody Cy.003 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	208	Total 1556	C 965	N 264	O 323	S 4	0	0	0
3	F	207	Total 1552	C 963	N 263	O 322	S 4	0	0	0
3	I	208	Total 1556	C 965	N 264	O 323	S 4	0	0	0
3	L	207	Total 1552	C 963	N 263	O 322	S 4	0	0	0

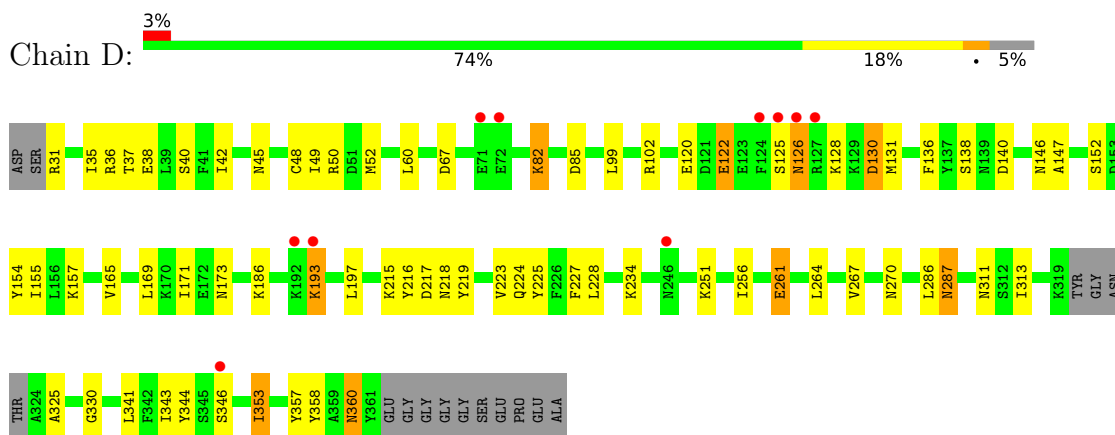
3 Residue-property plots [i](#)

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.

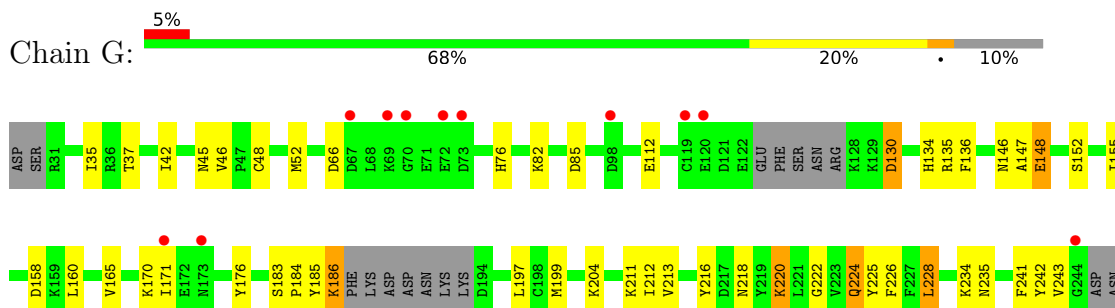
- Molecule 1: Cysteine-rich protective antigen

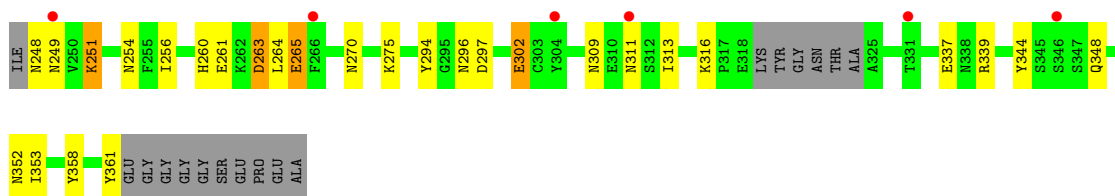


- Molecule 1: Cysteine-rich protective antigen

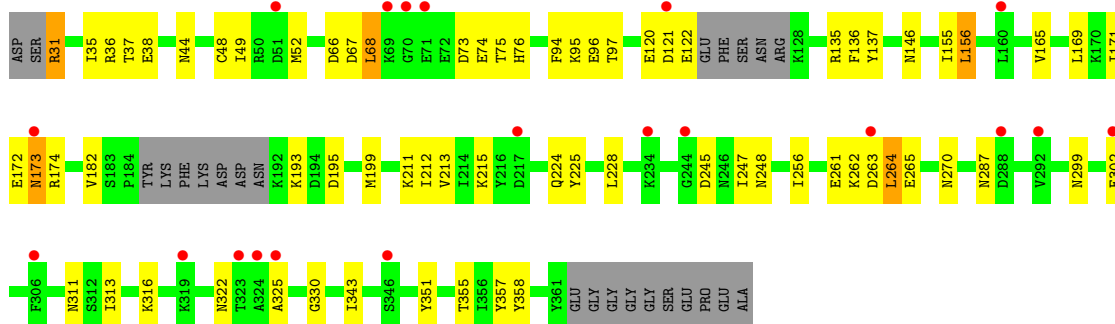
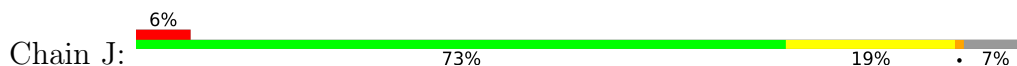


- Molecule 1: Cysteine-rich protective antigen





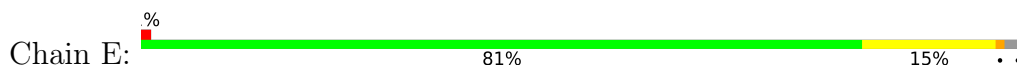
- Molecule 1: Cysteine-rich protective antigen



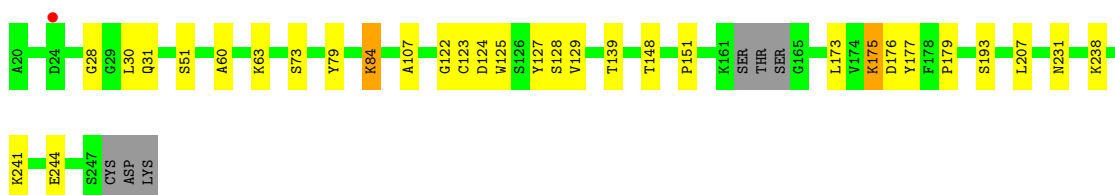
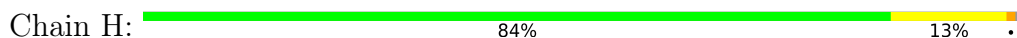
- Molecule 2: Monoclonal antibody Cy.003 heavy chain



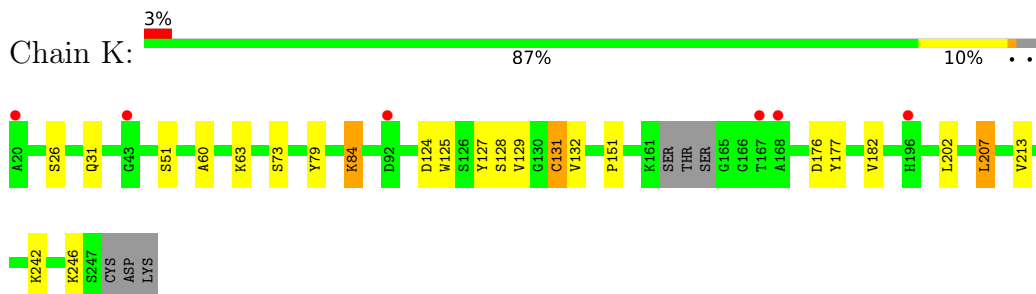
- Molecule 2: Monoclonal antibody Cy.003 heavy chain



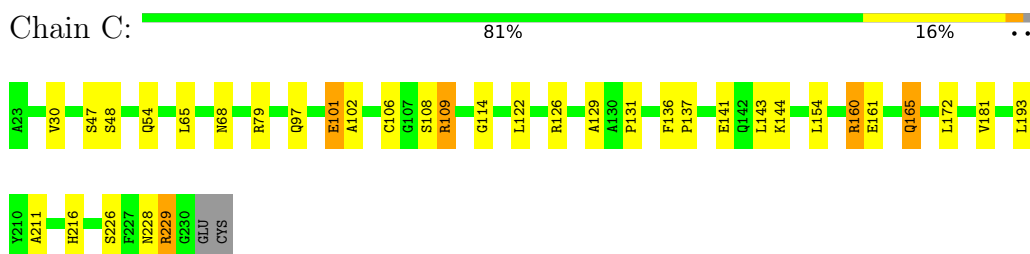
- Molecule 2: Monoclonal antibody Cy.003 heavy chain



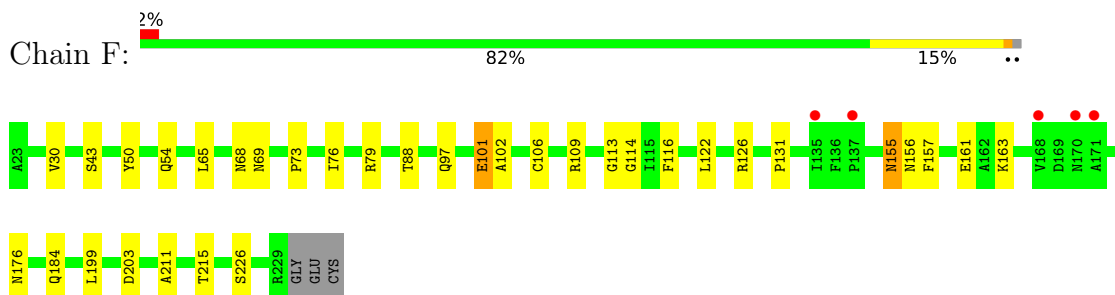
- Molecule 2: Monoclonal antibody Cy.003 heavy chain



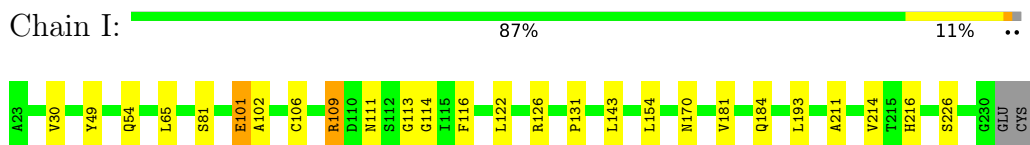
- Molecule 3: Monoclonal antibody Cy.003 light chain



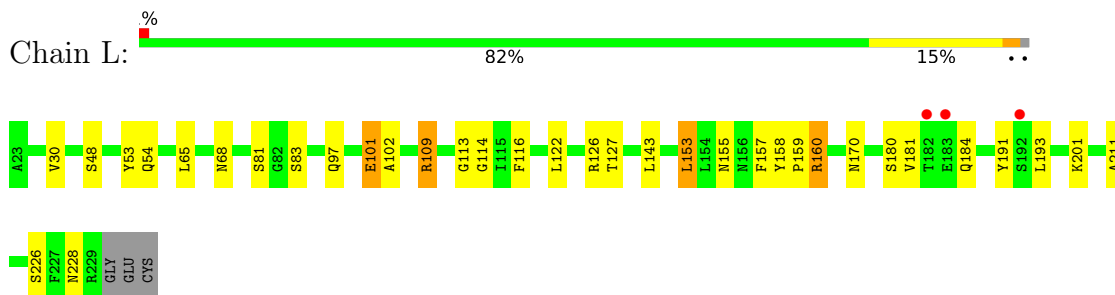
- Molecule 3: Monoclonal antibody Cy.003 light chain



- Molecule 3: Monoclonal antibody Cy.003 light chain



- Molecule 3: Monoclonal antibody Cy.003 light chain



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	354.30Å 71.05Å 164.79Å 90.00° 91.16° 90.00°	Depositor
Resolution (Å)	88.56 – 3.14 88.56 – 3.14	Depositor EDS
% Data completeness (in resolution range)	99.8 (88.56-3.14) 99.8 (88.56-3.14)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.61 (at 3.13Å)	Xtrriage
Refinement program	BUSTER 2.10.3 (19-MAR-2020)	Depositor
R, R_{free}	0.219 , 0.239 0.226 , 0.248	Depositor DCC
R_{free} test set	3461 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å ²)	59.8	Xtrriage
Anisotropy	0.791	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.014 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	23353	wwPDB-VP
Average B, all atoms (Å ²)	89.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.63% 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	0.44	0/2734	0.64	0/3691
1	D	0.42	0/2787	0.64	0/3761
1	G	0.48	0/2637	0.66	0/3558
1	J	0.44	0/2705	0.64	0/3652
2	B	0.42	0/1653	0.62	0/2250
2	E	0.45	0/1659	0.65	0/2258
2	H	0.46	0/1674	0.66	0/2277
2	K	0.51	0/1674	0.68	0/2277
3	C	0.46	0/1588	0.62	0/2158
3	F	0.38	0/1584	0.60	0/2153
3	I	0.50	0/1588	0.64	0/2158
3	L	0.44	0/1584	0.62	0/2153
All	All	0.45	0/23867	0.64	0/32346

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	2674	0	2560	50	0
1	D	2725	0	2622	43	0
1	G	2580	0	2479	49	0
1	J	2646	0	2549	40	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1616	0	1566	15	0
2	E	1622	0	1571	33	0
2	H	1637	0	1589	31	0
2	K	1637	0	1589	30	0
3	C	1556	0	1502	25	0
3	F	1552	0	1499	21	0
3	I	1556	0	1502	20	0
3	L	1552	0	1499	37	0
All	All	23353	0	22527	346	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 346 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:A:55:ILE:HG13	1:A:335:ILE:HG22	1.26	1.18
1:G:186:LYS:HB2	1:G:222:GLY:O	1.43	1.16
2:K:132:VAL:HG12	3:L:53:TYR:OH	1.59	1.03
3:F:163:LYS:HG3	3:F:215:THR:HB	1.41	1.02
3:L:160:ARG:HB3	3:L:191:TYR:CD2	1.98	0.98

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	313/343 (91%)	292 (93%)	21 (7%)	0	100 100
1	D	323/343 (94%)	302 (94%)	21 (6%)	0	100 100
1	G	300/343 (88%)	271 (90%)	29 (10%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	J	313/343 (91%)	286 (91%)	27 (9%)	0	100	100
2	B	218/231 (94%)	211 (97%)	7 (3%)	0	100	100
2	E	219/231 (95%)	212 (97%)	7 (3%)	0	100	100
2	H	221/231 (96%)	212 (96%)	9 (4%)	0	100	100
2	K	221/231 (96%)	211 (96%)	10 (4%)	0	100	100
3	C	206/210 (98%)	194 (94%)	12 (6%)	0	100	100
3	F	205/210 (98%)	194 (95%)	11 (5%)	0	100	100
3	I	206/210 (98%)	194 (94%)	12 (6%)	0	100	100
3	L	205/210 (98%)	196 (96%)	9 (4%)	0	100	100
All	All	2950/3136 (94%)	2775 (94%)	175 (6%)	0	100	100

There are no Ramachandran outliers to report.

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	300/316 (95%)	293 (98%)	7 (2%)	50	75
1	D	306/316 (97%)	285 (93%)	21 (7%)	15	43
1	G	290/316 (92%)	265 (91%)	25 (9%)	10	35
1	J	297/316 (94%)	278 (94%)	19 (6%)	17	46
2	B	177/186 (95%)	171 (97%)	6 (3%)	37	67
2	E	178/186 (96%)	170 (96%)	8 (4%)	27	59
2	H	180/186 (97%)	174 (97%)	6 (3%)	38	68
2	K	180/186 (97%)	173 (96%)	7 (4%)	32	63
3	C	177/179 (99%)	167 (94%)	10 (6%)	21	50
3	F	177/179 (99%)	169 (96%)	8 (4%)	27	59
3	I	177/179 (99%)	169 (96%)	8 (4%)	27	59
3	L	177/179 (99%)	166 (94%)	11 (6%)	18	47

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2616/2724 (96%)	2480 (95%)	136 (5%)	23 53

5 of 136 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	J	299	ASN
2	K	31	GLN
3	L	109	ARG
2	E	133	ASP
2	E	131	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	228	ASN
2	K	101	GLN
1	G	224	GLN
1	J	202	HIS
1	G	173	ASN

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](#)

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	321/343 (93%)	0.48	7 (2%) 62 42	53, 95, 155, 222	0
1	D	327/343 (95%)	0.33	10 (3%) 49 27	55, 92, 155, 186	0
1	G	310/343 (90%)	0.51	17 (5%) 25 11	54, 90, 133, 182	0
1	J	319/343 (93%)	0.64	20 (6%) 20 9	49, 95, 156, 187	0
2	B	222/231 (96%)	0.17	0 100 100	50, 74, 101, 112	0
2	E	223/231 (96%)	0.37	3 (1%) 77 61	36, 110, 150, 160	0
2	H	225/231 (97%)	0.09	1 (0%) 92 86	41, 68, 89, 118	0
2	K	225/231 (97%)	0.39	8 (3%) 42 22	42, 73, 133, 160	0
3	C	208/210 (99%)	0.14	0 100 100	47, 76, 104, 114	0
3	F	207/210 (98%)	0.48	5 (2%) 59 38	50, 104, 200, 211	0
3	I	208/210 (99%)	0.02	0 100 100	40, 66, 85, 103	0
3	L	207/210 (98%)	0.12	3 (1%) 75 59	40, 79, 117, 122	0
All	All	3002/3136 (95%)	0.33	74 (2%) 57 37	36, 82, 150, 222	0

The worst 5 of 74 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	J	71	GLU	5.6
1	G	67	ASP	5.5
1	G	346	SER	5.3
1	J	346	SER	4.6
3	F	168	VAL	4.3

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.