



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2017 – 02:47 PM EDT

PDB ID : 5WUX  
Title : TNFalpha-certolizumab Fab  
Authors : Heo, Y.S.; Lee, J.U.  
Deposited on : 2016-12-21  
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

<http://wwpdb.org/validation/2016/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)	:	1.9-1692
EDS	:	rb-20029077
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20029077

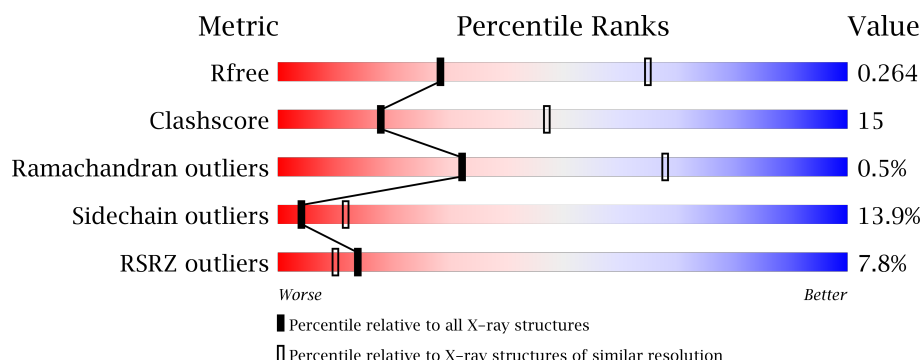
# 1 Overall quality at a glance

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.




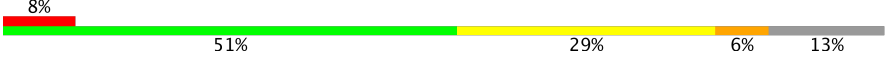

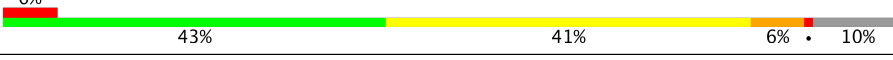
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	100719	1586 (2.90-2.90)
Clashscore	112137	1807 (2.90-2.90)
Ramachandran outliers	110173	1768 (2.90-2.90)
Sidechain outliers	110143	1770 (2.90-2.90)
RSRZ outliers	101464	1596 (2.90-2.90)

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. 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	224	<div> <div>2%</div> <div> <div></div> <div>63%</div> <div>27%</div> <div>•</div> <div>6%</div> </div> </div>
1	C	224	<div> <div>10%</div> <div> <div></div> <div>52%</div> <div>36%</div> <div>•</div> <div>10%</div> </div> </div>
1	H	224	<div> <div>7%</div> <div> <div></div> <div>58%</div> <div>33%</div> <div>•</div> <div>5%</div> </div> </div>
2	B	214	<div> <div>7%</div> <div> <div></div> <div>59%</div> <div>32%</div> <div>7%</div> <div>•</div> </div> </div>
2	D	214	<div> <div>17%</div> <div> <div></div> <div>54%</div> <div>31%</div> <div>8%</div> <div>•</div> <div>6%</div> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	L	214	 3% 69% 27% •
3	E	157	 8% 51% 29% 6% 13%
3	F	157	 6% 55% 31% • 10%
3	G	157	 6% 43% 41% 6% • 10%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 12793 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 heavy.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	210	Total	C	N	O	S	0	0	0
			1586	1010	260	308	8			
1	C	202	Total	C	N	O	S	0	0	0
			1530	976	251	295	8			
1	H	212	Total	C	N	O	S	0	0	0
			1601	1019	263	311	8			

- Molecule 2 is a protein called light.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	212	Total	C	N	O	S	0	0	0
			1632	1028	270	329	5			
2	D	201	Total	C	N	O	S	0	0	0
			1538	969	251	313	5			
2	L	213	Total	C	N	O	S	0	0	0
			1641	1033	271	332	5			

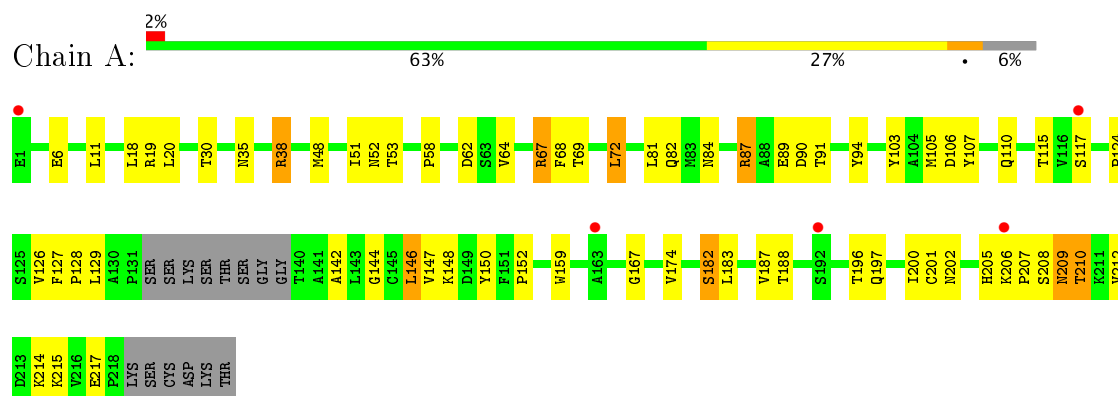
- Molecule 3 is a protein called Tumor necrosis factor alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	136	Total	C	N	O	S	0	0	0
			1057	681	177	197	2			
3	F	141	Total	C	N	O	S	0	0	0
			1104	708	190	204	2			
3	G	141	Total	C	N	O	S	0	0	0
			1104	708	190	204	2			

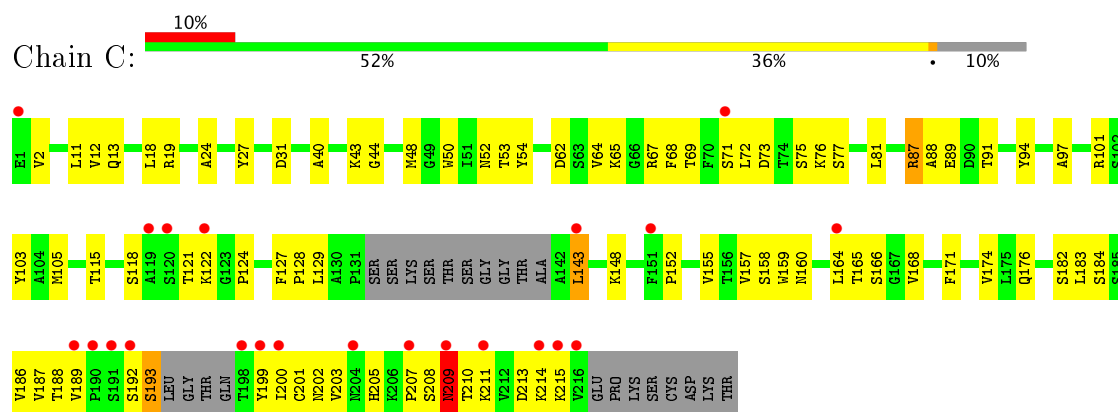
### 3 Residue-property plots [i](#)

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.

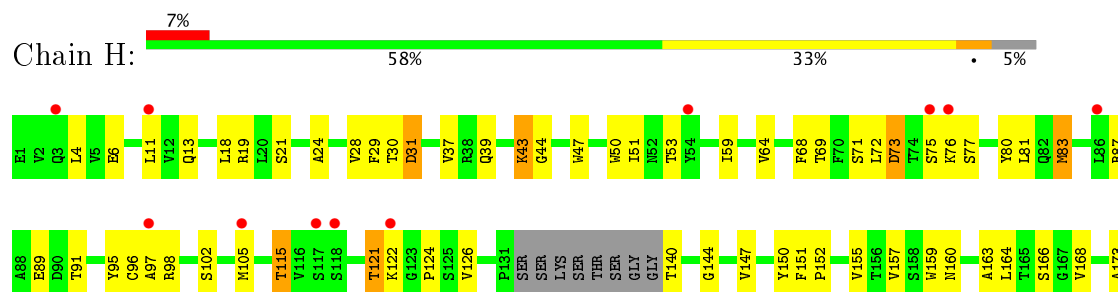
#### • Molecule 1: heavy



#### • Molecule 1: heavy

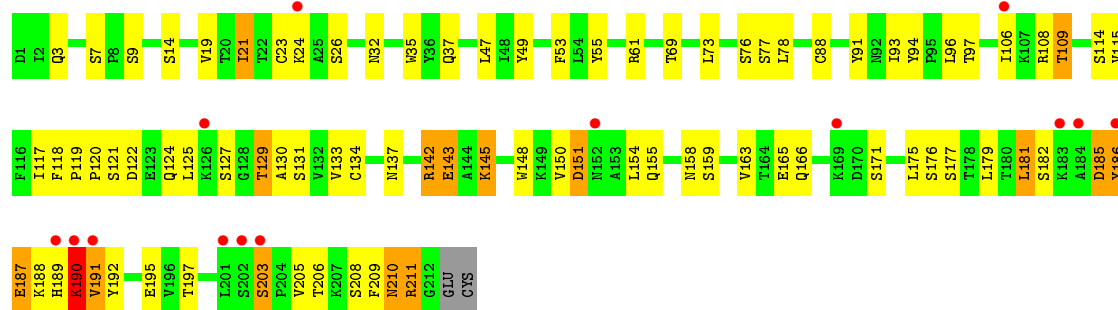


#### • Molecule 1: heavy

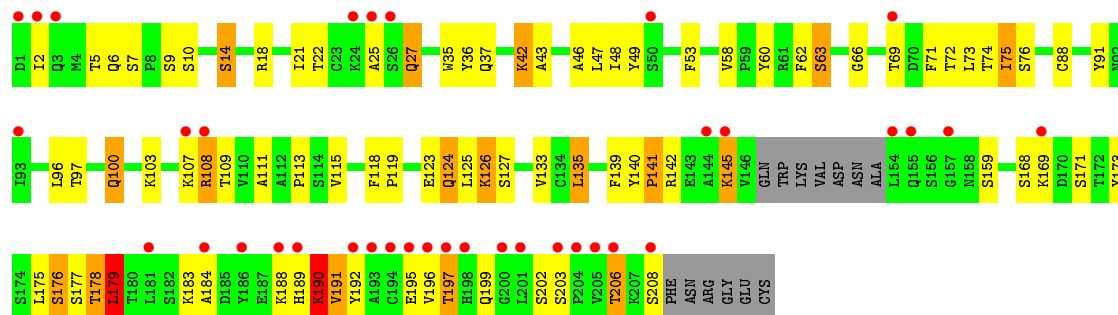




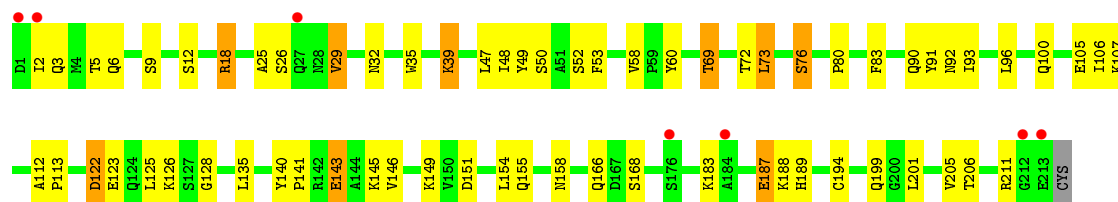
• Molecule 2: light



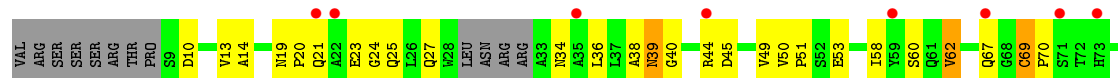
• Molecule 2: light

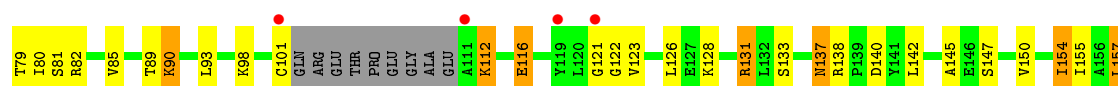


• Molecule 2: light

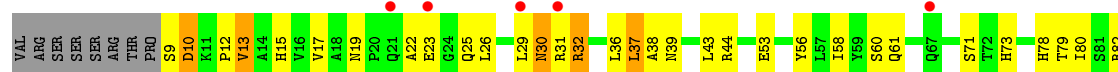


• Molecule 3: Tumor necrosis factor alpha

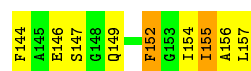
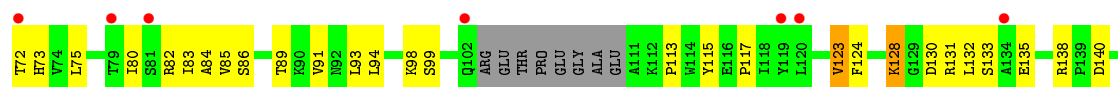
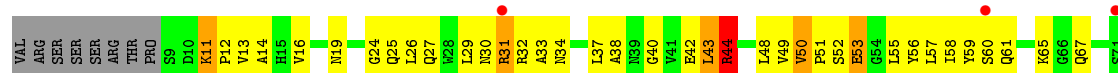




• Molecule 3: Tumor necrosis factor alpha



• Molecule 3: Tumor necrosis factor alpha



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	148.59Å 207.22Å 112.63Å 90.00° 118.81° 90.00°	Depositor
Resolution (Å)	30.00 – 2.90 31.05 – 2.89	Depositor EDS
% Data completeness (in resolution range)	95.1 (30.00-2.90) 95.2 (31.05-2.89)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.93 (at 2.90Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.226 , 0.266 0.224 , 0.264	Depositor DCC
$R_{free}$ test set	3188 reflections (5.08%)	DCC
Wilson B-factor (Å <sup>2</sup> )	79.8	Xtrriage
Anisotropy	0.561	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 62.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	12793	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	87.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.49% 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	0.37	0/1627	0.64	1/2219 (0.0%)
1	C	0.33	0/1569	0.63	0/2137
1	H	0.34	0/1642	0.57	0/2238
2	B	0.35	0/1669	0.60	0/2268
2	D	0.34	0/1571	0.66	1/2134 (0.0%)
2	L	0.34	0/1678	0.57	0/2280
3	E	0.38	0/1080	0.67	1/1468 (0.1%)
3	F	0.33	0/1128	0.61	0/1533
3	G	0.36	0/1128	0.63	0/1533
All	All	0.35	0/13092	0.62	3/17810 (0.0%)

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
1	A	0	3
1	C	0	2
2	D	0	1
3	E	0	1
3	G	0	1
All	All	0	8

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	179	LEU	CA-CB-CG	8.29	134.36	115.30
1	A	146	LEU	CA-CB-CG	8.04	133.78	115.30
3	E	157	LEU	CA-CB-CG	5.76	128.55	115.30

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	209	ASN	Peptide
1	A	210	THR	Peptide
1	A	72	LEU	Peptide
1	C	209	ASN	Peptide
1	C	64	VAL	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	1586	0	1542	37	0
1	C	1530	0	1487	52	0
1	H	1601	0	1560	49	0
2	B	1632	0	1586	59	0
2	D	1538	0	1499	51	0
2	L	1641	0	1592	42	0
3	E	1057	0	1052	38	0
3	F	1104	0	1104	33	0
3	G	1104	0	1104	55	0
All	All	12793	0	12526	383	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 15.

The worst 5 of 383 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)
2:B:151:ASP:H	2:B:191:VAL:HG12	1.25	1.01
3:E:131:ARG:HH11	3:E:131:ARG:HB2	1.25	1.01
1:C:128:PRO:HD3	1:C:214:LYS:HZ3	1.24	0.99
1:C:128:PRO:HD3	1:C:214:LYS:NZ	1.86	0.89
2:D:125:LEU:O	2:D:183:LYS:NZ	2.07	0.87

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	206/224 (92%)	198 (96%)	7 (3%)	1 (0%)	32	68
1	C	196/224 (88%)	189 (96%)	6 (3%)	1 (0%)	32	68
1	H	208/224 (93%)	198 (95%)	10 (5%)	0	100	100
2	B	210/214 (98%)	197 (94%)	12 (6%)	1 (0%)	32	68
2	D	197/214 (92%)	184 (93%)	11 (6%)	2 (1%)	18	51
2	L	211/214 (99%)	200 (95%)	11 (5%)	0	100	100
3	E	130/157 (83%)	123 (95%)	6 (5%)	1 (1%)	22	57
3	F	137/157 (87%)	134 (98%)	3 (2%)	0	100	100
3	G	137/157 (87%)	128 (93%)	7 (5%)	2 (2%)	12	39
All	All	1632/1785 (91%)	1551 (95%)	73 (4%)	8 (0%)	32	68

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	209	ASN
2	D	191	VAL
3	G	31	ARG
1	A	210	THR
2	B	190	LYS

### 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	174/186 (94%)	159 (91%)	15 (9%)	12	35
1	C	168/186 (90%)	146 (87%)	22 (13%)	5	14
1	H	176/186 (95%)	152 (86%)	24 (14%)	4	12
2	B	185/187 (99%)	153 (83%)	32 (17%)	2	7
2	D	176/187 (94%)	142 (81%)	34 (19%)	1	5
2	L	186/187 (100%)	171 (92%)	15 (8%)	14	38
3	E	114/133 (86%)	94 (82%)	20 (18%)	2	6
3	F	119/133 (90%)	103 (87%)	16 (13%)	4	13
3	G	119/133 (90%)	100 (84%)	19 (16%)	3	8
All	All	1417/1518 (93%)	1220 (86%)	197 (14%)	4	12

5 of 197 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	142	ARG
1	H	51	ILE
3	G	42	GLU
2	D	168	SER
2	D	202	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
3	E	39	ASN
3	E	67	GLN
3	F	39	ASN
1	C	13	GLN
3	E	137	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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 [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	210/224 (93%)	0.33	5 (2%) 59 55	54, 78, 128, 140	0
1	C	202/224 (90%)	0.62	22 (10%) 6 4	60, 92, 131, 149	0
1	H	212/224 (94%)	0.50	15 (7%) 17 12	54, 92, 123, 137	0
2	B	212/214 (99%)	0.52	14 (6%) 19 14	47, 80, 133, 166	0
2	D	201/214 (93%)	0.85	36 (17%) 2 1	57, 96, 157, 174	0
2	L	213/214 (99%)	0.21	7 (3%) 47 40	48, 72, 99, 122	0
3	E	136/157 (86%)	0.51	12 (8%) 11 7	54, 82, 128, 159	0
3	F	141/157 (89%)	0.51	9 (6%) 20 15	48, 72, 116, 146	0
3	G	141/157 (89%)	0.62	10 (7%) 17 12	54, 81, 119, 166	0
All	All	1668/1785 (93%)	0.51	130 (7%) 14 10	47, 83, 133, 174	0

The worst 5 of 130 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	G	71	SER	6.9
2	D	154	LEU	5.9
3	E	71	SER	5.6
1	C	199	TYR	5.5
2	B	184	ALA	5.1

### 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.