



## wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 01:53 PM BST

PDB ID : 3J8Z  
EMDB ID: : EMD-5990  
Title : Cryo-EM reconstruction of quasi-HPV16 complex with H16.1A Fab  
Authors : Guan, J.; Hafenstein, S.  
Deposited on : 2014-11-20  
Resolution : 14.00 Å(reported)  
Based on PDB ID : ?

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.  
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.  
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/EMValidationReportHelp>

---

MolProbity : 4.02b-467  
Mogul : unknown  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk27241

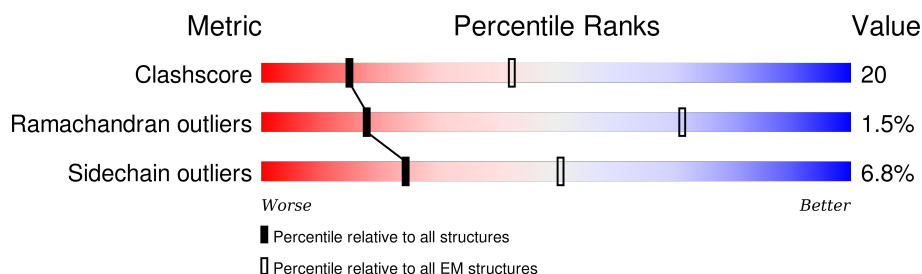
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 14.00 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)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	455	55% 31% 6% 7%
1	B	455	56% 30% 6% 7%
1	C	455	56% 30% 6% 7%
1	D	455	56% 31% 6% 7%
1	E	455	56% 30% 6% 7%
2	J	115	95% 5%
2	K	115	94% 5%
2	L	115	95% 5%
2	M	115	95% 5%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	114	 93%7%
3	G	114	 93%7%
3	H	114	 90%10%
3	I	114	 92%8%

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 30498 atoms, of which 6784 are hydrogens and 0 are deuteriums.

In the tables below, 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 L1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	421	Total	C	N	O	S	0	0
			3322	2116	556	630	20		
1	B	421	Total	C	N	O	S	0	0
			3322	2116	556	630	20		
1	C	421	Total	C	N	O	S	0	0
			3322	2116	556	630	20		
1	D	421	Total	C	N	O	S	0	0
			3322	2116	556	630	20		
1	E	421	Total	C	N	O	S	0	0
			3322	2116	556	630	20		

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	ALA	-	EXPRESSION TAG	UNP Q4VRM0
A	177	GLN	ASN	CONFLICT	UNP Q4VRM0
A	181	GLN	ASN	CONFLICT	UNP Q4VRM0
A	472	LEU	ALA	CONFLICT	UNP Q4VRM0
B	20	ALA	-	EXPRESSION TAG	UNP Q4VRM0
B	177	GLN	ASN	CONFLICT	UNP Q4VRM0
B	181	GLN	ASN	CONFLICT	UNP Q4VRM0
B	472	LEU	ALA	CONFLICT	UNP Q4VRM0
C	20	ALA	-	EXPRESSION TAG	UNP Q4VRM0
C	177	GLN	ASN	CONFLICT	UNP Q4VRM0
C	181	GLN	ASN	CONFLICT	UNP Q4VRM0
C	472	LEU	ALA	CONFLICT	UNP Q4VRM0
D	20	ALA	-	EXPRESSION TAG	UNP Q4VRM0
D	177	GLN	ASN	CONFLICT	UNP Q4VRM0
D	181	GLN	ASN	CONFLICT	UNP Q4VRM0
D	472	LEU	ALA	CONFLICT	UNP Q4VRM0
E	20	ALA	-	EXPRESSION TAG	UNP Q4VRM0
E	177	GLN	ASN	CONFLICT	UNP Q4VRM0
E	181	GLN	ASN	CONFLICT	UNP Q4VRM0
E	472	LEU	ALA	CONFLICT	UNP Q4VRM0

- Molecule 2 is a protein called H16.1A light chain.

Mol	Chain	Residues	Atoms						AltConf	Trace
2	L	115	Total	C	H	N	O	S	0	0
			1742	555	855	148	179	5		
2	J	115	Total	C	H	N	O	S	0	0
			1742	555	855	148	179	5		
2	K	115	Total	C	H	N	O	S	0	0
			1742	555	855	148	179	5		
2	M	115	Total	C	H	N	O	S	0	0
			1742	555	855	148	179	5		

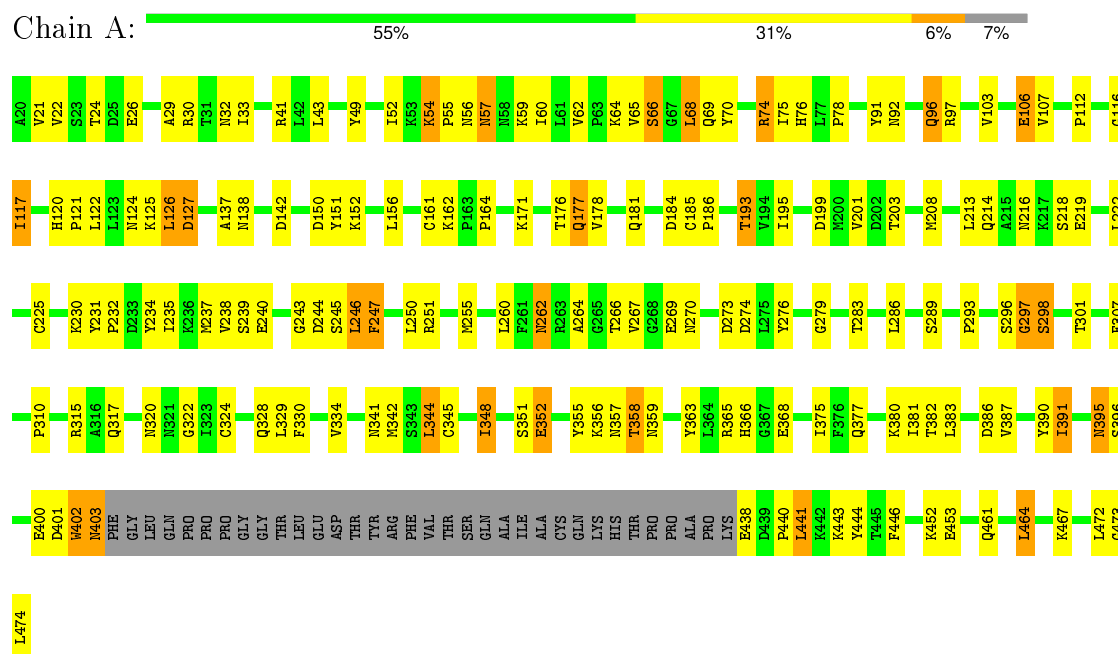
- Molecule 3 is a protein called H16.1A heavy chain.

Mol	Chain	Residues	Atoms						AltConf	Trace
3	H	114	Total	C	H	N	O	S	0	0
			1730	570	841	142	172	5		
3	F	114	Total	C	H	N	O	S	0	0
			1730	570	841	142	172	5		
3	G	114	Total	C	H	N	O	S	0	0
			1730	570	841	142	172	5		
3	I	114	Total	C	H	N	O	S	0	0
			1730	570	841	142	172	5		

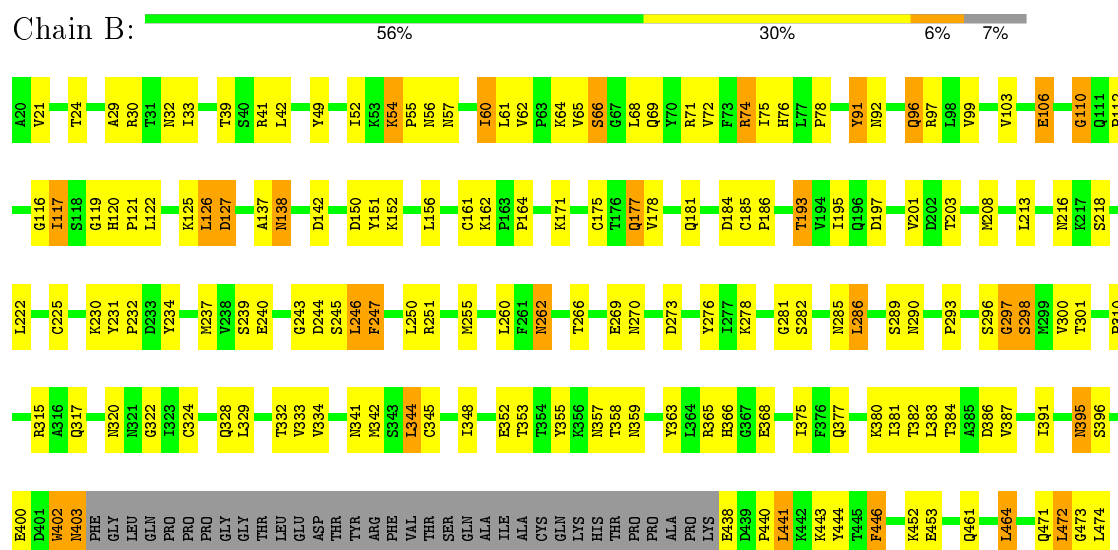
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.

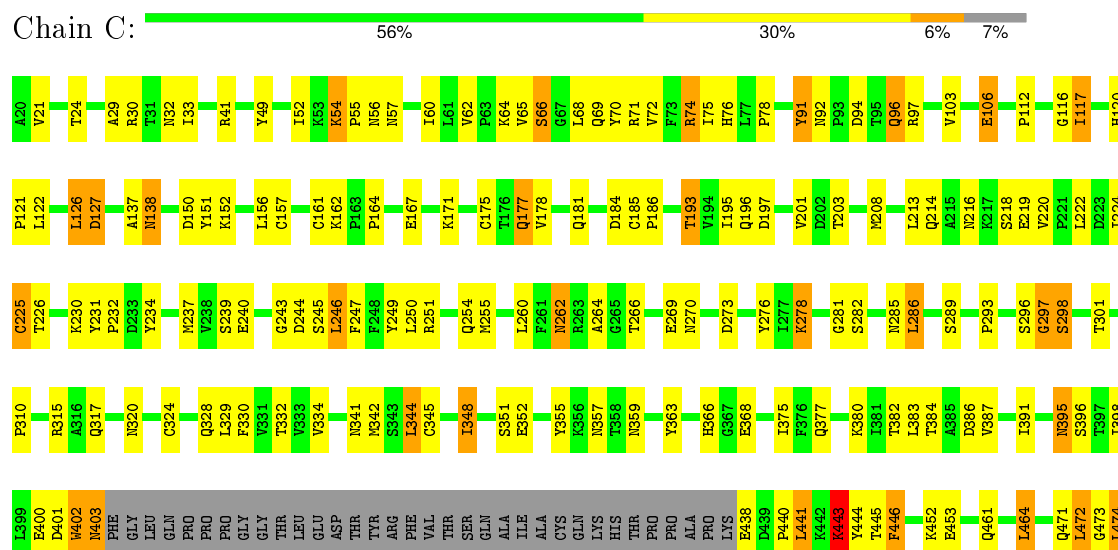
#### • Molecule 1: L1



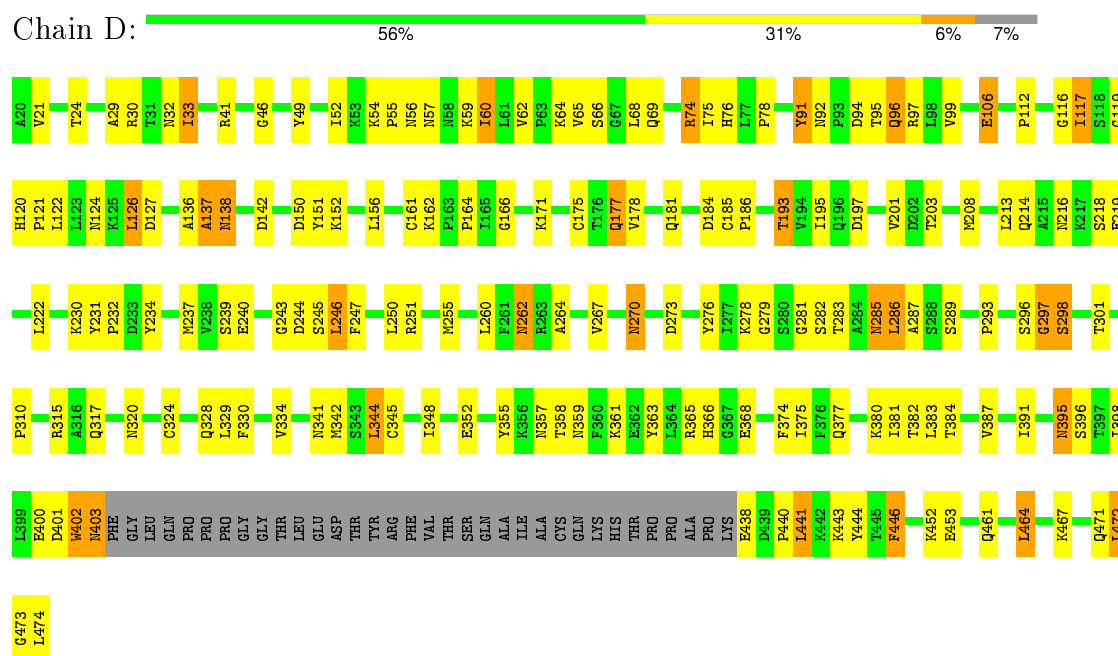
#### • Molecule 1: L1



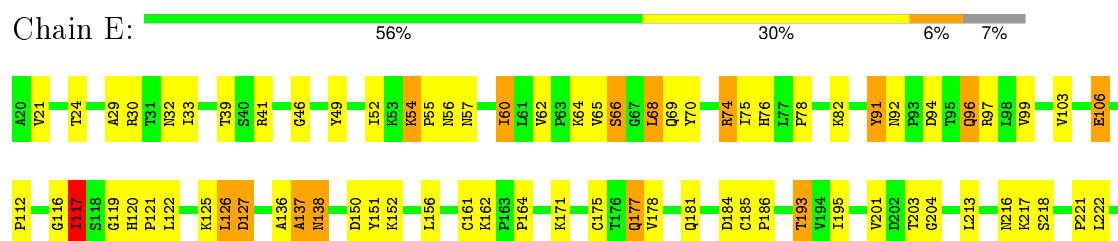
- Molecule 1: L1

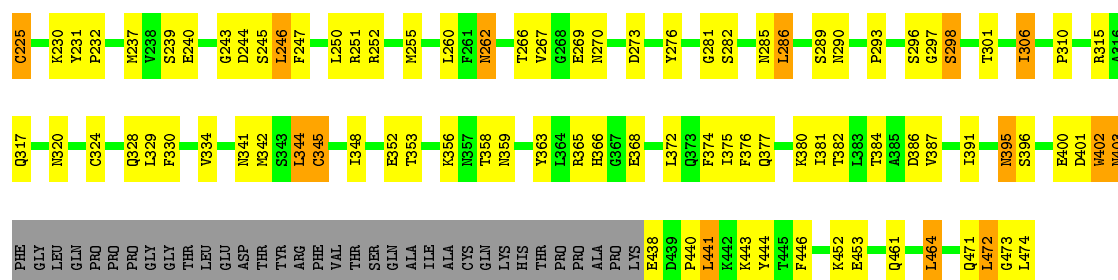


- Molecule 1: L1



- Molecule 1: L1





- Molecule 2: H16.1A light chain

Chain L: 95% 5%



- Molecule 2: H16.1A light chain

Chain J: 95% 5%



- Molecule 2: H16.1A light chain

Chain K: 94% 5%



- Molecule 2: H16.1A light chain

Chain M: 95% 5%



- Molecule 3: H16.1A heavy chain

Chain H: 90% 10%



- Molecule 3: H16.1A heavy chain

Chain F: 93% 7%






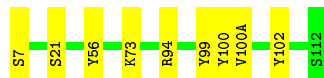
- Molecule 3: H16.1A heavy chain

Chain G:  93% 7%



- Molecule 3: H16.1A heavy chain

Chain I:  92% 8%



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of images	2300	Depositor
Resolution determination method	FSC 0.5	Depositor
CTF correction method	auto3dem	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	200	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	50000	Depositor
Image detector	KODAK SO-163 FILM	Depositor

## 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  > 2$	RMSZ	# $ Z  > 2$
1	A	0.58	6/3407 (0.2%)	0.72	3/4632 (0.1%)
1	B	0.49	0/3407	0.70	2/4632 (0.0%)
1	C	0.53	2/3407 (0.1%)	0.71	1/4632 (0.0%)
1	D	0.55	3/3407 (0.1%)	0.71	3/4632 (0.1%)
1	E	0.53	1/3407 (0.0%)	0.72	0/4632
2	J	0.64	0/906	0.75	1/1223 (0.1%)
2	K	0.65	0/906	0.75	2/1223 (0.2%)
2	L	0.64	0/906	0.75	1/1223 (0.1%)
2	M	0.64	0/906	0.75	1/1223 (0.1%)
3	F	0.59	1/915 (0.1%)	0.73	2/1241 (0.2%)
3	G	0.59	1/915 (0.1%)	0.73	2/1241 (0.2%)
3	H	0.59	1/915 (0.1%)	0.73	2/1241 (0.2%)
3	I	0.59	1/915 (0.1%)	0.73	2/1241 (0.2%)
All	All	0.56	16/24319 (0.1%)	0.72	22/33016 (0.1%)

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	59	LYS	CB-CG	-8.85	1.28	1.52
1	A	443	LYS	CD-CE	-6.54	1.34	1.51
1	A	57	ASN	CG-OD1	-6.32	1.10	1.24
1	D	285	ASN	CG-OD1	-6.04	1.10	1.24
3	I	102	TYR	C-N	6.04	1.48	1.34

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	94	ARG	O-C-N	-7.42	110.82	122.70
3	G	94	ARG	O-C-N	-7.39	110.88	122.70
3	H	94	ARG	O-C-N	-7.37	110.91	122.70
3	I	94	ARG	O-C-N	-7.36	110.92	122.70
2	J	91	HIS	CA-CB-CG	7.14	125.75	113.60

There are no chirality outliers.

There are no planarity outliers.

## 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	3322	0	3227	131	0
1	B	3322	0	3225	193	0
1	C	3322	0	3227	230	0
1	D	3322	0	3226	224	0
1	E	3322	0	3225	189	0
2	J	887	855	855	35	0
2	K	887	855	855	32	0
2	L	887	855	855	33	0
2	M	887	855	855	38	0
3	F	889	841	838	57	0
3	G	889	841	838	64	0
3	H	889	841	838	31	0
3	I	889	841	837	22	0
All	All	23714	6784	22901	918	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 20.

The worst 5 of 918 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:C:285:ASN:HB3	3:G:99:TYR:CD1	1.32	1.63
1:B:282:SER:CB	2:M:30(C):ARG:HB2	1.26	1.57
1:C:282:SER:CB	2:K:30(C):ARG:HB2	1.10	1.52
1:D:282:SER:HB3	2:J:30(C):ARG:CB	1.09	1.52
1:C:138:ASN:ND2	3:G:100:TYR:CE1	1.76	1.51

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 PDB entries followed by that with respect to all EM entries.

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	417/455 (92%)	378 (91%)	29 (7%)	10 (2%)	7	47
1	B	417/455 (92%)	378 (91%)	29 (7%)	10 (2%)	7	47
1	C	417/455 (92%)	375 (90%)	34 (8%)	8 (2%)	10	52
1	D	417/455 (92%)	377 (90%)	30 (7%)	10 (2%)	7	47
1	E	417/455 (92%)	379 (91%)	30 (7%)	8 (2%)	10	52
2	J	113/115 (98%)	109 (96%)	4 (4%)	0	100	100
2	K	113/115 (98%)	109 (96%)	4 (4%)	0	100	100
2	L	113/115 (98%)	109 (96%)	4 (4%)	0	100	100
2	M	113/115 (98%)	109 (96%)	4 (4%)	0	100	100
3	F	112/114 (98%)	108 (96%)	4 (4%)	0	100	100
3	G	112/114 (98%)	108 (96%)	4 (4%)	0	100	100
3	H	112/114 (98%)	108 (96%)	4 (4%)	0	100	100
3	I	112/114 (98%)	108 (96%)	4 (4%)	0	100	100
All	All	2985/3191 (94%)	2755 (92%)	184 (6%)	46 (2%)	18	57

5 of 46 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	137	ALA
1	A	298	SER
1	A	402	TRP
1	B	137	ALA
1	B	298	SER

### 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 PDB entries followed by that with respect to all EM

entries.

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	368/396 (93%)	334 (91%)	34 (9%)	11	43
1	B	368/396 (93%)	335 (91%)	33 (9%)	12	44
1	C	368/396 (93%)	329 (89%)	39 (11%)	8	36
1	D	368/396 (93%)	335 (91%)	33 (9%)	12	44
1	E	368/396 (93%)	330 (90%)	38 (10%)	9	37
2	J	99/99 (100%)	99 (100%)	0	100	100
2	K	99/99 (100%)	99 (100%)	0	100	100
2	L	99/99 (100%)	99 (100%)	0	100	100
2	M	99/99 (100%)	99 (100%)	0	100	100
3	F	89/89 (100%)	89 (100%)	0	100	100
3	G	89/89 (100%)	89 (100%)	0	100	100
3	H	89/89 (100%)	89 (100%)	0	100	100
3	I	89/89 (100%)	89 (100%)	0	100	100
All	All	2592/2732 (95%)	2415 (93%)	177 (7%)	24	57

5 of 177 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	213	LEU
1	C	441	LEU
1	E	317	GLN
1	C	246	LEU
1	C	315	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 66 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	259	HIS
1	C	461	GLN
1	E	403	ASN
1	C	262	ASN
1	C	366	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.