



## wwPDB EM Validation Summary Report ⓘ

Feb 24, 2025 – 12:11 pm GMT

PDB ID : 9HVM  
EMDB ID : EMD-52438  
Title : In-cell Structure of Pyrenoid Rubisco  
Authors : Nadav, E.; Zhen, H.; Maud, D.; Alireza, R.; Juan R, P.; Peijun, P.  
Deposited on : 2024-12-30  
Resolution : 8.10 Å(reported)  
Based on initial models : 1GK8, 1EJ7

This is a wwPDB EM 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/EMValidationReportHelp>  
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:

EMDB validation analysis : **FAILED**  
MolProbity : 4.02b-467  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
MapQ : **FAILED**  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.41

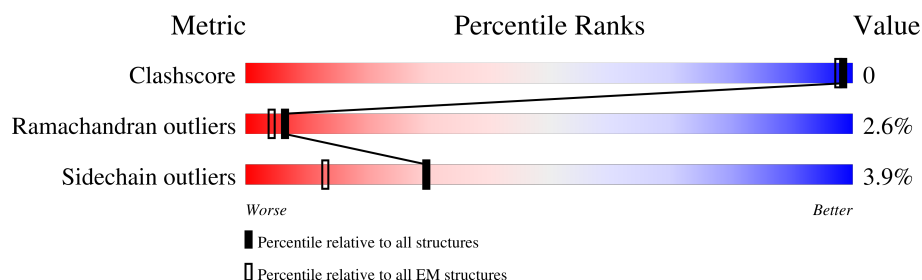
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 8.10 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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

Mol	Chain	Length	Quality of chain
1	A	469	85% 12% .
1	C	469	84% 13% .
1	E	469	85% 13% .
1	G	469	84% 14% .
1	I	469	84% 14% .
1	K	469	85% 13% .
1	M	469	85% 12% .
1	O	469	85% 13% .
2	B	132	92% 6% .

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	D	132	 93%6% •
2	F	132	 93%6% •
2	H	132	 92%7% •
2	J	132	 92%8%
2	L	132	 93%6% •
2	N	132	 93%6% •
2	P	132	 92%8%

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 37816 atoms, of which 0 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 Ribulose biphosphate carboxylase large chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	469	Total	C	N	O	S	0	0
			3647	2308	642	673	24		
1	C	469	Total	C	N	O	S	0	0
			3647	2308	642	673	24		
1	E	469	Total	C	N	O	S	0	0
			3647	2308	642	673	24		
1	G	469	Total	C	N	O	S	0	0
			3647	2308	642	673	24		
1	I	469	Total	C	N	O	S	0	0
			3647	2308	642	673	24		
1	K	469	Total	C	N	O	S	0	0
			3647	2308	642	673	24		
1	M	469	Total	C	N	O	S	0	0
			3647	2308	642	673	24		
1	O	469	Total	C	N	O	S	0	0
			3647	2308	642	673	24		

- Molecule 2 is a protein called Ribulose biphosphate carboxylase small subunit, chloroplastic 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	132	Total	C	N	O	S	0	0
			1080	701	176	192	11		
2	D	132	Total	C	N	O	S	0	0
			1080	701	176	192	11		
2	F	132	Total	C	N	O	S	0	0
			1080	701	176	192	11		
2	H	132	Total	C	N	O	S	0	0
			1080	701	176	192	11		
2	J	132	Total	C	N	O	S	0	0
			1080	701	176	192	11		
2	L	132	Total	C	N	O	S	0	0
			1080	701	176	192	11		

*Continued on next page...*

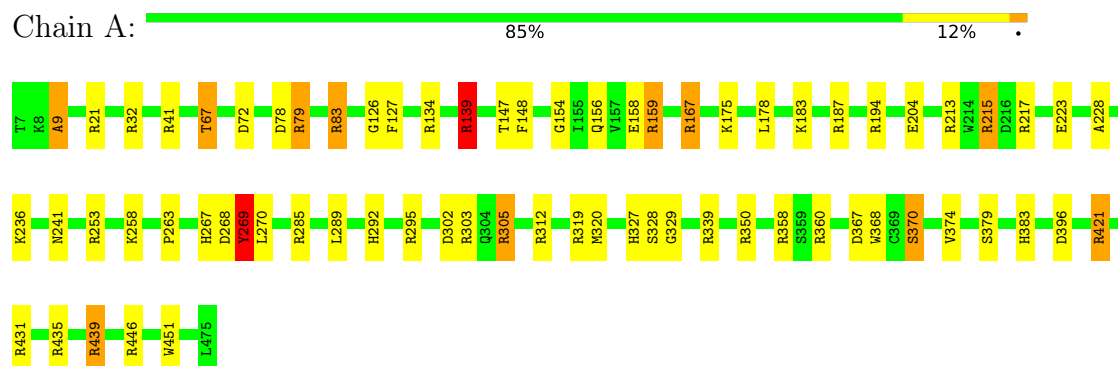
*Continued from previous page...*

Mol	Chain	Residues	Atoms					AltConf	Trace
2	N	132	Total	C	N	O	S	0	0
			1080	701	176	192	11		
2	P	132	Total	C	N	O	S	0	0
			1080	701	176	192	11		

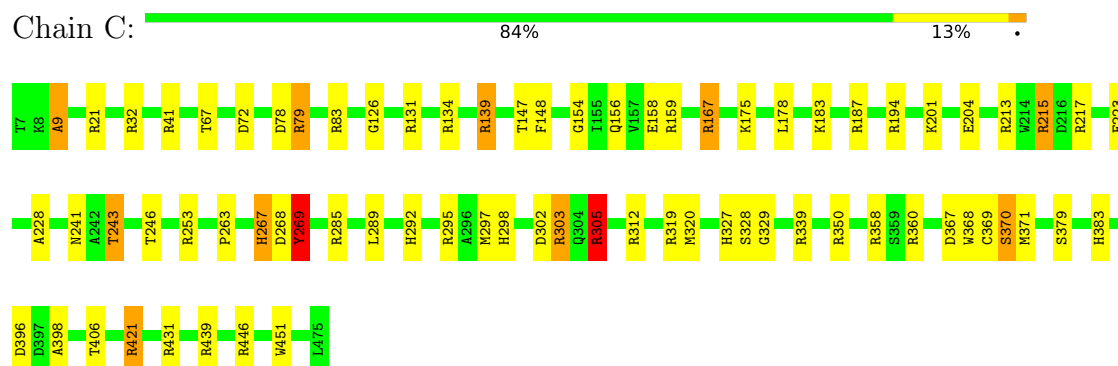
### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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. 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.

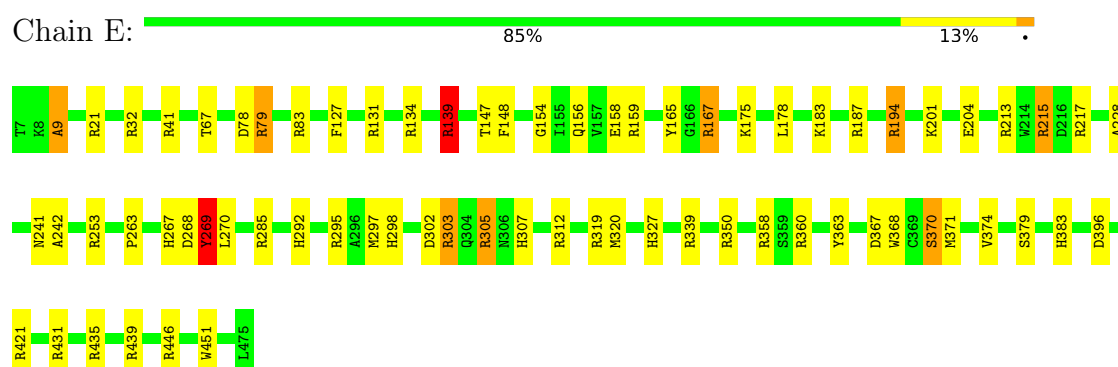
- Molecule 1: Ribulose biphosphate carboxylase large chain



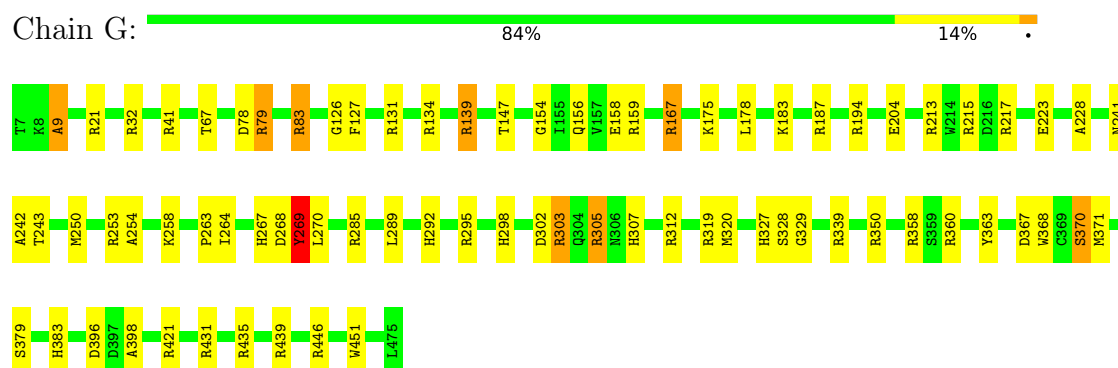
- Molecule 1: Ribulose biphosphate carboxylase large chain



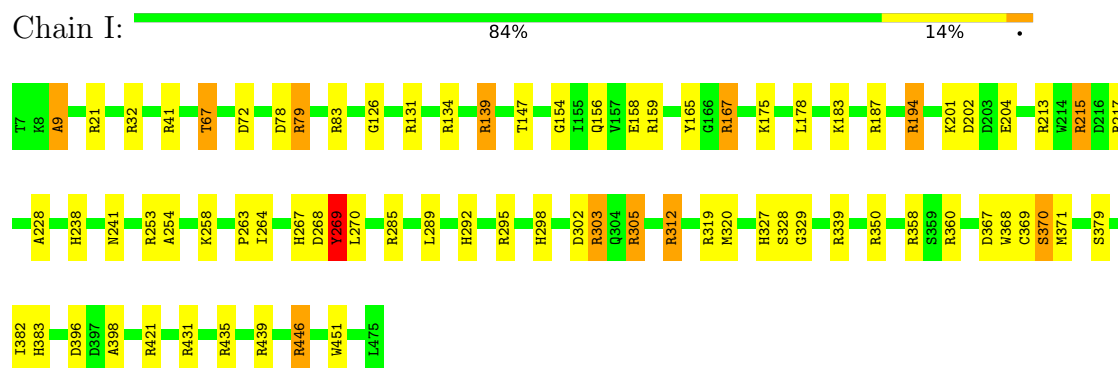
- Molecule 1: Ribulose biphosphate carboxylase large chain



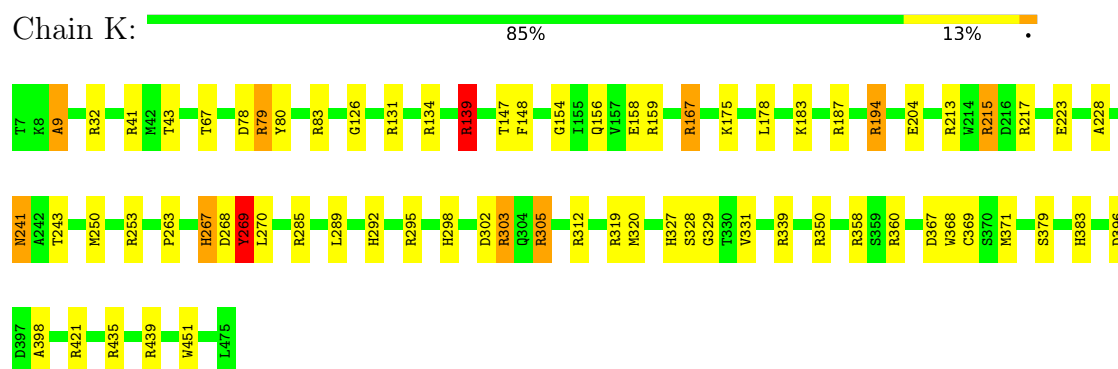
- Molecule 1: Ribulose biphosphate carboxylase large chain



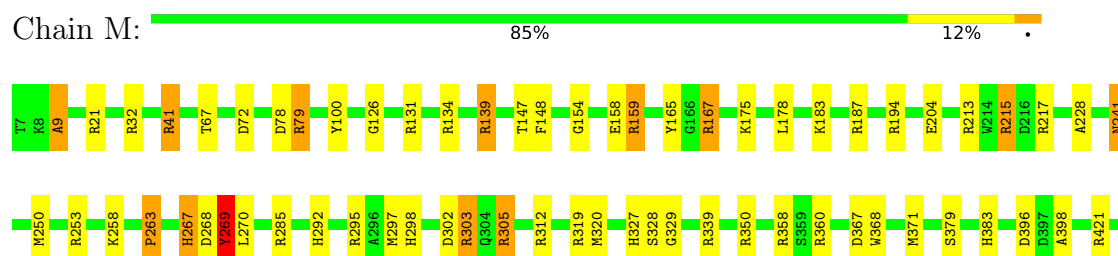
- Molecule 1: Ribulose biphosphate carboxylase large chain



- Molecule 1: Ribulose biphosphate carboxylase large chain



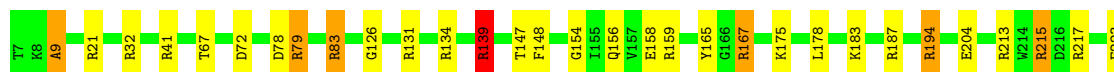
- Molecule 1: Ribulose biphosphate carboxylase large chain





- Molecule 1: Ribulose biphosphate carboxylase large chain

Chain O: 85% 13%



- Molecule 2: Ribulose biphosphate carboxylase small subunit, chloroplastic 1

Chain B: 92% 6%



- Molecule 2: Ribulose biphosphate carboxylase small subunit, chloroplastic 1

Chain D: 93% 6%



- Molecule 2: Ribulose biphosphate carboxylase small subunit, chloroplastic 1

Chain F: 93% 6%



- Molecule 2: Ribulose biphosphate carboxylase small subunit, chloroplastic 1

Chain H: 92% 7%



- Molecule 2: Ribulose biphosphate carboxylase small subunit, chloroplastic 1

Chain J: 92% 8%





- Molecule 2: Ribulose biphosphate carboxylase small subunit, chloroplastic 1



- Molecule 2: Ribulose biphosphate carboxylase small subunit, chloroplastic 1



- Molecule 2: Ribulose biphosphate carboxylase small subunit, chloroplastic 1



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SUBTOMOGRAM AVERAGING	Depositor
Imposed symmetry	POINT, D4	Depositor
Number of subtomograms used	17713	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	120	Depositor
Minimum defocus (nm)	2500	Depositor
Maximum defocus (nm)	4500	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (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  >5	RMSZ	# Z  >5
1	A	0.72	0/3733	1.27	46/5054 (0.9%)
1	C	0.72	0/3733	1.28	46/5054 (0.9%)
1	E	0.72	0/3733	1.29	46/5054 (0.9%)
1	G	0.72	0/3733	1.28	51/5054 (1.0%)
1	I	0.72	0/3733	1.25	43/5054 (0.9%)
1	K	0.72	0/3733	1.22	37/5054 (0.7%)
1	M	0.72	0/3733	1.24	45/5054 (0.9%)
1	O	0.72	0/3733	1.27	44/5054 (0.9%)
2	B	0.73	0/1111	1.14	10/1509 (0.7%)
2	D	0.74	0/1111	1.10	7/1509 (0.5%)
2	F	0.74	0/1111	1.10	7/1509 (0.5%)
2	H	0.74	0/1111	1.09	7/1509 (0.5%)
2	J	0.74	0/1111	1.10	6/1509 (0.4%)
2	L	0.74	0/1111	1.09	7/1509 (0.5%)
2	N	0.74	0/1111	1.11	7/1509 (0.5%)
2	P	0.74	0/1111	1.08	4/1509 (0.3%)
All	All	0.72	0/38752	1.23	413/52504 (0.8%)

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	A	0	19
1	C	0	16
1	E	0	20
1	G	0	18
1	I	0	20
1	K	0	19
1	M	0	18
1	O	0	19
2	B	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
2	D	0	1
2	F	0	1
2	H	0	2
2	J	0	1
2	L	0	2
2	N	0	2
2	P	0	2
All	All	0	161

There are no bond length outliers.

The worst 5 of 413 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	253	ARG	NE-CZ-NH1	13.27	126.93	120.30
1	M	253	ARG	NE-CZ-NH1	12.66	126.63	120.30
1	A	253	ARG	NE-CZ-NH1	12.51	126.56	120.30
1	C	446	ARG	NE-CZ-NH1	12.37	126.48	120.30
1	C	421	ARG	NE-CZ-NH1	11.99	126.30	120.30

There are no chirality outliers.

5 of 161 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	139	ARG	Sidechain
1	A	147	THR	Peptide
1	A	178	LEU	Peptide
1	A	215	ARG	Sidechain
1	A	83	ARG	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	A	3647	0	3570	0	0
1	C	3647	0	3570	2	0
1	E	3647	0	3570	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	3647	0	3570	2	0
1	I	3647	0	3570	3	0
1	K	3647	0	3570	2	0
1	M	3647	0	3570	2	0
1	O	3647	0	3570	1	0
2	B	1080	0	1052	0	0
2	D	1080	0	1052	0	0
2	F	1080	0	1052	0	0
2	H	1080	0	1052	0	0
2	J	1080	0	1052	1	0
2	L	1080	0	1052	0	0
2	N	1080	0	1052	0	0
2	P	1080	0	1052	0	0
All	All	37816	0	36976	14	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 0.

The worst 5 of 14 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:E:307:HIS:CE1	1:G:307:HIS:CE1	2.99	0.51
1:I:202:ASP:OD2	1:I:238:HIS:HE1	1.97	0.48
1:M:298:HIS:CE1	1:M:303:ARG:HD2	2.50	0.47
1:G:298:HIS:CE1	1:G:303:ARG:HD2	2.50	0.47
1:I:298:HIS:CE1	1:I:303:ARG:HD2	2.51	0.46

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	467/469 (100%)	420 (90%)	33 (7%)	14 (3%)	3	23
1	C	467/469 (100%)	420 (90%)	30 (6%)	17 (4%)	3	20
1	E	467/469 (100%)	418 (90%)	36 (8%)	13 (3%)	4	24
1	G	467/469 (100%)	419 (90%)	33 (7%)	15 (3%)	3	21
1	I	467/469 (100%)	416 (89%)	36 (8%)	15 (3%)	3	21
1	K	467/469 (100%)	425 (91%)	28 (6%)	14 (3%)	3	23
1	M	467/469 (100%)	420 (90%)	33 (7%)	14 (3%)	3	23
1	O	467/469 (100%)	418 (90%)	34 (7%)	15 (3%)	3	21
2	B	130/132 (98%)	118 (91%)	10 (8%)	2 (2%)	8	40
2	D	130/132 (98%)	120 (92%)	9 (7%)	1 (1%)	16	55
2	F	130/132 (98%)	120 (92%)	9 (7%)	1 (1%)	16	55
2	H	130/132 (98%)	121 (93%)	9 (7%)	0	100	100
2	J	130/132 (98%)	120 (92%)	9 (7%)	1 (1%)	16	55
2	L	130/132 (98%)	118 (91%)	11 (8%)	1 (1%)	16	55
2	N	130/132 (98%)	121 (93%)	9 (7%)	0	100	100
2	P	130/132 (98%)	120 (92%)	8 (6%)	2 (2%)	8	40
All	All	4776/4808 (99%)	4314 (90%)	337 (7%)	125 (3%)	6	26

5 of 125 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	9	ALA
1	A	67	THR
1	A	78	ASP
1	A	269	TYR
1	A	305	ARG

### 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	375/375 (100%)	357 (95%)	18 (5%)	21	43
1	C	375/375 (100%)	358 (96%)	17 (4%)	23	45
1	E	375/375 (100%)	360 (96%)	15 (4%)	27	47
1	G	375/375 (100%)	359 (96%)	16 (4%)	25	46
1	I	375/375 (100%)	357 (95%)	18 (5%)	21	43
1	K	375/375 (100%)	357 (95%)	18 (5%)	21	43
1	M	375/375 (100%)	359 (96%)	16 (4%)	25	46
1	O	375/375 (100%)	358 (96%)	17 (4%)	23	45
2	B	116/116 (100%)	114 (98%)	2 (2%)	56	72
2	D	116/116 (100%)	114 (98%)	2 (2%)	56	72
2	F	116/116 (100%)	114 (98%)	2 (2%)	56	72
2	H	116/116 (100%)	114 (98%)	2 (2%)	56	72
2	J	116/116 (100%)	114 (98%)	2 (2%)	56	72
2	L	116/116 (100%)	114 (98%)	2 (2%)	56	72
2	N	116/116 (100%)	113 (97%)	3 (3%)	41	59
2	P	116/116 (100%)	113 (97%)	3 (3%)	41	59
All	All	3928/3928 (100%)	3775 (96%)	153 (4%)	30	48

5 of 153 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	79	ARG
1	O	223	GLU
1	M	175	LYS
2	N	58	GLU
2	P	49	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
2	L	99	ASN
2	P	99	ASN
2	H	99	ASN
1	K	241	ASN
1	K	267	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 oligosaccharides 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.