



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 24, 2022 – 04:14 PM JST

PDB ID : 8GPU
Title : YFV_E_YD6Fab_prefusion
Authors : Li, Y.; Wu, L.; Qi, J.; Yan, J.; Gao, G.F.
Deposited on : 2022-08-27
Resolution : 2.79 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

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

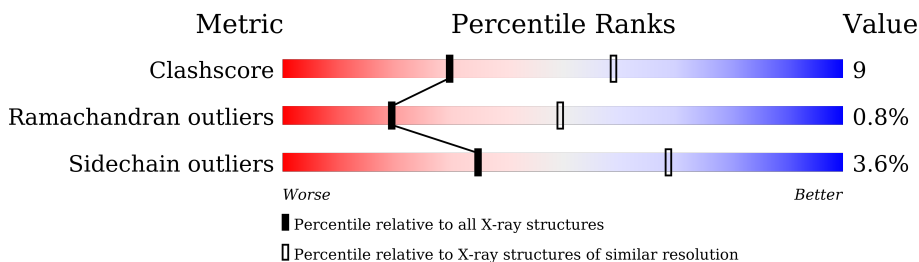
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.79 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))
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)











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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 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 $\leq 5\%$.

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	398	
1	B	398	
1	E	398	
1	I	398	
1	M	398	
1	P	398	
2	C	217	
2	F	217	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	217	 80% 18% .
2	J	217	 72% 23% . .
2	N	217	 75% 24% .
2	Q	217	 81% 18%
3	D	217	 72% 24% . .
3	G	217	 83% 15% .
3	K	217	 76% 20% . .
3	L	217	 71% 26% . .
3	O	217	 80% 17% . .
3	R	217	 76% 19% . .

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 37132 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	392	Total	C	N	O	S	0	1	0
			2997	1877	518	582	20			
1	E	392	Total	C	N	O	S	0	0	0
			2985	1868	515	582	20			
1	B	388	Total	C	N	O	S	0	1	0
			2965	1861	512	572	20			
1	I	388	Total	C	N	O	S	0	0	0
			2952	1853	508	571	20			
1	M	388	Total	C	N	O	S	0	0	0
			2950	1851	508	571	20			
1	P	388	Total	C	N	O	S	0	0	0
			2950	1851	508	571	20			


- Molecule 2 is a protein called YD6Fab_H.


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	216	Total	C	N	O	S	0	0	0
			1623	1028	270	318	7			
2	F	216	Total	C	N	O	S	0	0	0
			1623	1028	270	318	7			
2	C	216	Total	C	N	O	S	0	0	0
			1623	1028	270	318	7			
2	J	211	Total	C	N	O	S	0	0	0
			1591	1010	264	310	7			
2	N	216	Total	C	N	O	S	0	0	0
			1623	1028	270	318	7			
2	Q	216	Total	C	N	O	S	0	0	0
			1623	1028	270	318	7			

- Molecule 3 is a protein called YD6Fab_L.

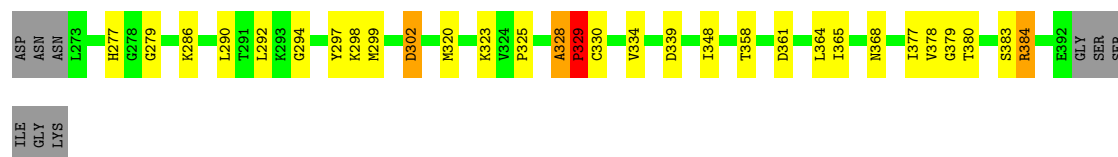
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	214	Total	C	N	O	S	0	0	0
			1601	1004	267	325	5			
3	G	216	Total	C	N	O	S	0	0	0
			1614	1010	268	330	6			
3	D	213	Total	C	N	O	S	0	0	0
			1594	1000	266	323	5			
3	K	214	Total	C	N	O	S	0	0	0
			1601	1004	267	325	5			
3	O	217	Total	C	N	O	S	0	0	0
			1623	1015	270	332	6			
3	R	213	Total	C	N	O	S	0	0	0
			1594	1000	266	323	5			

Note EDS failed to run properly.

- Chain A: 

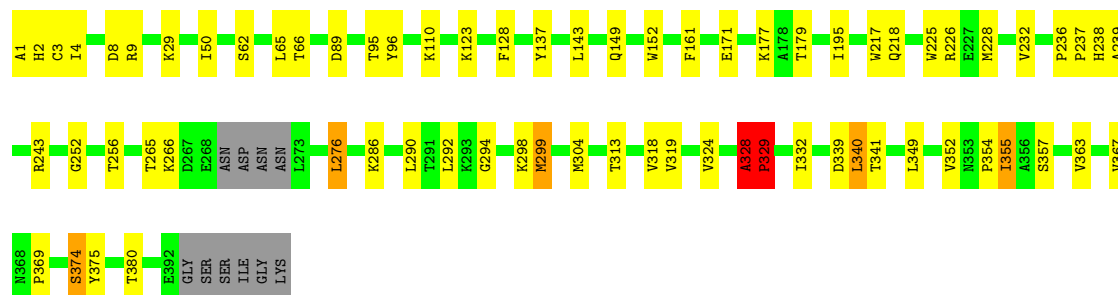
- Chain E:  74% 23% ..
- | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|-----|
| A1 | H2 | I4 | D8 | R9 | V15 | K29 | K34 | K38 | P39 | S40 | Q46 | A56 | B57 | V61 | S62 | A63 | V64 | L65 | T66 | C74 | T77 | A80 | H81 | L82 | A83 | E84 | D89 | R94 | R99 | M103 | K118 | F119 | T120 | G121 | A122 | K123 | S124 | M125 | E129 | G130 | |
| K275 | L276 | K286 | L290 | K293 | K303 | M304 | V307 | K308 | H315 | V319 | K320 | K321 | V322 | K323 | V324 | P325 | A328 | P329 | I335 | D338 | D339 | L349 | V350 | I355 | A356 | S357 | T358 | M359 | V367 | F371 | I377 | S383 | K384 | L385 | E392 | GLY | SER | SER | ILE | GLY | LYS |

- Chain B:
-
- 72% 23%
- A1 H2 C3 D8 R9 T24 L25 E26 K29 K38 P39 L45 V48 C60 V61 S62 L65 T66 H67 V68 K69 L70 S76 E79 E84 R94 T95 Y96 R99 G100 W101 G102 M103 K110 K118 G121 A122 K123 S124 G146 Q149 R150
- W151 W152 K157 F161 E171 F172 K177 A178 T179 Q185 T186 T187 V188 D189 F190 G191 M192 S193 Y194 S202 R207 L215 Q218 S219 W225 V232 E235 P236 P237 H238 A239 T242 R243 Q250 E251 G252 T256 R263 V264 T265 K266 D267 E268 R269



- Molecule 1: Envelope protein

Chain I: 80% 16% ..



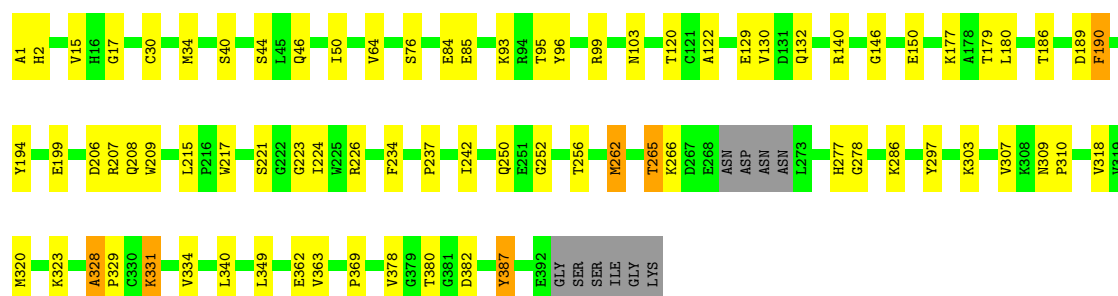
- Molecule 1: Envelope protein

Chain M: 76% 21% ..



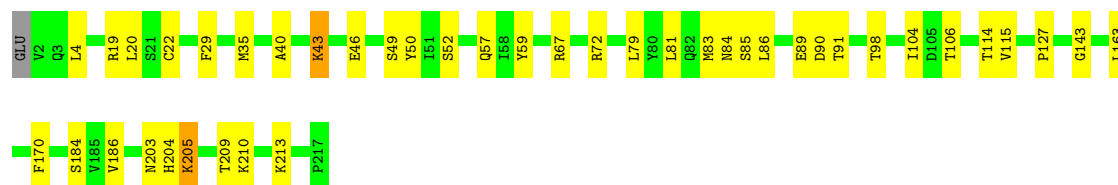
- Molecule 1: Envelope protein

Chain P: 78% 18% ..



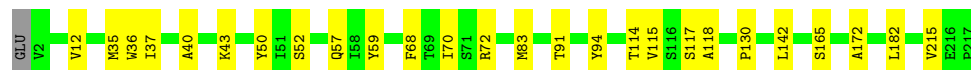
- Molecule 2: YD6Fab_H

Chain H: 80% 18% .



• Molecule 2: YD6Fab_H

Chain F: 88% 12%



• Molecule 2: YD6Fab_H

Chain C: 70% 27%



• Molecule 2: YD6Fab_H

Chain J: 72% 23%



• Molecule 2: YD6Fab_H

Chain N: 75% 24%



• Molecule 2: YD6Fab_H

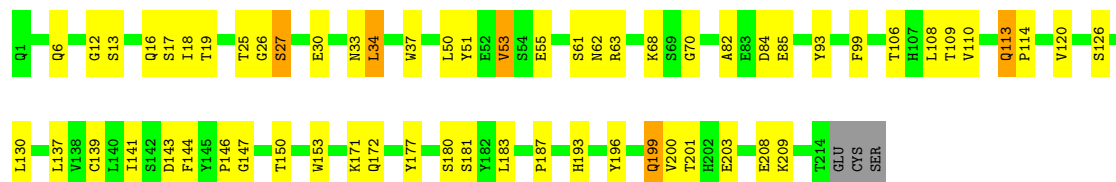
Chain Q: 81% 18%





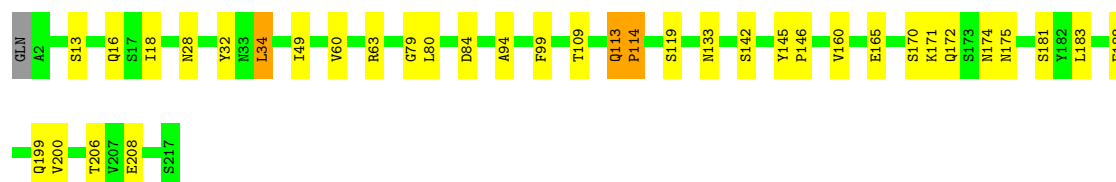
• Molecule 3: YD6Fab_L

Chain L: 71% 26% ..



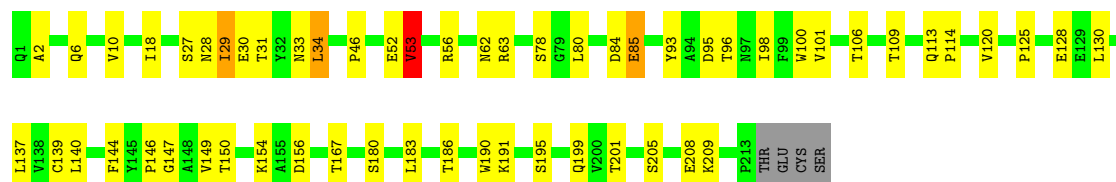
• Molecule 3: YD6Fab_L

Chain G: 83% 15% .



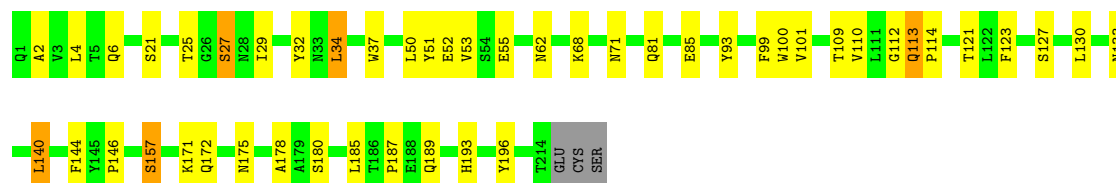
• Molecule 3: YD6Fab_L

Chain D: 72% 24% ..



• Molecule 3: YD6Fab_L

Chain K: 76% 20% ..



• Molecule 3: YD6Fab_L

Chain O: 80% 17% ..





● Molecule 3: YD6Fab_L



4 Data and refinement statistics

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	157.26Å 278.00Å 354.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.57 – 2.79	Depositor
% Data completeness (in resolution range)	85.4 (42.57-2.79)	Depositor
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.03 (at 2.81Å)	Xtriage
Refinement program	PHENIX 1.19_4092	Depositor
R, R_{free}	0.244 , 0.279	Depositor
Wilson B-factor (Å ²)	49.9	Xtriage
Anisotropy	0.138	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.006 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.006 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
Total number of atoms	37132	wwPDB-VP
Average B, all atoms (Å ²)	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.78% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.28	0/3061	0.52	0/4153
1	B	0.27	0/3025	0.54	0/4102
1	E	0.28	0/3045	0.55	2/4132 (0.0%)
1	I	0.26	0/3011	0.52	1/4084 (0.0%)
1	M	0.29	0/3009	0.55	1/4081 (0.0%)
1	P	0.28	0/3009	0.54	0/4081
2	C	0.31	0/1662	0.57	0/2262
2	F	0.29	0/1662	0.55	1/2262 (0.0%)
2	H	0.28	0/1662	0.53	0/2262
2	J	0.29	0/1629	0.56	0/2217
2	N	0.27	0/1662	0.56	0/2262
2	Q	0.28	0/1662	0.52	0/2262
3	D	0.30	0/1634	0.54	0/2232
3	G	0.31	0/1654	0.56	0/2258
3	K	0.33	0/1641	0.58	0/2242
3	L	0.33	0/1641	0.57	0/2242
3	O	0.32	0/1663	0.59	1/2270 (0.0%)
3	R	0.33	0/1634	0.55	0/2232
All	All	0.29	0/37966	0.55	6/51636 (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	B	0	1
1	I	0	1
3	K	0	1
All	All	0	3

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	O	114	PRO	N-CA-CB	-7.17	94.70	103.30
1	E	329	PRO	CA-N-CD	-6.08	102.99	111.50
1	I	329	PRO	CA-N-CD	-5.52	103.77	111.50
1	M	382	ASP	CB-CA-C	5.31	121.02	110.40
1	E	329	PRO	N-CA-CB	-5.21	96.87	102.60

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	328	ALA	Mainchain
1	I	328	ALA	Mainchain
3	K	112	GLY	Mainchain

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	2997	0	2934	54	0
1	B	2965	0	2908	77	0
1	E	2985	0	2911	53	0
1	I	2952	0	2893	39	0
1	M	2950	0	2886	48	0
1	P	2950	0	2886	56	0
2	C	1623	0	1595	46	0
2	F	1623	0	1595	15	0
2	H	1623	0	1595	26	0
2	J	1591	0	1561	38	0
2	N	1623	0	1595	39	0
2	Q	1623	0	1595	22	0
3	D	1594	0	1541	38	0
3	G	1614	0	1553	20	0
3	K	1601	0	1548	41	0
3	L	1601	0	1548	38	0
3	O	1623	0	1564	44	0
3	R	1594	0	1541	32	0
All	All	37132	0	36249	689	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 9.

The worst 5 of 689 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:P:328:ALA:CB	1:P:329:PRO:HD2	1.72	1.19
2:C:76:LYS:O	2:C:78:LEU:HD23	1.43	1.18
3:O:113:GLN:CG	3:O:114:PRO:HD3	1.76	1.14
1:P:328:ALA:HB1	1:P:329:PRO:CD	1.79	1.13
3:O:113:GLN:CB	3:O:114:PRO:HD3	1.78	1.12

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	391/398 (98%)	369 (94%)	19 (5%)	3 (1%)	19	49
1	B	385/398 (97%)	363 (94%)	19 (5%)	3 (1%)	19	49
1	E	390/398 (98%)	367 (94%)	19 (5%)	4 (1%)	15	44
1	I	384/398 (96%)	363 (94%)	19 (5%)	2 (0%)	29	61
1	M	384/398 (96%)	365 (95%)	17 (4%)	2 (0%)	29	61
1	P	384/398 (96%)	364 (95%)	19 (5%)	1 (0%)	41	72
2	C	214/217 (99%)	203 (95%)	10 (5%)	1 (0%)	29	61
2	F	214/217 (99%)	208 (97%)	6 (3%)	0	100	100
2	H	214/217 (99%)	205 (96%)	9 (4%)	0	100	100
2	J	207/217 (95%)	197 (95%)	8 (4%)	2 (1%)	15	44
2	N	214/217 (99%)	202 (94%)	12 (6%)	0	100	100
2	Q	214/217 (99%)	205 (96%)	9 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	D	211/217 (97%)	199 (94%)	9 (4%)	3 (1%)	11	34
3	G	214/217 (99%)	197 (92%)	15 (7%)	2 (1%)	17	46
3	K	212/217 (98%)	194 (92%)	16 (8%)	2 (1%)	17	46
3	L	212/217 (98%)	198 (93%)	10 (5%)	4 (2%)	8	26
3	O	215/217 (99%)	201 (94%)	8 (4%)	6 (3%)	5	17
3	R	211/217 (97%)	196 (93%)	13 (6%)	2 (1%)	17	46
All	All	4870/4992 (98%)	4596 (94%)	237 (5%)	37 (1%)	19	49

5 of 37 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	329	PRO
3	L	26	GLY
1	E	267	ASP
1	E	268	GLU
1	E	329	PRO

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	327/330 (99%)	315 (96%)	12 (4%)	34	68
1	B	322/330 (98%)	306 (95%)	16 (5%)	24	56
1	E	324/330 (98%)	312 (96%)	12 (4%)	34	68
1	I	320/330 (97%)	311 (97%)	9 (3%)	43	77
1	M	319/330 (97%)	309 (97%)	10 (3%)	40	74
1	P	319/330 (97%)	307 (96%)	12 (4%)	33	67
2	C	182/183 (100%)	173 (95%)	9 (5%)	25	57
2	F	182/183 (100%)	180 (99%)	2 (1%)	73	92
2	H	182/183 (100%)	174 (96%)	8 (4%)	28	61
2	J	178/183 (97%)	168 (94%)	10 (6%)	21	51

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	N	182/183 (100%)	176 (97%)	6 (3%)	38	72
2	Q	182/183 (100%)	180 (99%)	2 (1%)	73	92
3	D	179/183 (98%)	172 (96%)	7 (4%)	32	66
3	G	182/183 (100%)	173 (95%)	9 (5%)	25	57
3	K	180/183 (98%)	176 (98%)	4 (2%)	52	83
3	L	180/183 (98%)	174 (97%)	6 (3%)	38	72
3	O	183/183 (100%)	176 (96%)	7 (4%)	33	67
3	R	179/183 (98%)	171 (96%)	8 (4%)	27	60
All	All	4102/4176 (98%)	3953 (96%)	149 (4%)	35	69

5 of 149 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	N	38	ARG
3	R	34	LEU
2	N	183	SER
1	P	199	GLU
3	G	188	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	P	144	HIS
3	R	41	HIS
3	D	28	ASN
2	J	203	ASN
1	M	277	HIS

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 monosaccharides 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](#)

EDS failed to run properly - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates [i](#)

EDS failed to run properly - this section is therefore empty.

6.4 Ligands [i](#)

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers [i](#)

EDS failed to run properly - this section is therefore empty.