



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 01:18 pm BST

PDB ID : 3GJF  
Title : Rational development of high-affinity T-cell receptor-like antibodies  
Authors : Stewart-Jones, G.; Wadle, A.; Hombach, A.; Shenderov, E.; Held, G.; Fischer, E.; Kleber, S.; Stenner-Liewen, F.; Bauer, S.; McMichael, A.; Knuth, A.; Abken, H.; Hombach, A.A.; Cerundolo, V.; Jones, E.Y.; Renner, C.  
Deposited on : 2009-03-08  
Resolution : 1.90 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

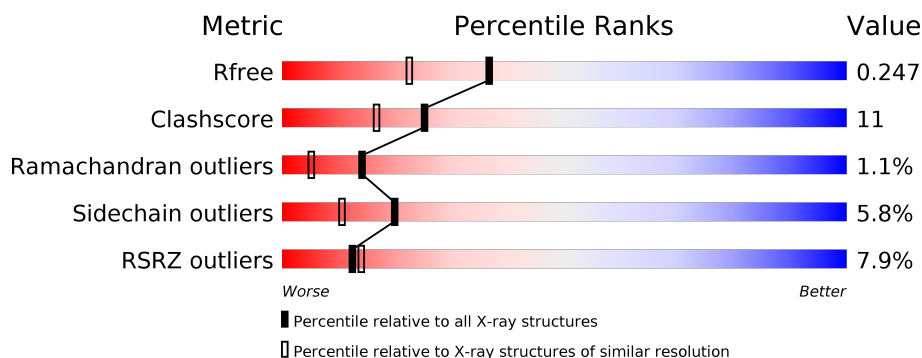
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	276	<div> <div>4%</div> <div> <div></div> <div>84%</div> <div>14%</div> <div>.</div> </div> </div>
1	D	276	<div> <div>3%</div> <div> <div></div> <div>83%</div> <div>15%</div> <div>.</div> </div> </div>
2	B	100	<div> <div>5%</div> <div> <div></div> <div>81%</div> <div>14%</div> <div>5%</div> </div> </div>
2	E	100	<div> <div>5%</div> <div> <div></div> <div>78%</div> <div>18%</div> <div>.</div> </div> </div>
3	C	9	<div> <div>11%</div> <div> <div></div> <div>56%</div> <div>33%</div> <div>11%</div> </div> </div>
3	F	9	<div> <div></div> <div> <div></div> <div>89%</div> <div>11%</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	K	212	<div><div></div><div>4%</div><div>84%</div><div>14%</div><div></div><div></div></div>
4	L	212	<div><div></div><div>15%</div><div>75%</div><div>14%</div><div>5%</div><div></div><div></div></div>
5	H	220	<div><div></div><div>22%</div><div>81%</div><div>16%</div><div></div><div></div></div>
5	M	220	<div><div></div><div>5%</div><div>90%</div><div>8%</div><div></div><div></div></div>

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 14331 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 HLA class I histocompatibility antigen, A-2 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	276	Total	C	N	O	S	0	0	0
			2253	1408	410	426	9			
1	D	276	Total	C	N	O	S	0	0	0
			2253	1408	410	426	9			

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	100	Total	C	N	O	S	0	0	0
			836	533	141	158	4			
2	E	100	Total	C	N	O	S	0	0	0
			836	533	141	158	4			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	EXPRESSION TAG	UNP P61769
E	0	MET	-	EXPRESSION TAG	UNP P61769

- Molecule 3 is a protein called NYESO-1 peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	S	0	0	0
			75	51	11	12	1			
3	F	9	Total	C	N	O	S	0	0	0
			75	51	11	12	1			

- Molecule 4 is a protein called Antibody light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	L	206	Total	C	N	O	S	0	0	0
			1559	976	262	317	4			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	K	212	Total	C	N	O	S	0	0	0
			1600	1001	269	326	4			

- Molecule 5 is a protein called Antibody heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	H	220	Total	C	N	O	S	0	0	0
			1621	1020	269	325	7			
5	M	220	Total	C	N	O	S	0	0	0
			1620	1020	269	324	7			

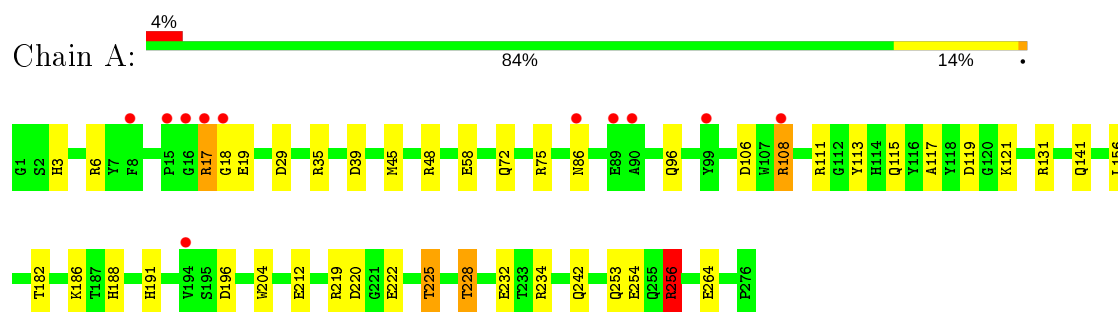
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	289	Total	O	0	0
			289	289		
6	B	144	Total	O	0	0
			144	144		
6	C	9	Total	O	0	0
			9	9		
6	D	291	Total	O	0	0
			291	291		
6	E	104	Total	O	0	0
			104	104		
6	F	7	Total	O	0	0
			7	7		
6	L	172	Total	O	0	0
			172	172		
6	K	178	Total	O	0	0
			178	178		
6	H	189	Total	O	0	0
			189	189		
6	M	220	Total	O	0	0
			220	220		

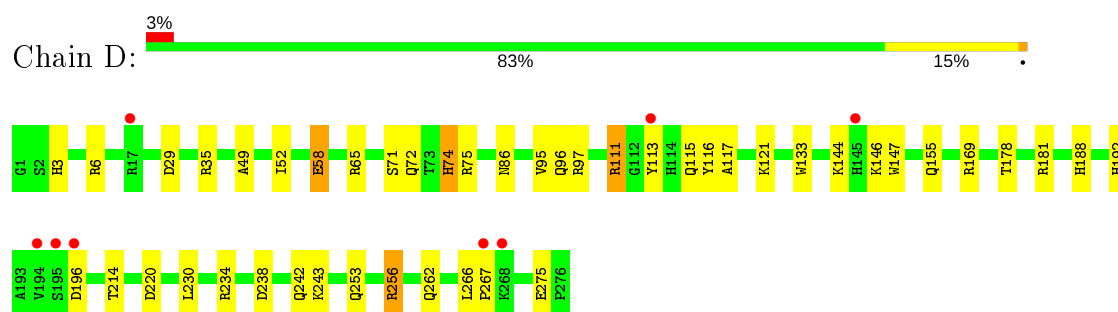
### 3 Residue-property plots [i](#)

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.

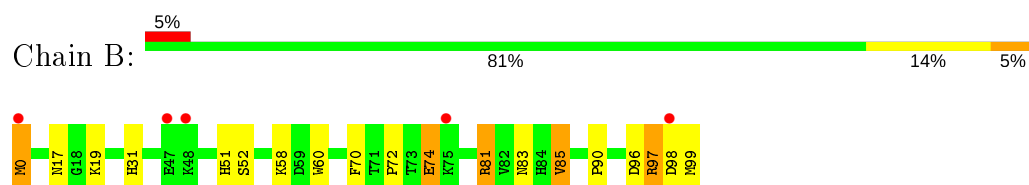
- Molecule 1: HLA class I histocompatibility antigen, A-2 alpha chain



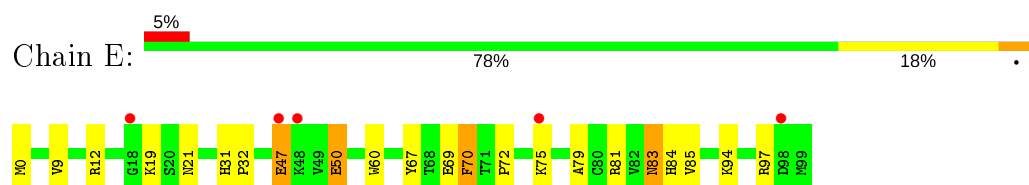
- Molecule 1: HLA class I histocompatibility antigen, A-2 alpha chain



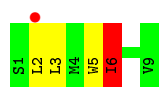
- Molecule 2: Beta-2-microglobulin



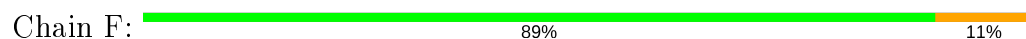
- Molecule 2: Beta-2-microglobulin



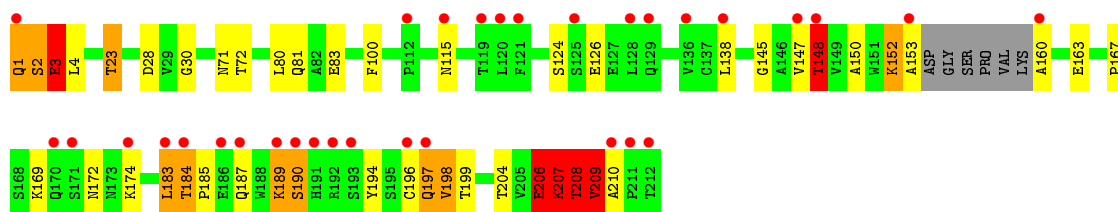
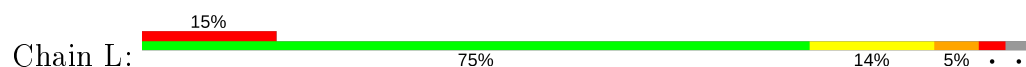
- Molecule 3: NYESO-1 peptide



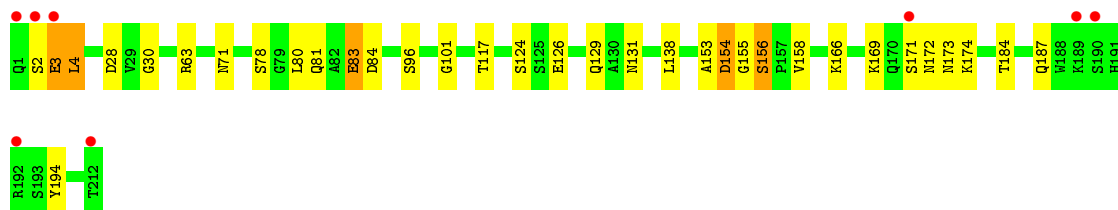
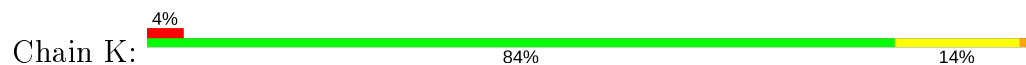
- Molecule 3: NYESO-1 peptide



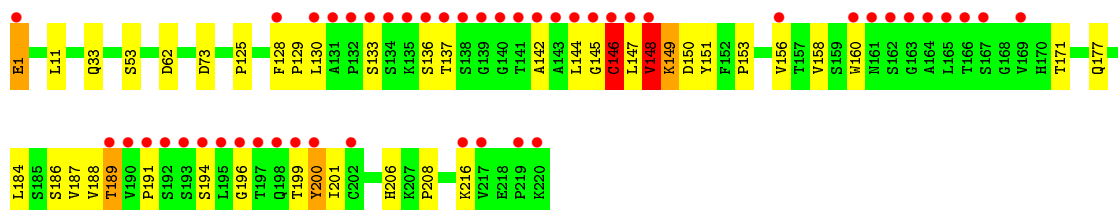
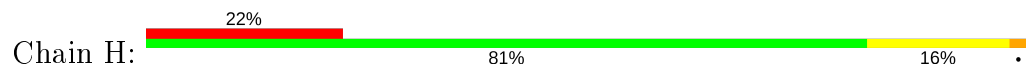
- Molecule 4: Antibody light chain



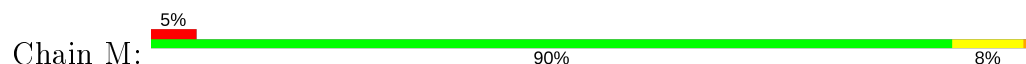
- Molecule 4: Antibody light chain

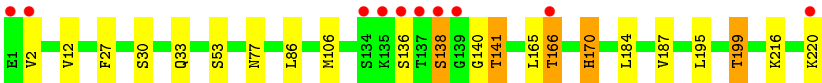


- Molecule 5: Antibody heavy chain



- Molecule 5: Antibody heavy chain







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.60Å 111.84Å 124.72Å 90.00° 93.49° 90.00°	Depositor
Resolution (Å)	25.00 – 1.90 29.98 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.7 (25.00-1.90) 99.7 (29.98-1.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.99 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.202 , 0.251 0.199 , 0.247	Depositor DCC
$R_{free}$ test set	7718 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	28.7	Xtriage
Anisotropy	0.206	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 46.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	14331	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.53% 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 ⓘ

### 5.1 Standard geometry ⓘ

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	0.68	0/2319	0.78	2/3149 (0.1%)
1	D	0.67	0/2319	0.76	2/3149 (0.1%)
2	B	0.75	0/859	0.71	0/1162
2	E	0.63	0/859	0.68	0/1162
3	C	0.65	0/76	1.15	1/103 (1.0%)
3	F	0.68	0/76	0.77	0/103
4	K	0.65	0/1641	0.73	1/2240 (0.0%)
4	L	0.64	0/1598	0.78	1/2180 (0.0%)
5	H	0.67	0/1658	0.85	3/2257 (0.1%)
5	M	0.66	0/1657	0.73	0/2255
All	All	0.67	0/13062	0.76	10/17760 (0.1%)

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	B	0	1
4	L	0	2
5	H	0	1
All	All	0	4

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	65	ARG	NE-CZ-NH2	-9.18	115.71	120.30
1	A	256	ARG	NE-CZ-NH1	-7.76	116.42	120.30
1	D	65	ARG	NE-CZ-NH1	6.25	123.42	120.30
5	H	150	ASP	N-CA-C	-6.00	94.80	111.00
5	H	146	CYS	C-N-CA	5.65	135.82	121.70

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	0	MET	Peptide
5	H	148	VAL	Peptide
4	L	2	SER	Peptide
4	L	206	GLU	Peptide

## 5.2 Too-close contacts

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	2253	0	2103	60	0
1	D	2253	0	2103	45	0
2	B	836	0	803	26	0
2	E	836	0	803	19	0
3	C	75	0	83	3	0
3	F	75	0	83	2	0
4	K	1600	0	1534	28	0
4	L	1559	0	1492	44	0
5	H	1621	0	1592	35	0
5	M	1620	0	1592	18	0
6	A	289	0	0	21	0
6	B	144	0	0	9	0
6	C	9	0	0	0	0
6	D	291	0	0	12	0
6	E	104	0	0	6	0
6	F	7	0	0	0	0
6	H	189	0	0	11	0
6	K	178	0	0	5	0
6	L	172	0	0	6	0
6	M	220	0	0	9	0
All	All	14331	0	12188	267	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 11.

The worst 5 of 267 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:106:ASP:OD1	1:A:108:ARG:HD2	1.40	1.21
1:A:35:ARG:NH1	1:A:48:ARG:HD3	1.56	1.20
1:A:256:ARG:CG	1:A:256:ARG:HH11	1.54	1.19
4:L:115:ASN:HB3	6:L:1463:HOH:O	1.51	1.10
1:D:256:ARG:HG3	1:D:256:ARG:HH11	1.08	1.10

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	274/276 (99%)	264 (96%)	9 (3%)	1 (0%)	34	24
1	D	274/276 (99%)	267 (97%)	7 (3%)	0	100	100
2	B	98/100 (98%)	97 (99%)	1 (1%)	0	100	100
2	E	98/100 (98%)	96 (98%)	2 (2%)	0	100	100
3	C	7/9 (78%)	7 (100%)	0	0	100	100
3	F	7/9 (78%)	7 (100%)	0	0	100	100
4	K	210/212 (99%)	202 (96%)	6 (3%)	2 (1%)	15	6
4	L	202/212 (95%)	182 (90%)	11 (5%)	9 (4%)	2	0
5	H	218/220 (99%)	203 (93%)	10 (5%)	5 (2%)	6	1
5	M	218/220 (99%)	214 (98%)	4 (2%)	0	100	100
All	All	1606/1634 (98%)	1539 (96%)	50 (3%)	17 (1%)	14	5

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	L	190	SER
4	L	197	GLN
4	L	208	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
5	H	149	LYS
5	H	156	VAL

### 5.3.2 Protein sidechains ⓘ

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	232/232 (100%)	219 (94%)	13 (6%)	21	11
1	D	232/232 (100%)	223 (96%)	9 (4%)	32	23
2	B	95/95 (100%)	89 (94%)	6 (6%)	18	8
2	E	95/95 (100%)	87 (92%)	8 (8%)	11	4
3	C	9/9 (100%)	8 (89%)	1 (11%)	6	2
3	F	9/9 (100%)	8 (89%)	1 (11%)	6	2
4	K	179/179 (100%)	170 (95%)	9 (5%)	24	15
4	L	174/179 (97%)	158 (91%)	16 (9%)	9	3
5	H	182/182 (100%)	174 (96%)	8 (4%)	28	19
5	M	182/182 (100%)	172 (94%)	10 (6%)	21	12
All	All	1389/1394 (100%)	1308 (94%)	81 (6%)	20	10

5 of 81 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	97	ARG
4	L	163	GLU
5	M	141	THR
3	F	6	ILE
4	L	28	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 44 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	74	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	242	GLN
5	H	33	GLN
1	D	155	GLN
1	D	188	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	276/276 (100%)	0.15	11 (3%) 38 41	16, 28, 43, 57	0
1	D	276/276 (100%)	0.08	8 (2%) 51 54	17, 28, 43, 55	0
2	B	100/100 (100%)	0.36	5 (5%) 28 32	19, 27, 45, 51	0
2	E	100/100 (100%)	0.33	5 (5%) 28 32	17, 34, 50, 57	0
3	C	9/9 (100%)	1.08	1 (11%) 5 6	23, 28, 33, 35	0
3	F	9/9 (100%)	0.94	0 100 100	20, 27, 31, 32	0
4	K	212/212 (100%)	0.30	8 (3%) 40 43	19, 31, 50, 55	0
4	L	206/212 (97%)	0.93	32 (15%) 2 2	16, 36, 69, 84	1 (0%)
5	H	220/220 (100%)	1.14	48 (21%) 0 0	17, 32, 82, 90	0
5	M	220/220 (100%)	0.25	10 (4%) 33 36	18, 29, 43, 66	0
All	All	1628/1634 (99%)	0.44	128 (7%) 12 14	16, 30, 58, 90	1 (0%)

The worst 5 of 128 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	H	195	LEU	10.7
5	H	190	VAL	10.7
4	L	196	CYS	9.3
4	L	212	THR	9.0
5	M	137	THR	8.7

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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.