wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 23, 2025 – 02:20 PM JST

PDB ID	:	8YUN / pdb_00008yun
Title	:	The crystal structure of HNBP001-HCP
Authors	:	Chen, C.Z.; Xia, Q.F.
Deposited on	:	2024-03-27
Resolution	:	2.70 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

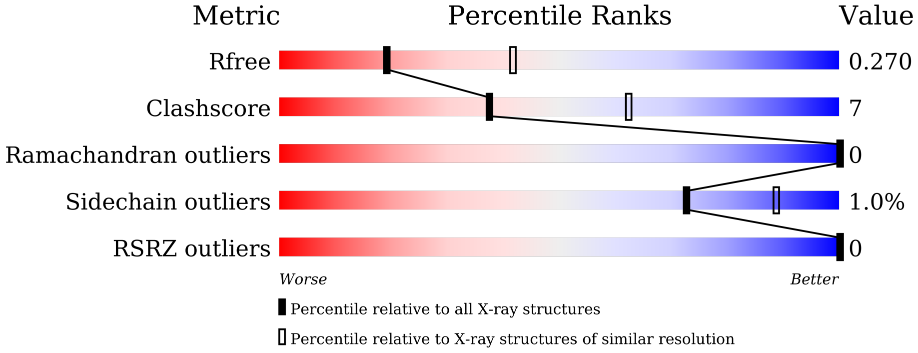
MolProbity	:	4-5-2 with Phenix2.0rc1
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.44

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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}	164625	3333 (2.70-2.70)
Clashscore	180529	3684 (2.70-2.70)
Ramachandran outliers	177936	3633 (2.70-2.70)
Sidechain outliers	177891	3633 (2.70-2.70)
RSRZ outliers	164620	3333 (2.70-2.70)

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 <=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	161	
1	B	161	
1	C	161	
1	D	161	
1	E	161	
1	F	161	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	161	 73% 16% 12%
1	H	161	 67% 23% 10%
1	I	161	 71% 17% 12%
1	J	161	 70% 17% • 13%
1	K	161	 74% 14% 12%
1	L	161	 71% 16% • 12%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 13300 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 Hcp1 family type VI secretion system effector.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
1	A	140	Total C N O S 1089 680 194 212 3	0	0	0
1	B	142	Total C N O S 1103 688 194 218 3	0	0	0
1	C	141	Total C N O S 1094 682 192 217 3	0	0	0
1	D	142	Total C N O S 1105 688 196 218 3	0	0	0
1	E	145	Total C N O S 1130 703 200 224 3	0	0	0
1	G	142	Total C N O S 1106 690 197 216 3	0	0	0
1	K	142	Total C N O S 1103 688 194 218 3	0	0	0
1	H	145	Total C N O S 1130 703 200 224 3	0	0	0
1	I	142	Total C N O S 1103 688 194 218 3	0	0	0
1	F	141	Total C N O S 1095 684 193 215 3	0	0	0
1	L	141	Total C N O S 1095 684 193 215 3	0	0	0
1	J	140	Total C N O S 1086 678 191 214 3	0	0	0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	4	Total O 4 4	0	0
2	B	6	Total O 6 6	0	0

Continued on next page...

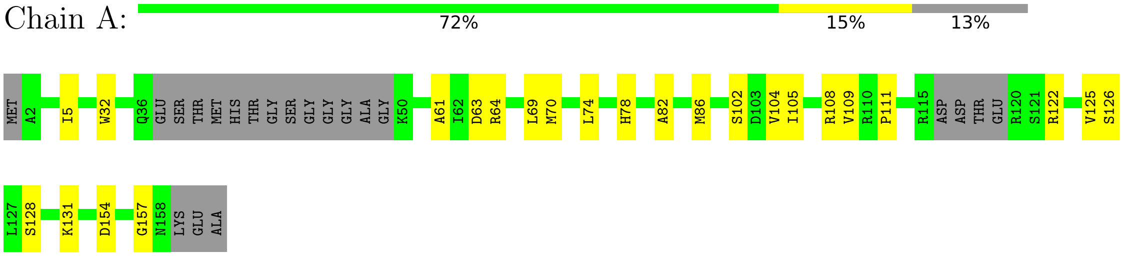
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	4	Total O 4 4	0	0
2	D	2	Total O 2 2	0	0
2	E	8	Total O 8 8	0	0
2	G	3	Total O 3 3	0	0
2	K	4	Total O 4 4	0	0
2	H	2	Total O 2 2	0	0
2	I	3	Total O 3 3	0	0
2	F	7	Total O 7 7	0	0
2	L	11	Total O 11 11	0	0
2	J	7	Total O 7 7	0	0

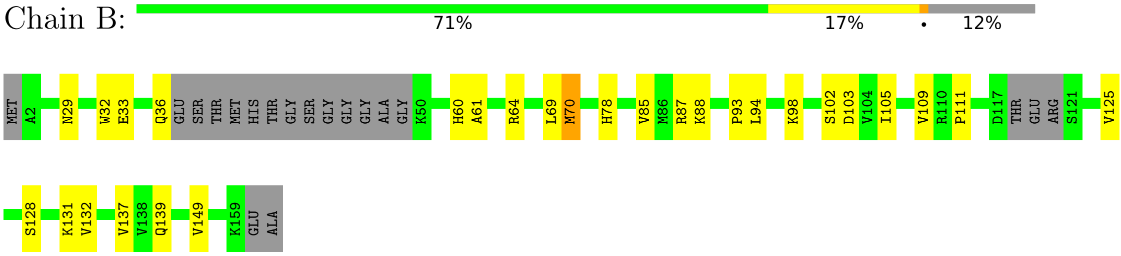
3 Residue-property plots ⓘ

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 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.

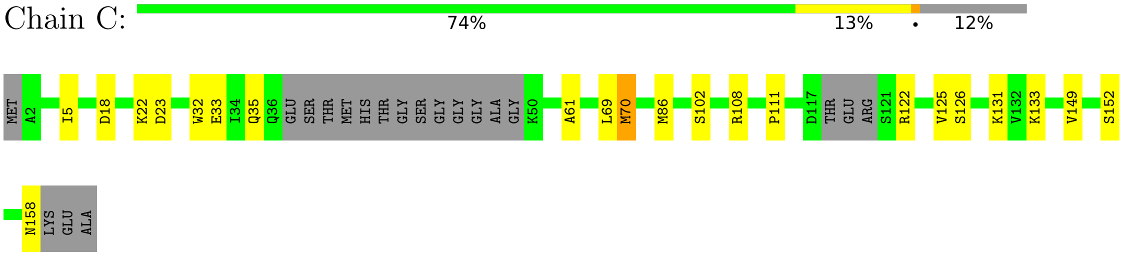
- Molecule 1: Hcp1 family type VI secretion system effector



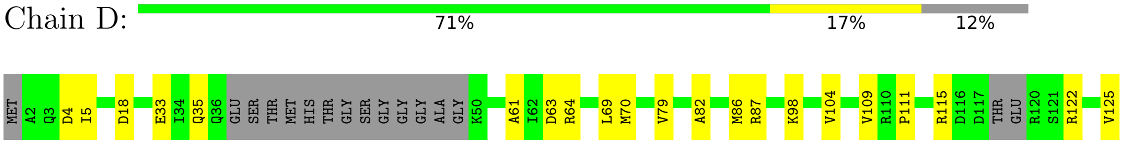
- Molecule 1: Hcp1 family type VI secretion system effector



- Molecule 1: Hcp1 family type VI secretion system effector

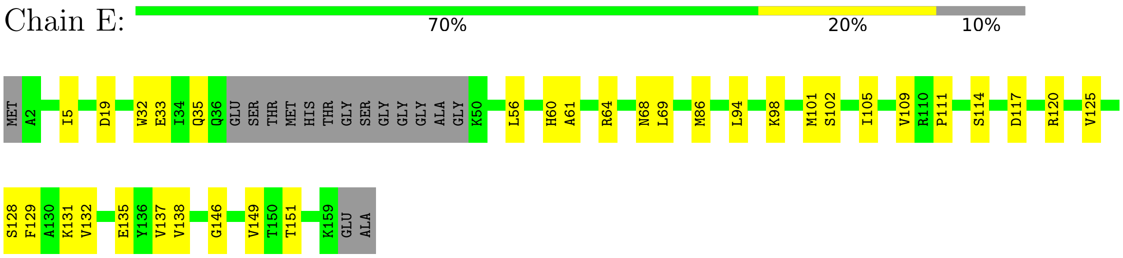


- Molecule 1: Hcp1 family type VI secretion system effector



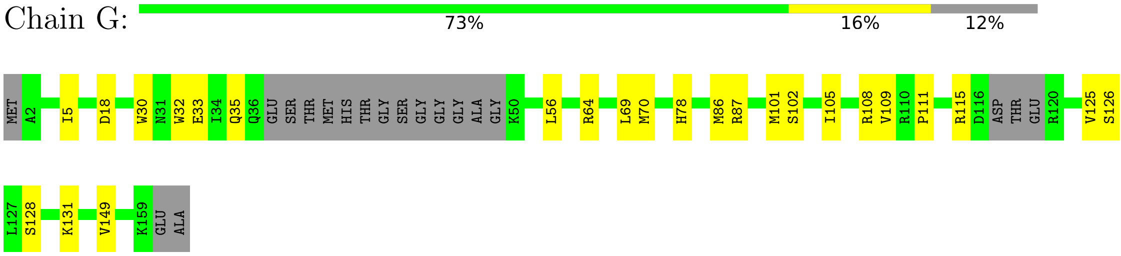
- Molecule 1: Hcp1 family type VI secretion system effector

Chain E:



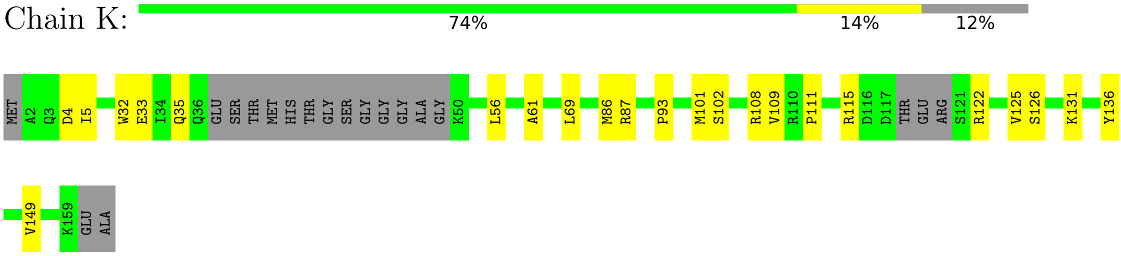
- Molecule 1: Hcp1 family type VI secretion system effector

Chain G:



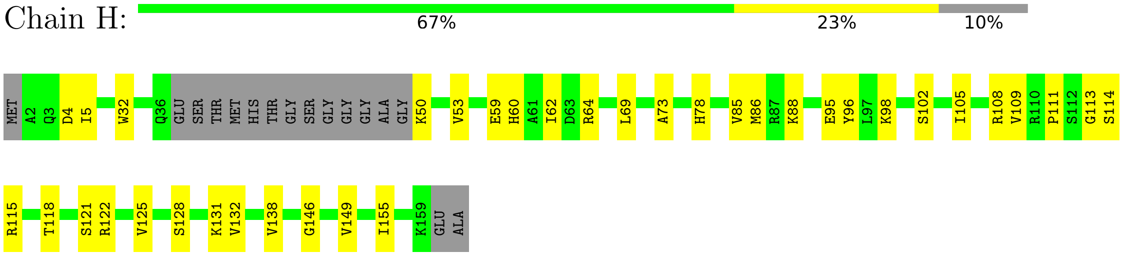
- Molecule 1: Hcp1 family type VI secretion system effector

Chain K:



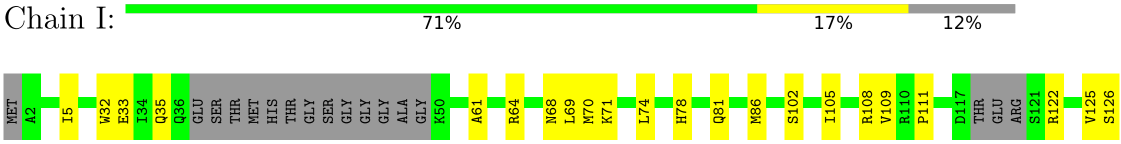
- Molecule 1: Hcp1 family type VI secretion system effector

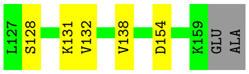
Chain H:



- Molecule 1: Hcp1 family type VI secretion system effector

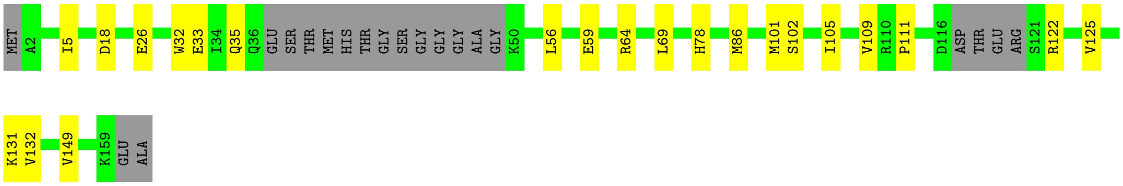
Chain I:





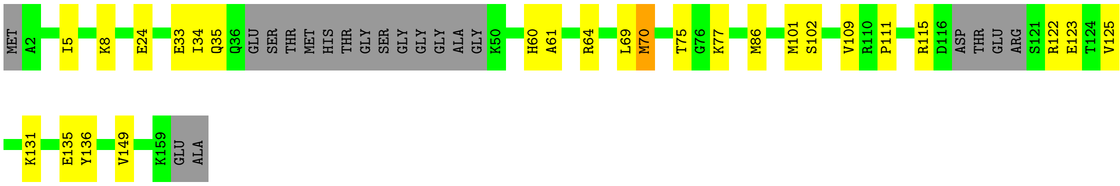
- Molecule 1: Hcp1 family type VI secretion system effector

Chain F:



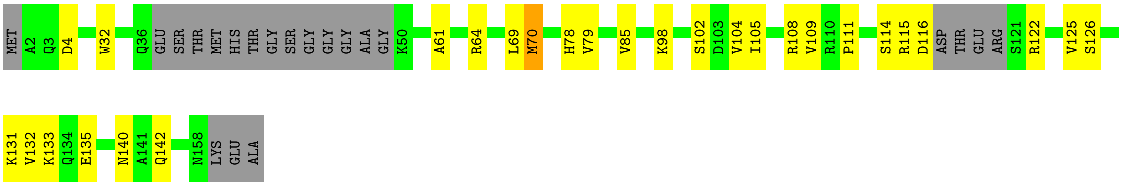
- Molecule 1: Hcp1 family type VI secretion system effector

Chain L:



- Molecule 1: Hcp1 family type VI secretion system effector

Chain J:



4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	82.08Å 82.21Å 100.12Å 89.99° 90.31° 119.73°	Depositor
Resolution (Å)	41.11 – 2.70 41.11 – 2.70	Depositor EDS
% Data completeness (in resolution range)	71.3 (41.11-2.70) 70.5 (41.11-2.70)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	2.91 (at 2.69Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.228 , 0.267 0.225 , 0.270	Depositor DCC
R_{free} test set	2344 reflections (5.25%)	wwPDB-VP
Wilson B-factor (Å^2)	38.1	Xtriage
Anisotropy	0.715	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 19.9	EDS
L-test for twinning[2]	$< \|L\| >$ = 0.49, $< L^2 >$ = 0.32	Xtriage
Estimated twinning fraction	0.145 for -k,h+k,l 0.145 for h+k,-h,l 0.145 for -h-k,h,l 0.145 for k,-h-k,l 0.149 for h,-h-k,-l 0.410 for -h-k,k,-l 0.147 for -h,-k,l 0.139 for -k,-h,-l 0.156 for k,h,-l 0.149 for -h,h+k,-l 0.148 for h+k,-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	13300	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.61% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	0.07	0/1103	0.25	0/1483
1	B	0.09	0/1117	0.25	0/1502
1	C	0.08	0/1108	0.26	0/1491
1	D	0.08	0/1119	0.24	0/1505
1	E	0.08	0/1145	0.25	0/1541
1	F	0.07	0/1109	0.24	0/1491
1	G	0.08	0/1120	0.24	0/1505
1	H	0.08	0/1145	0.26	0/1541
1	I	0.07	0/1117	0.23	0/1502
1	J	0.07	0/1100	0.24	0/1480
1	K	0.08	0/1117	0.24	0/1502
1	L	0.08	0/1109	0.24	0/1491
All	All	0.08	0/13409	0.24	0/18034

There are no bond length outliers.

There are no bond angle outliers.

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	1089	0	1089	17	0
1	B	1103	0	1097	18	0
1	C	1094	0	1084	13	0
1	D	1105	0	1097	17	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	1130	0	1124	22	0
1	F	1095	0	1093	16	0
1	G	1106	0	1106	17	0
1	H	1130	0	1124	26	0
1	I	1103	0	1097	16	0
1	J	1086	0	1080	18	0
1	K	1103	0	1097	15	0
1	L	1095	0	1093	19	0
2	A	4	0	0	2	0
2	B	6	0	0	0	0
2	C	4	0	0	1	0
2	D	2	0	0	1	0
2	E	8	0	0	2	0
2	F	7	0	0	1	0
2	G	3	0	0	0	0
2	H	2	0	0	2	0
2	I	3	0	0	0	0
2	J	7	0	0	0	0
2	K	4	0	0	0	0
2	L	11	0	0	3	0
All	All	13300	0	13181	181	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 7.

The worst 5 of 181 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:B:102:SER:HB2	1:B:131:LYS:HB2	1.71	0.71
1:J:102:SER:HB2	1:J:131:LYS:HB2	1.74	0.70
1:C:33:GLU:OE2	1:C:35:GLN:NE2	2.26	0.69
1:E:68:ASN:HA	2:E:202:HOH:O	1.92	0.68
1:F:102:SER:HB2	1:F:131:LYS:HB2	1.75	0.68

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	134/161 (83%)	132 (98%)	2 (2%)	0	100	100
1	B	136/161 (84%)	135 (99%)	1 (1%)	0	100	100
1	C	135/161 (84%)	134 (99%)	1 (1%)	0	100	100
1	D	136/161 (84%)	135 (99%)	1 (1%)	0	100	100
1	E	141/161 (88%)	136 (96%)	5 (4%)	0	100	100
1	F	135/161 (84%)	134 (99%)	1 (1%)	0	100	100
1	G	136/161 (84%)	134 (98%)	2 (2%)	0	100	100
1	H	141/161 (88%)	140 (99%)	1 (1%)	0	100	100
1	I	136/161 (84%)	134 (98%)	2 (2%)	0	100	100
1	J	134/161 (83%)	132 (98%)	2 (2%)	0	100	100
1	K	136/161 (84%)	133 (98%)	3 (2%)	0	100	100
1	L	135/161 (84%)	133 (98%)	2 (2%)	0	100	100
All	All	1635/1932 (85%)	1612 (99%)	23 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	119/133 (90%)	118 (99%)	1 (1%)	79	91
1	B	121/133 (91%)	120 (99%)	1 (1%)	79	91

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	120/133 (90%)	118 (98%)	2 (2%)	56	81
1	D	121/133 (91%)	119 (98%)	2 (2%)	56	81
1	E	124/133 (93%)	122 (98%)	2 (2%)	58	82
1	F	120/133 (90%)	120 (100%)	0	100	100
1	G	121/133 (91%)	120 (99%)	1 (1%)	79	91
1	H	124/133 (93%)	123 (99%)	1 (1%)	79	91
1	I	121/133 (91%)	119 (98%)	2 (2%)	56	81
1	J	119/133 (90%)	118 (99%)	1 (1%)	79	91
1	K	121/133 (91%)	121 (100%)	0	100	100
1	L	120/133 (90%)	118 (98%)	2 (2%)	56	81
All	All	1451/1596 (91%)	1436 (99%)	15 (1%)	73	89

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	151	THR
1	L	101	MET
1	G	70	MET
1	J	70	MET
1	I	81	GLN

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	L	81	GLN
1	J	158	ASN
1	G	81	GLN
1	K	158	ASN
1	H	139	GLN

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 oligosaccharides 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.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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, 95th 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($Å^2$)	Q<0.9
1	A	140/161 (86%)	-1.64	0	100	100	27, 37, 54, 69	0
1	B	142/161 (88%)	-1.59	0	100	100	28, 39, 58, 82	0
1	C	141/161 (87%)	-1.55	0	100	100	28, 40, 59, 75	0
1	D	142/161 (88%)	-1.56	0	100	100	26, 39, 55, 78	0
1	E	145/161 (90%)	-1.62	0	100	100	25, 39, 59, 78	0
1	F	141/161 (87%)	-1.57	0	100	100	28, 40, 59, 69	0
1	G	142/161 (88%)	-1.61	0	100	100	26, 39, 61, 74	0
1	H	145/161 (90%)	-1.57	0	100	100	26, 40, 59, 67	0
1	I	142/161 (88%)	-1.61	0	100	100	25, 40, 64, 78	0
1	J	140/161 (86%)	-1.59	0	100	100	25, 40, 59, 77	0
1	K	142/161 (88%)	-1.54	0	100	100	26, 39, 58, 79	0
1	L	141/161 (87%)	-1.54	0	100	100	27, 40, 60, 75	0
All	All	1703/1932 (88%)	-1.58	0	100	100	25, 40, 60, 82	0

There are no RSRZ outliers to report.

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no oligosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.