



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 5, 2021 – 12:11 PM JST

PDB ID : 7CJB  
Title : VHL recognizes hydroxyproline in RIPK1  
Authors : Liu, J.; Wang, Y.; Pan, L.  
Deposited on : 2020-07-09  
Resolution : 2.80 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](mailto: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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.22  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.22

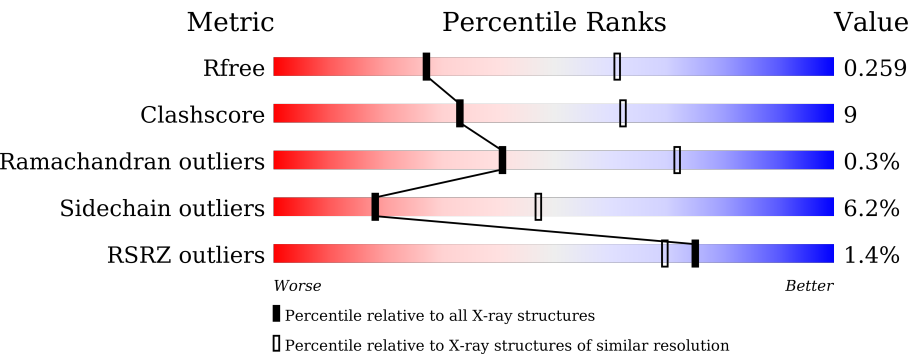
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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))
$R_{free}$	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (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  $\geq 3$ , 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  $\leq 5\%$ . 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	163	<div><div>%</div><div><div></div><div></div><div></div><div></div><div></div></div><div>67%20%•10%</div></div>
1	E	163	<div><div>2%</div><div><div></div><div></div><div></div><div></div><div></div></div><div>63%25%•9%</div></div>
1	I	163	<div><div></div><div><div></div><div></div><div></div><div></div><div></div></div><div>71%16%••10%</div></div>
1	M	163	<div><div></div><div><div></div><div></div><div></div><div></div><div></div></div><div>68%21%•10%</div></div>
2	B	122	<div><div>%</div><div><div></div><div></div><div></div><div></div><div></div></div><div>76%10%•13%</div></div>
2	F	122	<div><div>2%</div><div><div></div><div></div><div></div><div></div><div></div></div><div>67%14%•18%</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	J	122	<div><div></div><div>71%13%14%</div></div>
2	N	122	<div><div></div><div>66%16%15%</div></div>
3	C	100	<div><div></div><div>%78%10%10%</div></div>
3	G	100	<div><div></div><div>2%77%10%12%</div></div>
3	K	100	<div><div></div><div>2%78%9%11%</div></div>
3	O	100	<div><div></div><div>%77%9%13%</div></div>
4	D	15	<div><div></div><div>20%13%33%20%33%</div></div>
4	H	15	<div><div></div><div>7%27%20%53%</div></div>
4	L	15	<div><div></div><div>13%33%20%7%40%</div></div>
4	P	15	<div><div></div><div>27%20%53%</div></div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 11190 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 von Hippel-Lindau disease tumor suppressor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	147	Total	C	N	O	S	0	0	0
			1208	767	223	216	2			
1	E	149	Total	C	N	O	S	0	0	0
			1214	771	223	218	2			
1	I	147	Total	C	N	O	S	0	0	0
			1199	763	220	214	2			
1	M	147	Total	C	N	O	S	0	0	0
			1198	762	220	214	2			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	51	GLY	-	expression tag	UNP P40337
A	52	PRO	-	expression tag	UNP P40337
A	53	GLY	-	expression tag	UNP P40337
A	54	SER	-	expression tag	UNP P40337
E	51	GLY	-	expression tag	UNP P40337
E	52	PRO	-	expression tag	UNP P40337
E	53	GLY	-	expression tag	UNP P40337
E	54	SER	-	expression tag	UNP P40337
I	51	GLY	-	expression tag	UNP P40337
I	52	PRO	-	expression tag	UNP P40337
I	53	GLY	-	expression tag	UNP P40337
I	54	SER	-	expression tag	UNP P40337
M	51	GLY	-	expression tag	UNP P40337
M	52	PRO	-	expression tag	UNP P40337
M	53	GLY	-	expression tag	UNP P40337
M	54	SER	-	expression tag	UNP P40337

- Molecule 2 is a protein called Elongin-B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	106	Total	C	N	O	S	0	0	0
			828	526	140	159	3			
2	F	100	Total	C	N	O	S	0	0	0
			784	498	131	152	3			
2	J	105	Total	C	N	O	S	0	0	0
			822	523	139	157	3			
2	N	104	Total	C	N	O	S	0	0	0
			814	518	138	156	2			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-3	GLY	-	expression tag	UNP Q15370
B	-2	PRO	-	expression tag	UNP Q15370
B	-1	GLY	-	expression tag	UNP Q15370
B	0	SER	-	expression tag	UNP Q15370
B	60	ALA	CYS	conflict	UNP Q15370
B	89	SER	CYS	conflict	UNP Q15370
F	-3	GLY	-	expression tag	UNP Q15370
F	-2	PRO	-	expression tag	UNP Q15370
F	-1	GLY	-	expression tag	UNP Q15370
F	0	SER	-	expression tag	UNP Q15370
F	60	ALA	CYS	conflict	UNP Q15370
F	89	SER	CYS	conflict	UNP Q15370
J	-3	GLY	-	expression tag	UNP Q15370
J	-2	PRO	-	expression tag	UNP Q15370
J	-1	GLY	-	expression tag	UNP Q15370
J	0	SER	-	expression tag	UNP Q15370
J	60	ALA	CYS	conflict	UNP Q15370
J	89	SER	CYS	conflict	UNP Q15370
N	-3	GLY	-	expression tag	UNP Q15370
N	-2	PRO	-	expression tag	UNP Q15370
N	-1	GLY	-	expression tag	UNP Q15370
N	0	SER	-	expression tag	UNP Q15370
N	60	ALA	CYS	conflict	UNP Q15370
N	89	SER	CYS	conflict	UNP Q15370

- Molecule 3 is a protein called Elongin-C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	90	Total	C	N	O	S	0	0	0
			713	457	114	136	6			

*Continued on next page...*

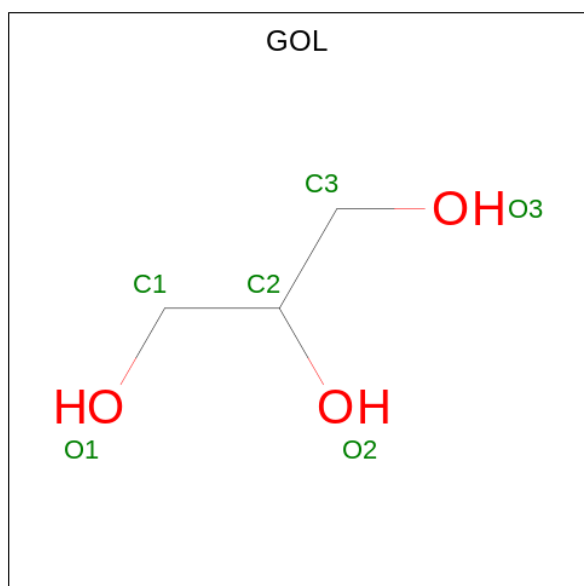
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	88	Total	C	N	O	S	0	0	0
			700	450	112	132	6			
3	K	89	Total	C	N	O	S	0	0	0
			704	453	113	132	6			
3	O	87	Total	C	N	O	S	0	0	0
			693	446	111	130	6			

- Molecule 4 is a protein called THR-LEU-TYR-TYR-MET-ALA-PRO-GLU-HIS-LEU-ASN-ASP-VAL-ASN-ALA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	10	Total	C	N	O	S	0	0	0
			88	58	13	16	1			
4	H	7	Total	C	N	O	S	0	1	0
			69	46	9	12	2			
4	L	9	Total	C	N	O	S	0	0	0
			80	54	11	14	1			
4	P	7	Total	C	N	O	S	0	0	0
			64	42	9	12	1			

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).

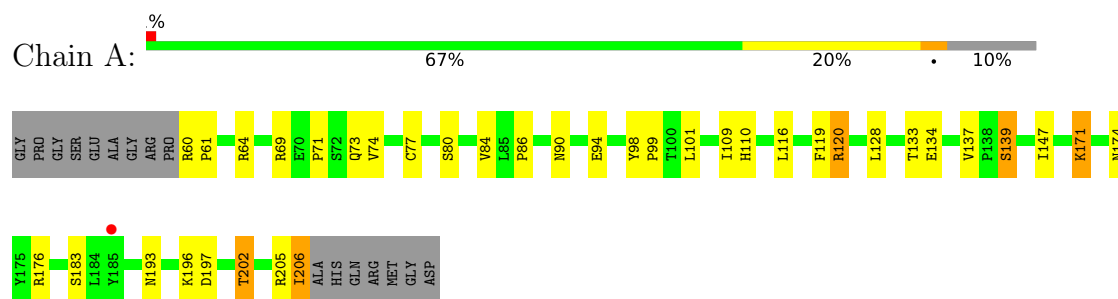


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	E	1	Total	C	O	0	0
			6	3	3		
5	I	1	Total	C	O	0	0
			6	3	3		

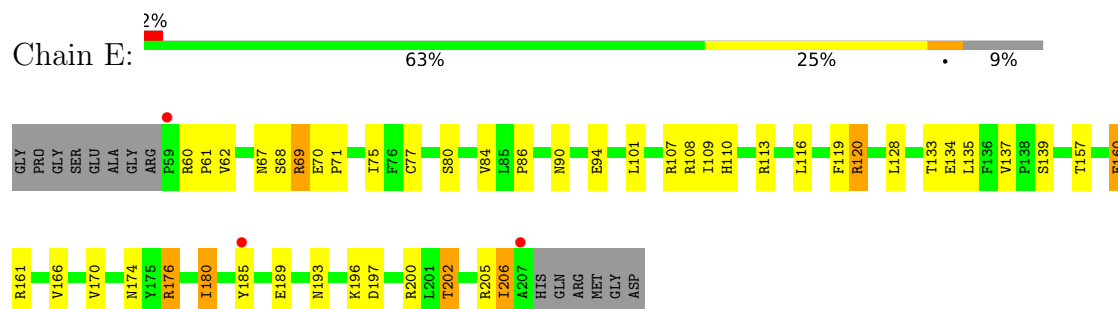
### 3 Residue-property plots

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.

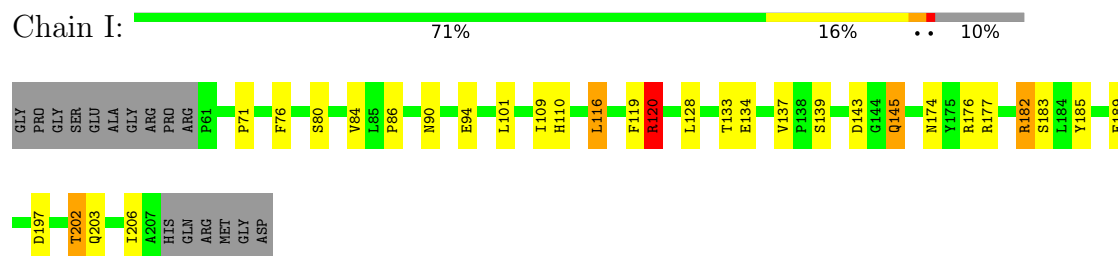
- Molecule 1: von Hippel-Lindau disease tumor suppressor



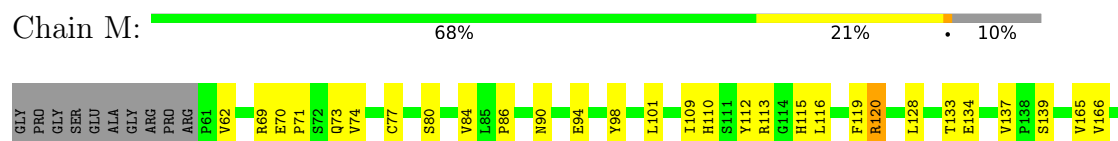
- Molecule 1: von Hippel-Lindau disease tumor suppressor



- Molecule 1: von Hippel-Lindau disease tumor suppressor

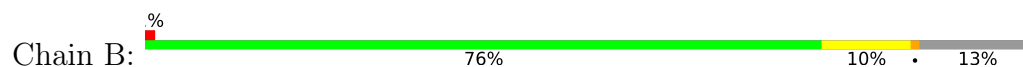


- Molecule 1: von Hippel-Lindau disease tumor suppressor

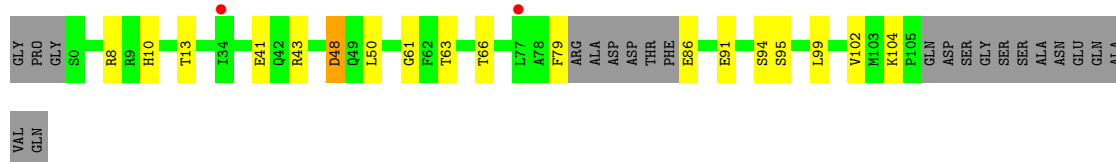




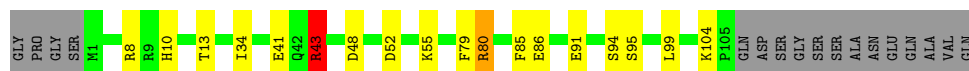
• Molecule 2: Elongin-B



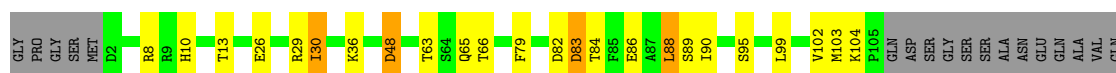
• Molecule 2: Elongin-B



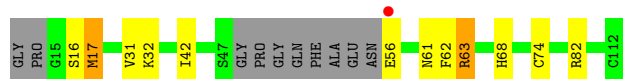
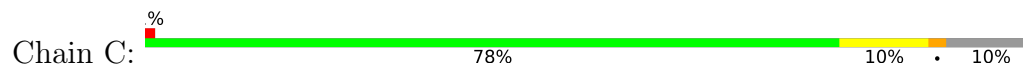
• Molecule 2: Elongin-B



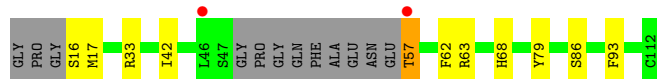
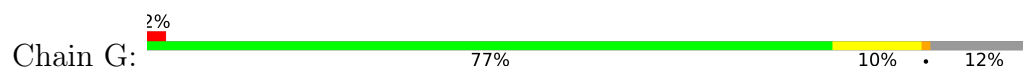
• Molecule 2: Elongin-B



• Molecule 3: Elongin-C

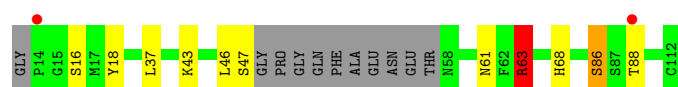
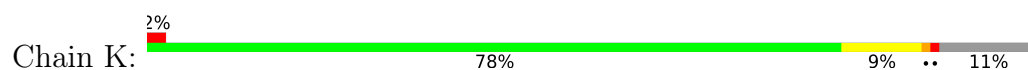


• Molecule 3: Elongin-C

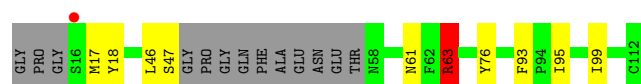
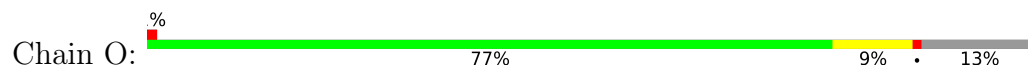


• Molecule 3: Elongin-C

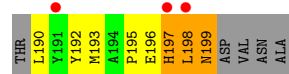




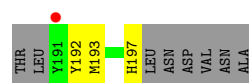
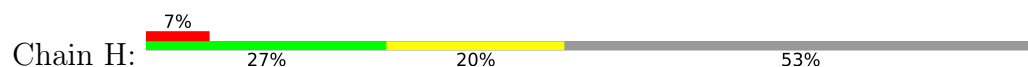
• Molecule 3: Elongin-C



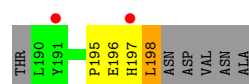
• Molecule 4: THR-LEU-TYR-TYR-MET-ALA-PRO-GLU-HIS-LEU-ASN-ASP-VAL-ASN-ALA



• Molecule 4: THR-LEU-TYR-TYR-MET-ALA-PRO-GLU-HIS-LEU-ASN-ASP-VAL-ASN-ALA



• Molecule 4: THR-LEU-TYR-TYR-MET-ALA-PRO-GLU-HIS-LEU-ASN-ASP-VAL-ASN-ALA



• Molecule 4: THR-LEU-TYR-TYR-MET-ALA-PRO-GLU-HIS-LEU-ASN-ASP-VAL-ASN-ALA



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	94.83Å 94.83Å 361.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.06 – 2.80 47.01 – 2.80	Depositor EDS
% Data completeness (in resolution range)	86.9 (47.06-2.80) 87.0 (47.01-2.80)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.26 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.8.0257	Depositor
R, $R_{free}$	0.220 , 0.259 0.221 , 0.259	Depositor DCC
$R_{free}$ test set	1783 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	42.7	Xtriage
Anisotropy	0.123	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 27.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	11190	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 28.99 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.6751e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: HYP, GOL

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	0.46	0/1239	0.87	2/1690 (0.1%)
1	E	0.42	0/1245	0.84	0/1698
1	I	0.48	0/1230	0.91	2/1678 (0.1%)
1	M	0.41	0/1229	0.87	1/1677 (0.1%)
2	B	0.48	0/845	0.88	1/1143 (0.1%)
2	F	0.47	0/799	0.85	0/1079
2	J	0.47	0/839	0.81	1/1135 (0.1%)
2	N	0.42	0/831	0.80	0/1125
3	C	0.55	0/727	0.82	1/980 (0.1%)
3	G	0.46	0/714	0.76	0/963
3	K	0.51	0/719	0.79	1/969 (0.1%)
3	O	0.46	0/707	0.77	1/953 (0.1%)
4	D	0.43	0/81	1.13	1/107 (0.9%)
4	H	0.48	0/65	0.84	0/84
4	L	0.47	0/73	0.78	0/96
4	P	0.39	0/57	0.72	0/74
All	All	0.46	0/11400	0.84	11/15451 (0.1%)

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	I	120	ARG	NE-CZ-NH2	-6.54	117.03	120.30
3	K	63	ARG	CG-CD-NE	6.45	125.34	111.80
3	C	63	ARG	NE-CZ-NH2	6.38	123.49	120.30
2	B	43	ARG	CG-CD-NE	-6.13	98.93	111.80
1	A	176	ARG	CG-CD-NE	6.01	124.42	111.80

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	1208	0	1212	30	0
1	E	1214	0	1213	41	0
1	I	1199	0	1203	25	0
1	M	1198	0	1201	30	0
2	B	828	0	832	8	0
2	F	784	0	793	12	0
2	J	822	0	827	13	0
2	N	814	0	815	12	0
3	C	713	0	709	5	0
3	G	700	0	700	6	0
3	K	704	0	704	8	0
3	O	693	0	693	6	0
4	D	88	0	79	13	0
4	H	69	0	60	7	0
4	L	80	0	73	6	0
4	P	64	0	51	3	0
5	E	6	0	8	1	0
5	I	6	0	8	0	0
All	All	11190	0	11181	192	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 192 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:M:62:VAL:CG1	1:M:202:THR:HG23	1.79	1.11
1:M:62:VAL:HG12	1:M:202:THR:HG23	1.29	1.10
1:M:166:VAL:CG1	1:M:188:LEU:HD21	1.83	1.07
1:E:193:ASN:HB3	1:E:196:LYS:HG2	1.35	1.02
1:A:193:ASN:HB3	1:A:196:LYS:HG2	1.46	0.97

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	145/163 (89%)	141 (97%)	4 (3%)	0	100	100
1	E	147/163 (90%)	144 (98%)	3 (2%)	0	100	100
1	I	145/163 (89%)	139 (96%)	6 (4%)	0	100	100
1	M	145/163 (89%)	142 (98%)	3 (2%)	0	100	100
2	B	104/122 (85%)	99 (95%)	4 (4%)	1 (1%)	15	44
2	F	96/122 (79%)	92 (96%)	4 (4%)	0	100	100
2	J	103/122 (84%)	98 (95%)	5 (5%)	0	100	100
2	N	102/122 (84%)	98 (96%)	3 (3%)	1 (1%)	15	44
3	C	86/100 (86%)	84 (98%)	1 (1%)	1 (1%)	13	39
3	G	84/100 (84%)	83 (99%)	1 (1%)	0	100	100
3	K	85/100 (85%)	83 (98%)	1 (1%)	1 (1%)	13	39
3	O	83/100 (83%)	82 (99%)	1 (1%)	0	100	100
4	D	7/15 (47%)	6 (86%)	1 (14%)	0	100	100
4	H	5/15 (33%)	5 (100%)	0	0	100	100
4	L	6/15 (40%)	6 (100%)	0	0	100	100
4	P	4/15 (27%)	3 (75%)	1 (25%)	0	100	100
All	All	1347/1600 (84%)	1305 (97%)	38 (3%)	4 (0%)	41	72

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	16	SER
3	K	16	SER
2	B	82	ASP
2	N	83	ASP

### 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	138/148 (93%)	128 (93%)	10 (7%)	14	38
1	E	137/148 (93%)	126 (92%)	11 (8%)	12	34
1	I	136/148 (92%)	129 (95%)	7 (5%)	24	55
1	M	136/148 (92%)	127 (93%)	9 (7%)	16	44
2	B	91/104 (88%)	89 (98%)	2 (2%)	52	83
2	F	88/104 (85%)	82 (93%)	6 (7%)	16	42
2	J	90/104 (86%)	85 (94%)	5 (6%)	21	51
2	N	89/104 (86%)	80 (90%)	9 (10%)	7	22
3	C	81/87 (93%)	76 (94%)	5 (6%)	18	47
3	G	80/87 (92%)	76 (95%)	4 (5%)	24	56
3	K	80/87 (92%)	77 (96%)	3 (4%)	33	67
3	O	79/87 (91%)	76 (96%)	3 (4%)	33	67
4	D	8/12 (67%)	6 (75%)	2 (25%)	0	2
4	H	6/12 (50%)	6 (100%)	0	100	100
4	L	7/12 (58%)	6 (86%)	1 (14%)	3	10
4	P	5/12 (42%)	5 (100%)	0	100	100
All	All	1251/1404 (89%)	1174 (94%)	77 (6%)	18	47

5 of 77 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	77	CYS
2	N	84	THR
1	M	101	LEU
2	N	10	HIS
3	O	63	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
3	G	35	HIS
4	H	197	HIS
3	O	35	HIS
3	K	27	HIS
3	C	58	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

4 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
4	HYP	L	195	4	6,8,9	1.04	1 (16%)	5,10,12	1.56	2 (40%)
4	HYP	H	195	4	6,8,9	0.52	0	5,10,12	0.85	0
4	HYP	D	195	4	6,8,9	0.57	0	5,10,12	0.97	1 (20%)
4	HYP	P	195	4	6,8,9	0.93	1 (16%)	5,10,12	1.84	2 (40%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	HYP	L	195	4	-	0/0/11/13	0/1/1/1
4	HYP	H	195	4	-	0/0/11/13	0/1/1/1
4	HYP	D	195	4	-	0/0/11/13	0/1/1/1
4	HYP	P	195	4	-	0/0/11/13	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	L	195	HYP	O-C	2.21	1.28	1.19
4	P	195	HYP	O-C	2.10	1.28	1.19

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	P	195	HYP	CG-CB-CA	2.80	107.49	103.96
4	P	195	HYP	CB-CG-CD	2.75	106.64	103.27
4	L	195	HYP	CG-CB-CA	2.55	107.18	103.96
4	L	195	HYP	CB-CG-CD	2.23	106.00	103.27
4	D	195	HYP	O-C-CA	-2.06	119.38	124.78

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
5	GOL	I	301	-	5,5,5	0.10	0	5,5,5	0.40	0
5	GOL	E	301	-	5,5,5	0.13	0	5,5,5	0.37	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns.



'-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	GOL	I	301	-	-	4/4/4/4	-
5	GOL	E	301	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	E	301	GOL	O1-C1-C2-C3
5	E	301	GOL	C1-C2-C3-O3
5	I	301	GOL	O1-C1-C2-C3
5	I	301	GOL	C1-C2-C3-O3
5	E	301	GOL	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	301	GOL	1	0

## 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 ⓘ

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	147/163 (90%)	-0.19	1 (0%) 87 84	18, 35, 91, 114	0
1	E	149/163 (91%)	-0.22	3 (2%) 65 56	23, 39, 101, 121	0
1	I	147/163 (90%)	-0.27	0 100 100	19, 36, 73, 109	0
1	M	147/163 (90%)	-0.12	0 100 100	33, 51, 94, 112	0
2	B	106/122 (86%)	-0.38	1 (0%) 84 80	18, 38, 83, 97	0
2	F	100/122 (81%)	0.15	2 (2%) 65 56	35, 66, 101, 108	0
2	J	105/122 (86%)	-0.40	0 100 100	24, 46, 85, 109	0
2	N	104/122 (85%)	-0.01	0 100 100	42, 69, 107, 127	0
3	C	90/100 (90%)	-0.35	1 (1%) 80 75	18, 37, 71, 103	0
3	G	88/100 (88%)	-0.06	2 (2%) 60 51	29, 55, 80, 99	0
3	K	89/100 (89%)	-0.30	2 (2%) 62 52	21, 36, 68, 93	0
3	O	87/100 (87%)	-0.03	1 (1%) 80 75	37, 52, 76, 89	0
4	D	9/15 (60%)	1.20	3 (33%) 0 0	48, 75, 96, 102	0
4	H	6/15 (40%)	1.15	1 (16%) 1 1	48, 57, 92, 130	0
4	L	8/15 (53%)	0.95	2 (25%) 0 0	50, 82, 98, 110	0
4	P	6/15 (40%)	0.55	0 100 100	71, 85, 113, 117	0
All	All	1388/1600 (86%)	-0.16	19 (1%) 75 70	18, 46, 93, 130	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	H	191	TYR	5.0
4	D	198	LEU	5.0
3	C	56	GLU	4.4
3	G	57	THR	3.3
1	A	185	TYR	3.2

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	HYP	P	195	8/9	0.83	0.22	63,66,68,68	0
4	HYP	L	195	8/9	0.87	0.20	43,45,49,50	0
4	HYP	H	195	8/9	0.97	0.11	41,43,44,45	0
4	HYP	D	195	8/9	0.98	0.18	46,47,49,49	0

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	GOL	E	301	6/6	0.77	0.27	55,59,61,62	0
5	GOL	I	301	6/6	0.93	0.24	55,58,60,60	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.