



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 26, 2020 – 07:07 pm BST

PDB ID : 6IW2  
Title : Crystal structure of 5A ScFv in complex with YFV-17D sE in prefusion state  
Authors : Lu, X.S.; Xiao, H.X.; Li, S.H.; Pang, X.F.  
Deposited on : 2018-12-04  
Resolution : 2.90 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 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 : 2.11  
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) : 2.11

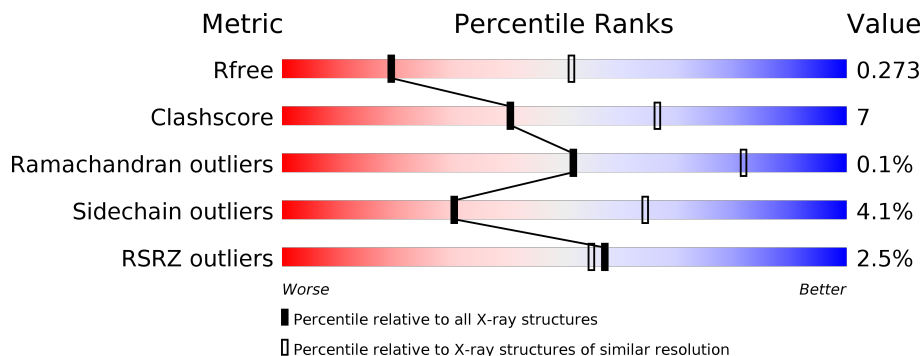
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 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	395	 2% (red), 76% (green), 19% (yellow), 2% (grey)
1	D	395	 2% (red), 75% (green), 21% (yellow), 2% (grey)
1	G	395	 2% (red), 77% (green), 20% (yellow), 2% (grey)
1	J	395	 3% (red), 79% (green), 17% (yellow), 2% (grey)
1	M	395	 2% (red), 78% (green), 18% (yellow), 2% (grey)
1	P	395	 3% (red), 78% (green), 19% (yellow), 2% (grey)

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	B	120	<p>2% 79% 21%</p>
2	E	120	<p>3% 71% 29%</p>
2	H	120	<p>3% 77% 23%</p>
2	K	120	<p>4% 81% 19%</p>
2	N	120	<p>2% 72% 28%</p>
2	Q	120	<p>3% 80% 20%</p>
3	C	107	<p>2% 85% 15%</p>
3	F	107	<p>7% 91% 9%</p>
3	I	107	<p>2% 82% 16%</p>
3	L	107	<p>5% 81% 18%</p>
3	O	107	<p>1% 84% 14%</p>
3	R	107	<p>4% 98%</p>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 28516 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 Envelope protein E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	386	2965	1863	514	569	19	0	0	0
1	D	388	2986	1877	518	572	19	0	1	0
1	G	388	2979	1872	516	572	19	0	0	0
1	J	387	2983	1877	516	571	19	0	1	0
1	M	387	2972	1868	515	570	19	0	0	0
1	P	387	2972	1868	515	570	19	0	0	0

- Molecule 2 is a protein called Heavy chain of monoclonal antibody 5A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	120	954	606	158	187	3	0	0	0
2	E	120	954	606	158	187	3	0	0	0
2	H	120	954	606	158	187	3	0	0	0
2	K	120	954	606	158	187	3	0	0	0
2	N	120	960	610	158	189	3	0	1	0
2	Q	120	954	606	158	187	3	0	0	0

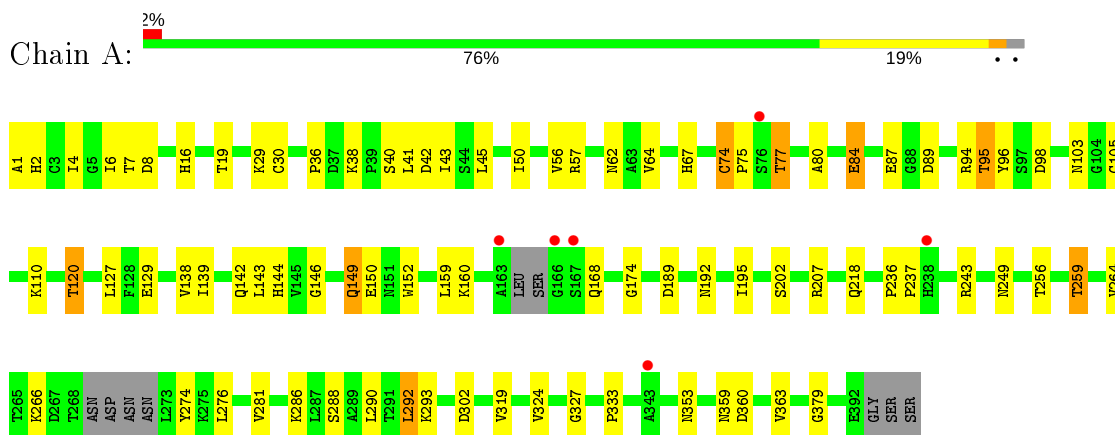
- Molecule 3 is a protein called Light chain of monoclonal antibody 5A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	107	Total	C	N	O	S	0	0	0
			823	523	134	163	3			
3	F	107	Total	C	N	O	S	0	0	0
			823	523	134	163	3			
3	I	107	Total	C	N	O	S	0	0	0
			823	523	134	163	3			
3	L	106	Total	C	N	O	S	0	0	0
			814	518	132	161	3			
3	O	107	Total	C	N	O	S	0	0	0
			823	523	134	163	3			
3	R	107	Total	C	N	O	S	0	0	0
			823	523	134	163	3			

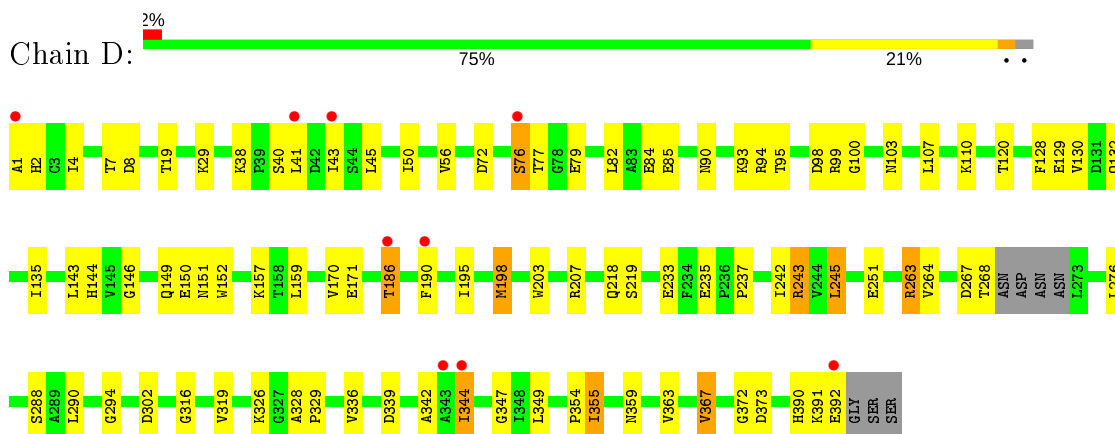
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

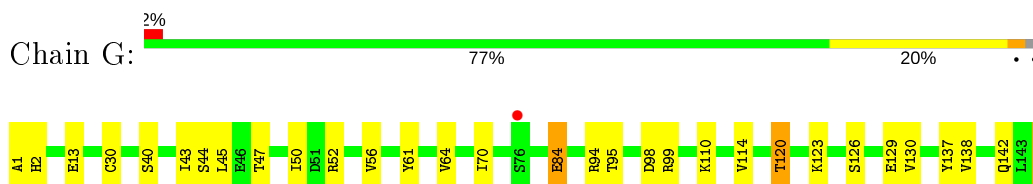
- Molecule 1: Envelope protein E

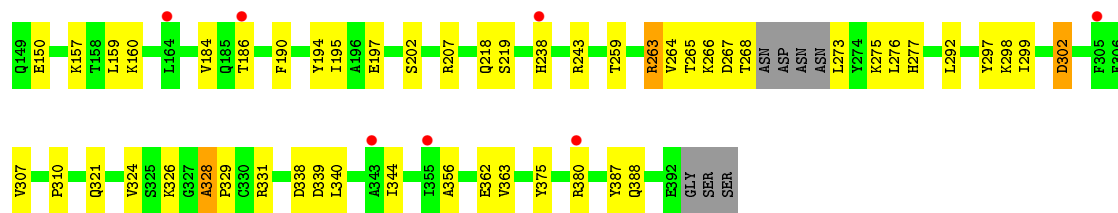


- Molecule 1: Envelope protein E

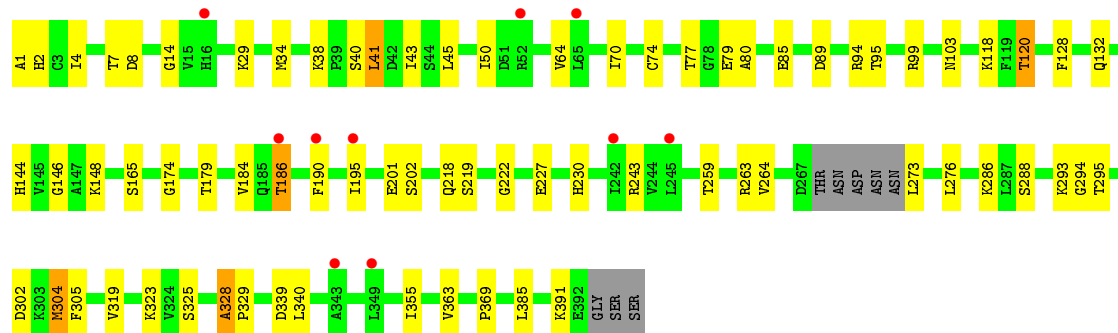
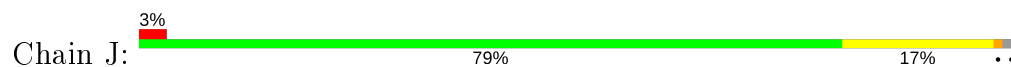


- Molecule 1: Envelope protein E

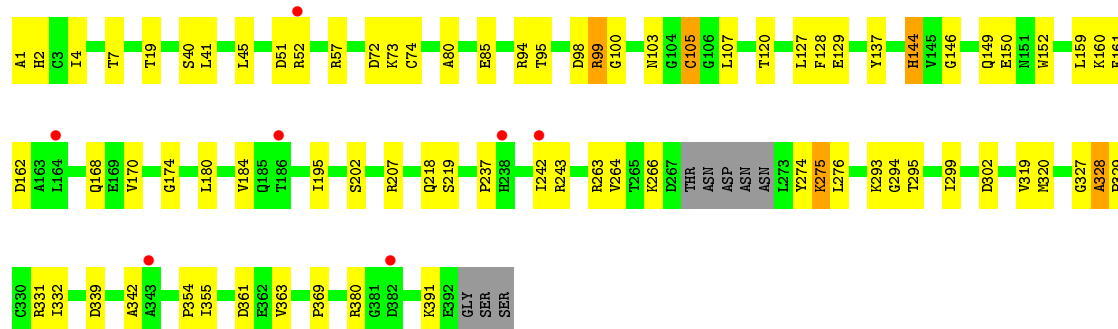
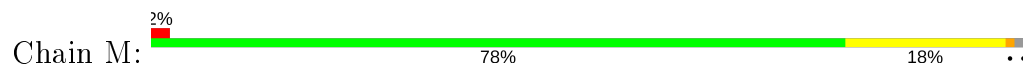




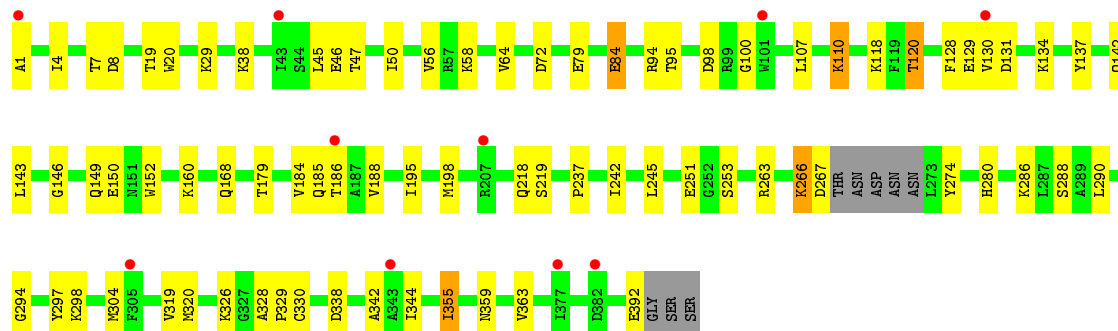
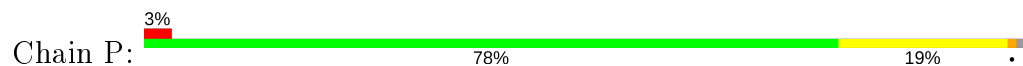
● Molecule 1: Envelope protein E



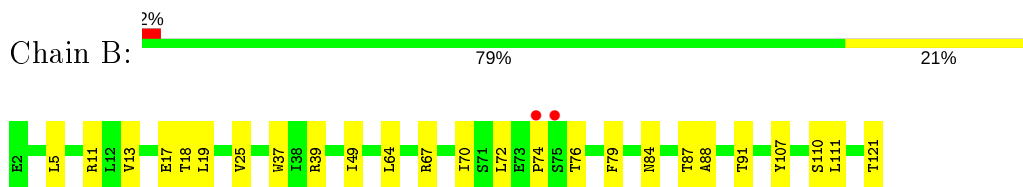
● Molecule 1: Envelope protein E



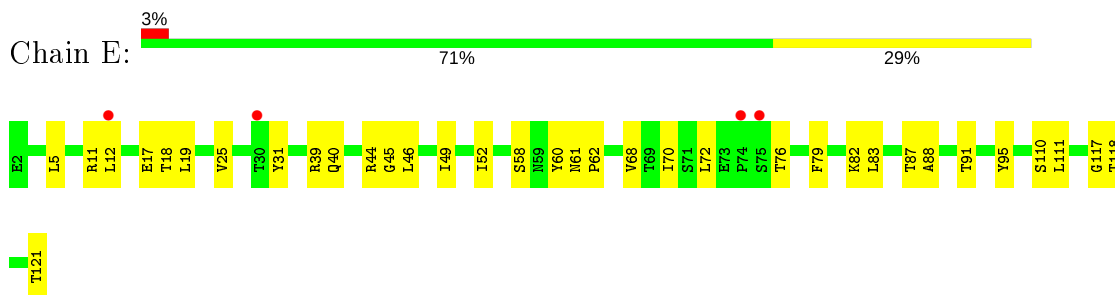
● Molecule 1: Envelope protein E



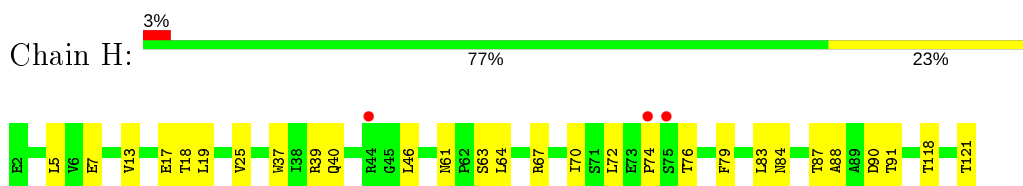
- Molecule 2: Heavy chain of monoclonal antibody 5A



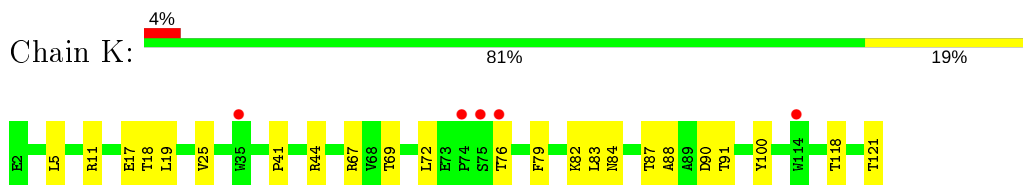
- Molecule 2: Heavy chain of monoclonal antibody 5A



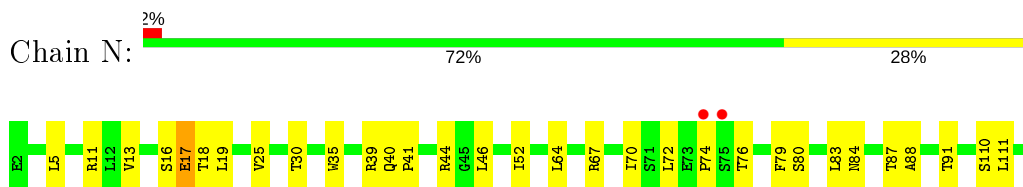
- Molecule 2: Heavy chain of monoclonal antibody 5A



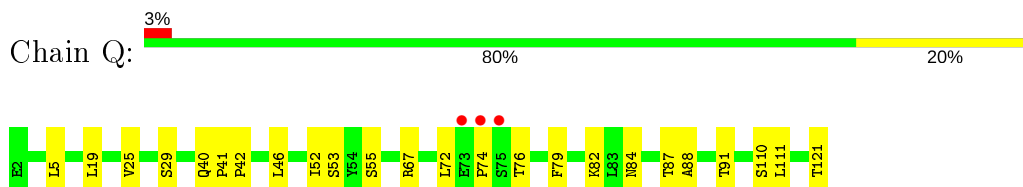
- Molecule 2: Heavy chain of monoclonal antibody 5A



- Molecule 2: Heavy chain of monoclonal antibody 5A

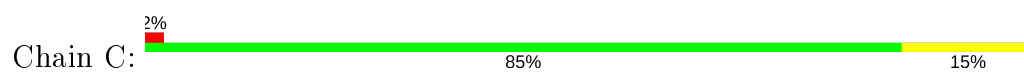


- Molecule 2: Heavy chain of monoclonal antibody 5A

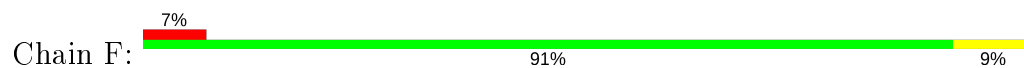


- Molecule 3: Light chain of monoclonal antibody 5A

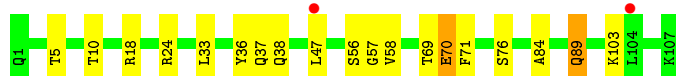
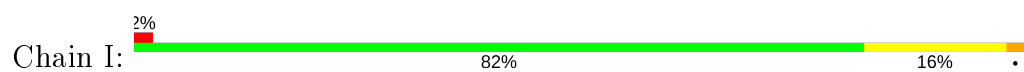




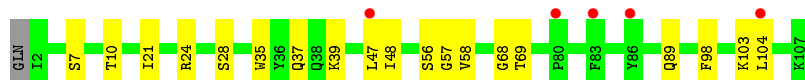
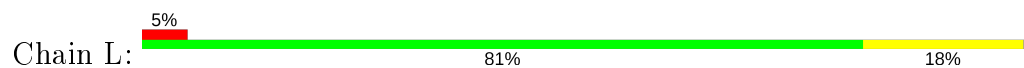
- Molecule 3: Light chain of monoclonal antibody 5A



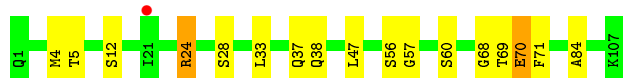
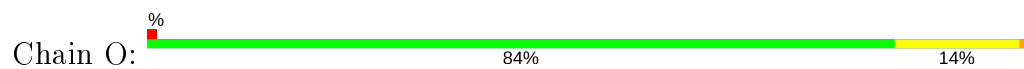
- Molecule 3: Light chain of monoclonal antibody 5A



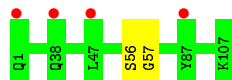
- Molecule 3: Light chain of monoclonal antibody 5A



- Molecule 3: Light chain of monoclonal antibody 5A



- Molecule 3: Light chain of monoclonal antibody 5A



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	95.97Å 99.24Å 123.10Å 96.39° 94.90° 100.36°	Depositor
Resolution (Å)	48.70 – 2.90 48.69 – 2.89	Depositor EDS
% Data completeness (in resolution range)	85.2 (48.70-2.90) 77.6 (48.69-2.89)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.50 (at 2.91Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.237 , 0.272 0.240 , 0.273	Depositor DCC
$R_{free}$ test set	4170 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.9	Xtrriage
Anisotropy	0.113	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 0.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.106 for -k,-h,-l	Xtrriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	28516	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.99% 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	0.26	0/3023	0.48	0/4098
1	D	0.31	0/3049	0.54	1/4135 (0.0%)
1	G	0.29	0/3038	0.54	0/4120
1	J	0.27	0/3043	0.52	1/4126 (0.0%)
1	M	0.27	0/3031	0.48	0/4110
1	P	0.26	0/3031	0.52	0/4110
2	B	0.23	0/982	0.47	0/1342
2	E	0.29	0/982	0.52	0/1342
2	H	0.23	0/982	0.45	0/1342
2	K	0.24	0/982	0.47	0/1342
2	N	0.23	0/991	0.48	0/1355
2	Q	0.22	0/982	0.46	0/1342
3	C	0.23	0/843	0.43	0/1144
3	F	0.23	0/843	0.43	0/1144
3	I	0.25	0/843	0.46	0/1144
3	L	0.26	0/834	0.45	0/1132
3	O	0.32	0/843	0.55	1/1144 (0.1%)
3	R	0.23	0/843	0.44	0/1144
All	All	0.27	0/29165	0.50	3/39616 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	O	24	ARG	NE-CZ-NH1	7.07	123.83	120.30
1	D	186	THR	CA-CB-CG2	-5.25	105.06	112.40
1	J	222	GLY	N-CA-C	-5.24	100.00	113.10

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	2965	0	2915	47	0
1	D	2986	0	2938	59	0
1	G	2979	0	2931	47	0
1	J	2983	0	2932	38	0
1	M	2972	0	2925	43	0
1	P	2972	0	2924	45	0
2	B	954	0	907	13	0
2	E	954	0	907	19	0
2	H	954	0	907	16	0
2	K	954	0	907	12	0
2	N	960	0	913	19	0
2	Q	954	0	907	13	0
3	C	823	0	807	8	0
3	F	823	0	807	4	0
3	I	823	0	807	13	0
3	L	814	0	796	10	0
3	O	823	0	807	9	0
3	R	823	0	807	2	0
All	All	28516	0	27844	405	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 7.

The worst 5 of 405 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:M:74:CYS:CB	1:M:105:CYS:SG	2.04	1.45
1:M:74:CYS:SG	1:M:105:CYS:SG	1.43	1.43
1:D:186:THR:HG21	1:D:190:PHE:CE2	2.04	0.93
1:D:186:THR:HG21	1:D:190:PHE:HE2	1.34	0.91
3:I:24:ARG:HH11	3:I:69:THR:HG22	1.36	0.88

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	380/395 (96%)	360 (95%)	20 (5%)	0	100	100
1	D	385/395 (98%)	358 (93%)	27 (7%)	0	100	100
1	G	384/395 (97%)	359 (94%)	24 (6%)	1 (0%)	41	71
1	J	384/395 (97%)	359 (94%)	24 (6%)	1 (0%)	41	71
1	M	383/395 (97%)	358 (94%)	23 (6%)	2 (0%)	29	61
1	P	383/395 (97%)	360 (94%)	23 (6%)	0	100	100
2	B	118/120 (98%)	113 (96%)	5 (4%)	0	100	100
2	E	118/120 (98%)	113 (96%)	5 (4%)	0	100	100
2	H	118/120 (98%)	113 (96%)	5 (4%)	0	100	100
2	K	118/120 (98%)	113 (96%)	5 (4%)	0	100	100
2	N	119/120 (99%)	116 (98%)	3 (2%)	0	100	100
2	Q	118/120 (98%)	113 (96%)	5 (4%)	0	100	100
3	C	105/107 (98%)	101 (96%)	4 (4%)	0	100	100
3	F	105/107 (98%)	100 (95%)	5 (5%)	0	100	100
3	I	105/107 (98%)	101 (96%)	4 (4%)	0	100	100
3	L	104/107 (97%)	99 (95%)	5 (5%)	0	100	100
3	O	105/107 (98%)	100 (95%)	5 (5%)	0	100	100
3	R	105/107 (98%)	99 (94%)	6 (6%)	0	100	100
All	All	3637/3732 (98%)	3435 (94%)	198 (5%)	4 (0%)	51	82

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	M	327	GLY
1	M	328	ALA
1	J	328	ALA
1	G	328	ALA

### 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	322/331 (97%)	305 (95%)	17 (5%)	22	54
1	D	325/331 (98%)	306 (94%)	19 (6%)	20	50
1	G	324/331 (98%)	301 (93%)	23 (7%)	14	40
1	J	324/331 (98%)	307 (95%)	17 (5%)	23	55
1	M	323/331 (98%)	306 (95%)	17 (5%)	22	54
1	P	323/331 (98%)	309 (96%)	14 (4%)	29	62
2	B	107/107 (100%)	105 (98%)	2 (2%)	57	84
2	E	107/107 (100%)	106 (99%)	1 (1%)	78	93
2	H	107/107 (100%)	106 (99%)	1 (1%)	78	93
2	K	107/107 (100%)	106 (99%)	1 (1%)	78	93
2	N	108/107 (101%)	104 (96%)	4 (4%)	34	68
2	Q	107/107 (100%)	106 (99%)	1 (1%)	78	93
3	C	92/92 (100%)	90 (98%)	2 (2%)	52	81
3	F	92/92 (100%)	90 (98%)	2 (2%)	52	81
3	I	92/92 (100%)	90 (98%)	2 (2%)	52	81
3	L	91/92 (99%)	90 (99%)	1 (1%)	73	92
3	O	92/92 (100%)	89 (97%)	3 (3%)	38	72
3	R	92/92 (100%)	92 (100%)	0	100	100
All	All	3135/3180 (99%)	3008 (96%)	127 (4%)	30	64

5 of 127 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	238	HIS
1	J	74	CYS
1	P	150	GLU
1	G	259	THR
1	G	302	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
2	H	61	ASN
1	J	132	GLN
1	J	390	HIS
1	M	67	HIS

### 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 carbohydrates 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

### 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	A	386/395 (97%)	0.04	6 (1%) 72 71	32, 46, 77, 95	0
1	D	388/395 (98%)	0.16	9 (2%) 60 58	32, 50, 87, 103	0
1	G	388/395 (98%)	0.20	8 (2%) 63 61	31, 48, 80, 111	0
1	J	387/395 (97%)	0.11	10 (2%) 56 52	31, 46, 73, 104	0
1	M	387/395 (97%)	0.05	7 (1%) 68 67	32, 47, 77, 101	0
1	P	387/395 (97%)	0.17	10 (2%) 56 52	31, 48, 83, 104	0
2	B	120/120 (100%)	-0.05	2 (1%) 70 69	30, 46, 71, 90	0
2	E	120/120 (100%)	0.13	4 (3%) 46 41	35, 58, 95, 107	0
2	H	120/120 (100%)	0.26	3 (2%) 57 55	35, 59, 83, 100	0
2	K	120/120 (100%)	0.41	5 (4%) 36 32	32, 64, 97, 106	0
2	N	120/120 (100%)	0.03	2 (1%) 70 69	32, 45, 68, 78	0
2	Q	120/120 (100%)	0.15	3 (2%) 57 55	35, 60, 85, 94	0
3	C	107/107 (100%)	0.20	2 (1%) 66 65	35, 52, 77, 89	0
3	F	107/107 (100%)	0.39	8 (7%) 14 11	40, 57, 84, 98	0
3	I	107/107 (100%)	0.30	2 (1%) 66 65	38, 53, 71, 76	0
3	L	106/107 (99%)	0.35	5 (4%) 31 28	39, 59, 82, 113	0
3	O	107/107 (100%)	0.02	1 (0%) 84 84	33, 46, 69, 82	0
3	R	107/107 (100%)	0.30	4 (3%) 41 37	35, 50, 76, 91	0
All	All	3684/3732 (98%)	0.15	91 (2%) 57 55	30, 49, 83, 113	0

The worst 5 of 91 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	75	SER	6.0
3	F	84	ALA	5.5
2	E	74	PRO	5.5

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	P	343	ALA	5.1
3	F	86	TYR	4.9

## 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 carbohydrates 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.