



wwPDB EM Map/Model Validation Report ⓘ

Apr 26, 2016 – 06:28 PM BST

PDB ID : 5FUA
EMDB ID: : EMD-3283
Title : Cryo-EM of BK polyomavirus
Authors : Hurdiss, D.L.; Morgan, E.L.; Thompson, R.F.; Prescott, E.L.; Panou, M.M.;
Macdonald, A.; Ranson, N.A.
Deposited on : 2016-01-22
Resolution : 7.60 Å(reported)

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
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) : trunk27461

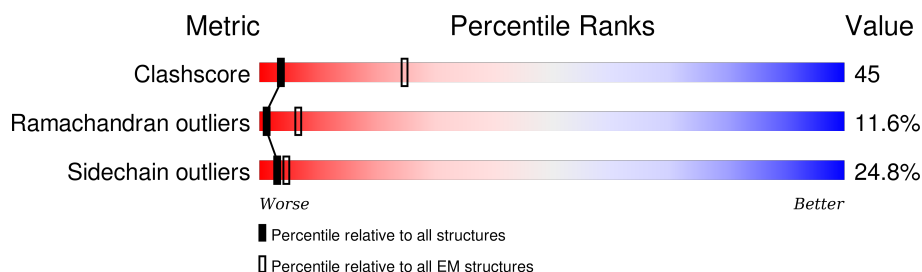
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 7.60 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 ≥ 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	1	362	
1	2	362	
1	3	362	
1	4	362	
1	5	362	
1	6	362	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 31890 atoms, of which 15848 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.

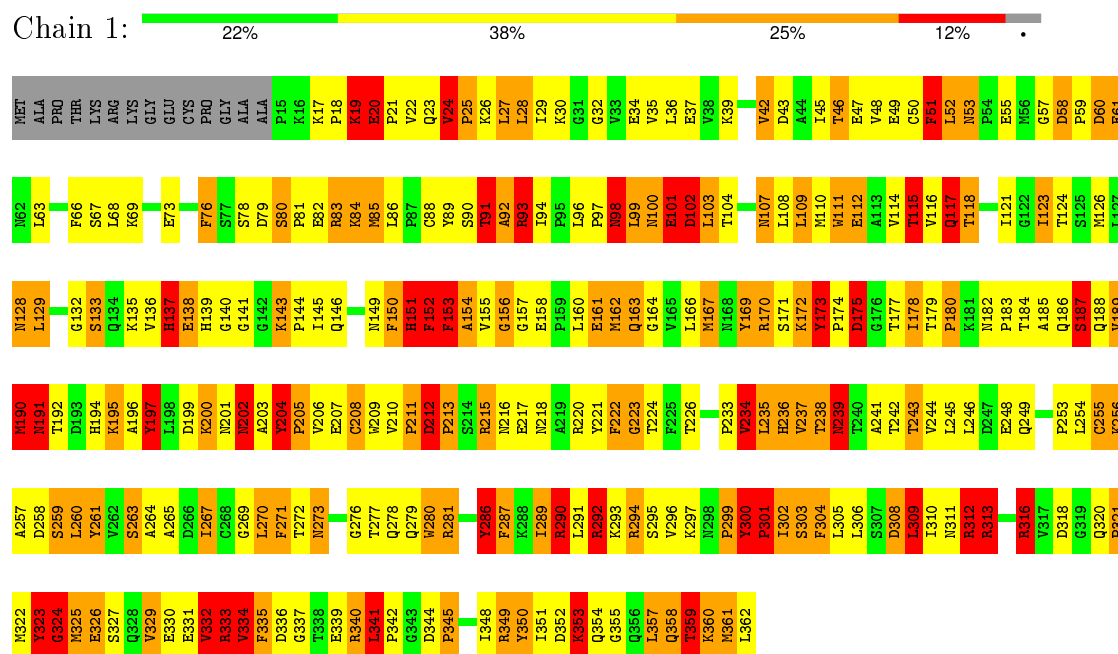
- Molecule 1 is a protein called MAJOR CAPSID PROTEIN VP1.

Mol	Chain	Residues	Atoms						AltConf	Trace
1	1	348	Total	C	H	N	O	S	0	0
			5406	1708	2689	470	522	17		
1	2	348	Total	C	H	N	O	S	0	0
			5406	1708	2689	470	522	17		
1	3	342	Total	C	H	N	O	S	0	0
			5297	1676	2630	462	513	16		
1	4	331	Total	C	H	N	O	S	0	0
			5106	1616	2535	445	495	15		
1	5	347	Total	C	H	N	O	S	0	0
			5392	1703	2682	469	521	17		
1	6	341	Total	C	H	N	O	S	0	0
			5283	1671	2623	461	512	16		

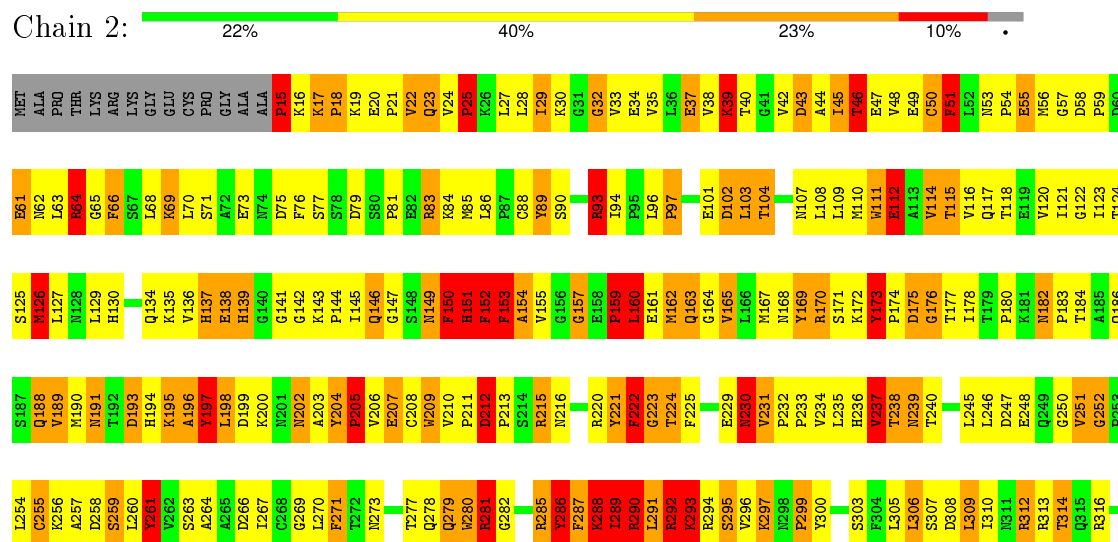
3 Residue-property plots

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: MAJOR CAPSID PROTEIN VP1



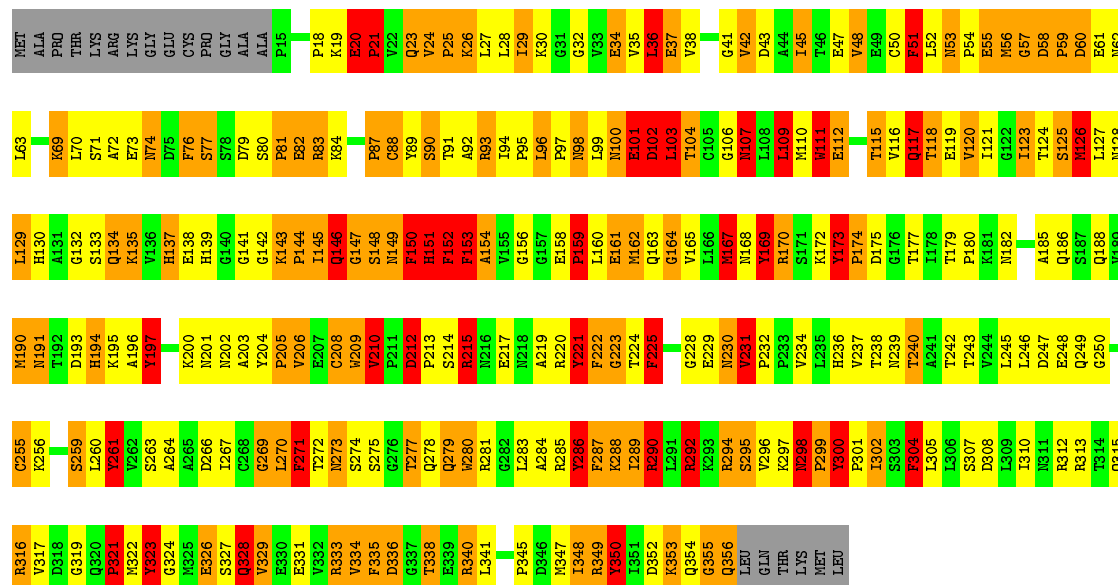
• Molecule 1: MAJOR CAPSID PROTEIN VP1





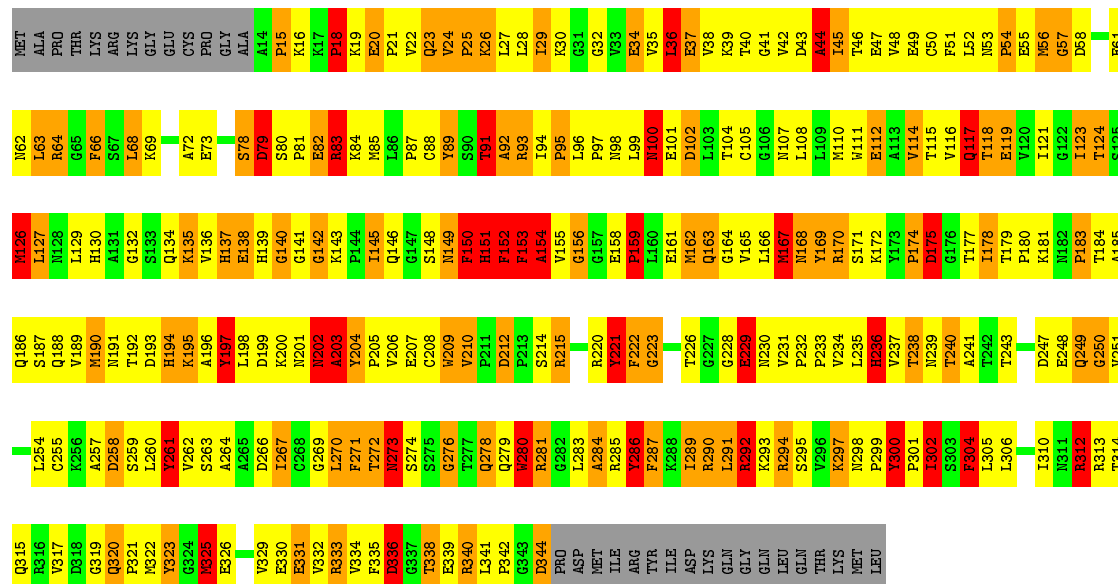
• Molecule 1: MAJOR CAPSID PROTEIN VP1

Chain 3: 23% 35% 26% 11% 6%



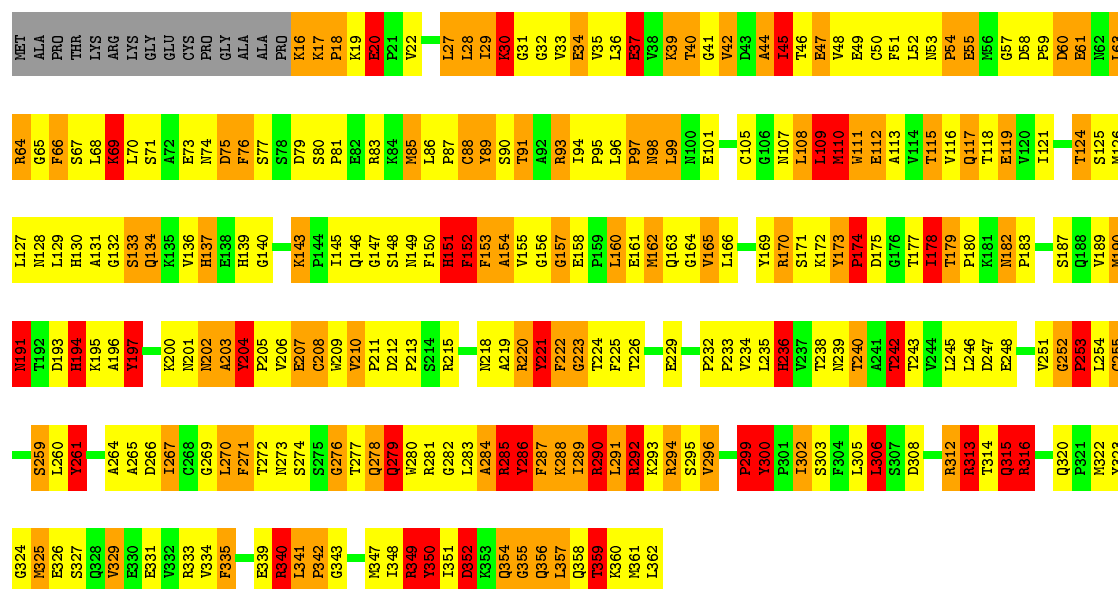
• Molecule 1: MAJOR CAPSID PROTEIN VP1

Chain 4: 19% 40% 23% 9% 9%



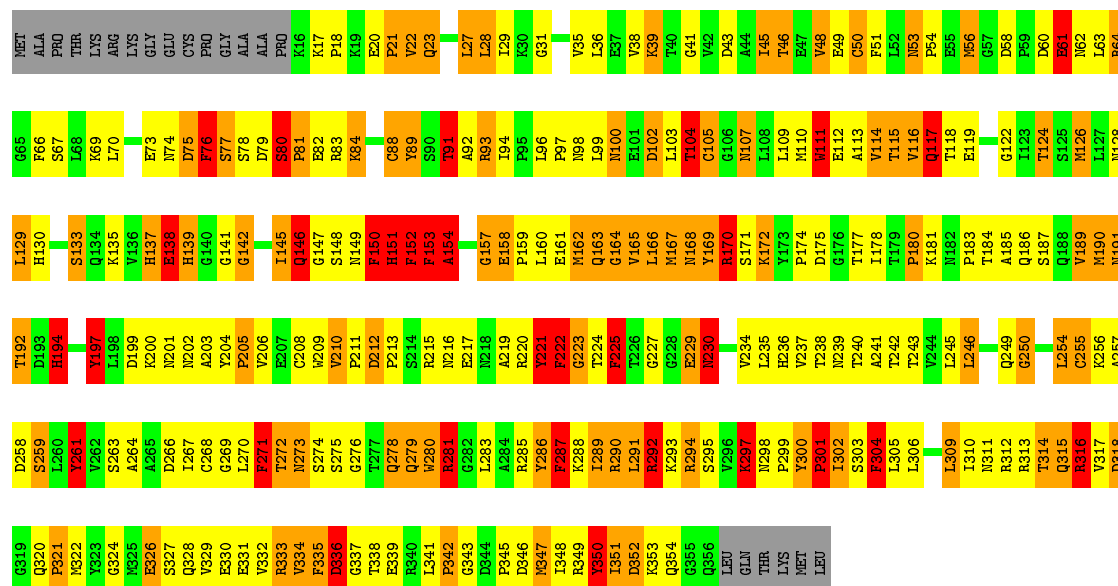
• Molecule 1: MAJOR CAPSID PROTEIN VP1

Chain 5: 22% 40% 24% 10% *



• Molecule 1: MAJOR CAPSID PROTEIN VP1

Chain 6: 23% 39% 24% 9% 6%



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	CTFFIND3, Not provided	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/\text{\AA}^2$)	40	Depositor
Minimum defocus (nm)	600	Depositor
Maximum defocus (nm)	4800	Depositor
Magnification	19000	Depositor
Image detector	GATAN K2 SUMMIT (4K X 4K)	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	1	1.62	18/2777 (0.6%)	2.15	121/3765 (3.2%)
1	2	1.62	11/2777 (0.4%)	2.02	91/3765 (2.4%)
1	3	1.59	18/2727 (0.7%)	2.05	96/3700 (2.6%)
1	4	1.62	10/2629 (0.4%)	2.16	118/3570 (3.3%)
1	5	1.65	22/2769 (0.8%)	2.10	105/3754 (2.8%)
1	6	1.57	16/2719 (0.6%)	2.04	93/3689 (2.5%)
All	All	1.61	95/16398 (0.6%)	2.08	624/22243 (2.8%)

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	1	0	40
1	2	0	40
1	3	0	37
1	4	0	34
1	5	0	36
1	6	0	33
All	All	0	220

The worst 5 of 95 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	5	31	GLY	CA-C	-8.09	1.39	1.51
1	1	350	TYR	CB-CG	-7.78	1.40	1.51
1	3	197	TYR	CB-CG	-7.47	1.40	1.51
1	4	223	GLY	CA-C	-7.38	1.40	1.51
1	2	223	GLY	CA-C	-7.19	1.40	1.51

The worst 5 of 624 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	5	350	TYR	CB-CG-CD1	-19.61	109.24	121.00
1	4	280	TRP	CB-CG-CD2	-15.46	106.50	126.60
1	5	350	TYR	CB-CG-CD2	-14.63	112.22	121.00
1	6	152	PHE	CB-CG-CD1	-13.64	111.25	120.80
1	2	261	TYR	CB-CG-CD1	-13.47	112.92	121.00

There are no chirality outliers.

5 of 220 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	1	20	GLU	Mainchain,Peptide
1	1	24	VAL	Mainchain,Peptide
1	1	25	PRO	Mainchain
1	1	51	PHE	Sidechain
1	1	76	PHE	Sidechain

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	1	2717	2689	2687	263	0
1	2	2717	2689	2687	281	0
1	3	2667	2630	2628	268	0
1	4	2571	2535	2532	253	0
1	5	2710	2682	2679	274	0
1	6	2660	2623	2620	238	0
All	All	16042	15848	15833	1429	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 45.

The worst 5 of 1429 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)
1:4:99:LEU:HD21	1:4:111:TRP:CH2	2.05	0.91
1:3:89:TYR:CD1	1:3:206:VAL:HG12	2.12	0.85

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:2:114:VAL:HG22	1:2:292:ARG:HH11	1.43	0.83
1:3:109:LEU:HD22	1:3:295:SER:HA	1.62	0.81
1:3:310:ILE:HA	1:3:313:ARG:HE	1.46	0.79

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 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	1	346/362 (96%)	248 (72%)	53 (15%)	45 (13%)	0	7
1	2	346/362 (96%)	247 (71%)	60 (17%)	39 (11%)	0	10
1	3	340/362 (94%)	243 (72%)	54 (16%)	43 (13%)	0	8
1	4	329/362 (91%)	233 (71%)	56 (17%)	40 (12%)	0	8
1	5	345/362 (95%)	261 (76%)	54 (16%)	30 (9%)	1	17
1	6	339/362 (94%)	241 (71%)	58 (17%)	40 (12%)	0	9
All	All	2045/2172 (94%)	1473 (72%)	335 (16%)	237 (12%)	1	9

5 of 237 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	1	19	LYS
1	1	78	SER
1	1	81	PRO
1	1	129	LEU
1	1	133	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	1	304/313 (97%)	221 (73%)	83 (27%)	0	4
1	2	304/313 (97%)	235 (77%)	69 (23%)	1	8
1	3	298/313 (95%)	220 (74%)	78 (26%)	0	5
1	4	287/313 (92%)	220 (77%)	67 (23%)	1	7
1	5	303/313 (97%)	234 (77%)	69 (23%)	1	8
1	6	297/313 (95%)	218 (73%)	79 (27%)	0	5
All	All	1793/1878 (96%)	1348 (75%)	445 (25%)	3	6

5 of 445 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	3	217	GLU
1	4	101	GLU
1	6	190	MET
1	3	256	LYS
1	3	348	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 49 such sidechains are listed below:

Mol	Chain	Res	Type
1	3	201	ASN
1	4	53	ASN
1	6	202	ASN
1	3	273	ASN
1	4	74	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 ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.