Full wwPDB X-ray Structure Validation Report ⓘ

May 23, 2024 – 12:31 PM EDT

PDB ID	:	9BJU
Title	:	Crystal structure of the complex between VHL, ElonginB, ElonginC, and compound 5
Authors	:	Murray, J.M.; Wu, H.; Fuhrmann, J.; Fairbrother, W.J.; DiPasquale, A.
Deposited on	:	2024-04-25
Resolution	:	2.47 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

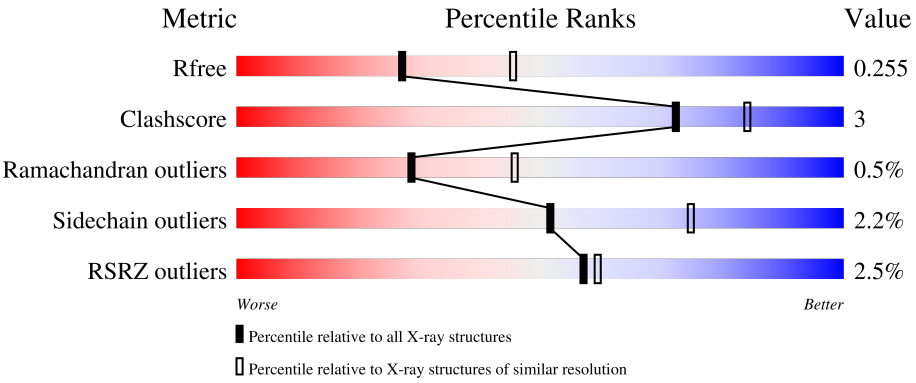
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.47 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.



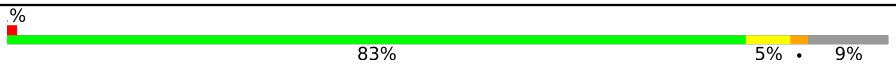
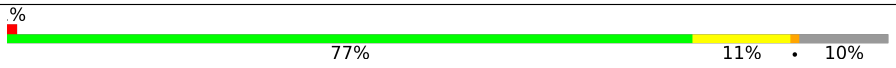
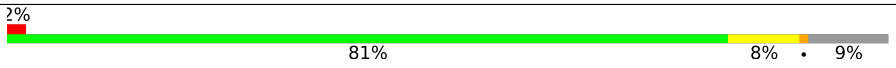
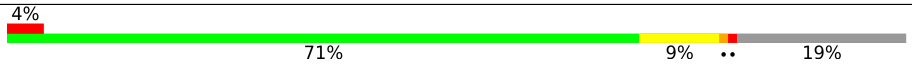
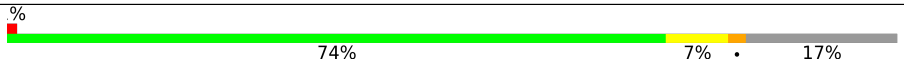
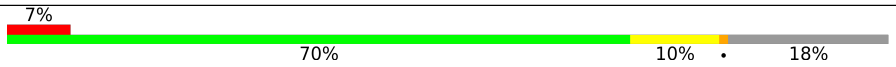
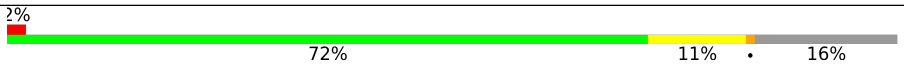
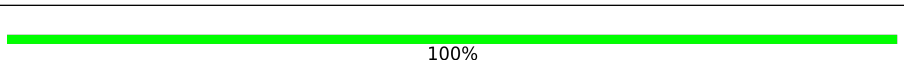
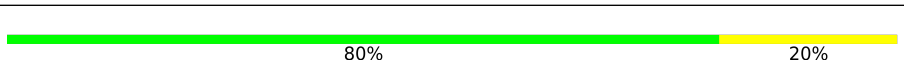
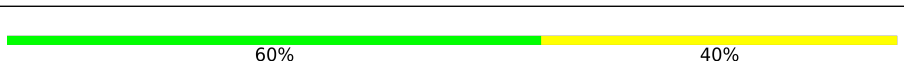
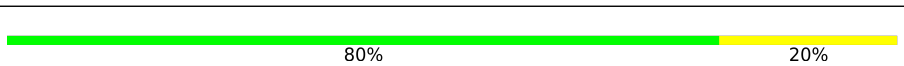
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5857 (2.50-2.46)
Clashscore	141614	6594 (2.50-2.46)
Ramachandran outliers	138981	6469 (2.50-2.46)
Sidechain outliers	138945	6471 (2.50-2.46)
RSRZ outliers	127900	5738 (2.50-2.46)

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	104	97%
1	D	104	3% 93% 5%
1	G	104	91% 8%
1	J	104	91% 8%
2	B	96	82% 9% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	96	
2	H	96	
2	K	96	
3	C	176	
3	L	176	
4	F	176	
4	I	176	
5	W	5	
5	X	5	
5	Y	5	
5	Z	5	

2 Entry composition (i)

There are 11 unique types of molecules in this entry. The entry contains 11148 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 Elongin-B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	104	Total	C	N	O	S	0	0	0
			822	520	138	159	5			
1	D	103	Total	C	N	O	S	0	0	0
			813	514	136	158	5			
1	G	103	Total	C	N	O	S	0	0	0
			813	514	136	158	5			
1	J	104	Total	C	N	O	S	0	0	0
			822	520	138	159	5			

- Molecule 2 is a protein called Elongin-C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	88	Total	C	N	O	S	0	0	0
			703	452	112	133	6			
2	E	87	Total	C	N	O	S	0	0	0
			694	447	111	130	6			
2	H	86	Total	C	N	O	S	0	0	0
			687	443	110	128	6			
2	K	87	Total	C	N	O	S	0	0	0
			694	447	111	130	6			

- Molecule 3 is a protein called von Hippel-Lindau disease tumor suppressor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	142	Total	C	N	O	S	0	1	0
			1181	750	218	211	2			
3	L	146	Total	C	N	O	S	0	0	0
			1196	759	219	216	2			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	38	MET	-	initiating methionine	UNP P40337
C	39	HIS	-	expression tag	UNP P40337
C	40	HIS	-	expression tag	UNP P40337
C	41	HIS	-	expression tag	UNP P40337
C	42	HIS	-	expression tag	UNP P40337
C	43	HIS	-	expression tag	UNP P40337
C	44	HIS	-	expression tag	UNP P40337
C	45	GLY	-	expression tag	UNP P40337
C	46	GLU	-	expression tag	UNP P40337
C	47	ASN	-	expression tag	UNP P40337
C	48	LEU	-	expression tag	UNP P40337
C	49	TYR	-	expression tag	UNP P40337
C	50	PHE	-	expression tag	UNP P40337
C	51	GLN	-	expression tag	UNP P40337
C	52	GLY	-	expression tag	UNP P40337
C	53	SER	-	expression tag	UNP P40337
L	38	MET	-	initiating methionine	UNP P40337
L	39	HIS	-	expression tag	UNP P40337
L	40	HIS	-	expression tag	UNP P40337
L	41	HIS	-	expression tag	UNP P40337
L	42	HIS	-	expression tag	UNP P40337
L	43	HIS	-	expression tag	UNP P40337
L	44	HIS	-	expression tag	UNP P40337
L	45	GLY	-	expression tag	UNP P40337
L	46	GLU	-	expression tag	UNP P40337
L	47	ASN	-	expression tag	UNP P40337
L	48	LEU	-	expression tag	UNP P40337
L	49	TYR	-	expression tag	UNP P40337
L	50	PHE	-	expression tag	UNP P40337
L	51	GLN	-	expression tag	UNP P40337
L	52	GLY	-	expression tag	UNP P40337
L	53	SER	-	expression tag	UNP P40337

- Molecule 4 is a protein called von Hippel-Lindau disease tumor suppressor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	144	Total	C	N	O	S	0	1	0
			1192	756	220	214	2			
4	I	147	Total	C	N	O	S	0	0	0
			1202	764	220	216	2			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	38	MET	-	initiating methionine	UNP P40337
F	39	HIS	-	expression tag	UNP P40337
F	40	HIS	-	expression tag	UNP P40337
F	41	HIS	-	expression tag	UNP P40337
F	42	HIS	-	expression tag	UNP P40337
F	43	HIS	-	expression tag	UNP P40337
F	44	HIS	-	expression tag	UNP P40337
F	45	GLY	-	expression tag	UNP P40337
F	46	GLU	-	expression tag	UNP P40337
F	47	ASN	-	expression tag	UNP P40337
F	48	LEU	-	expression tag	UNP P40337
F	49	TYR	-	expression tag	UNP P40337
F	50	PHE	-	expression tag	UNP P40337
F	51	GLN	-	expression tag	UNP P40337
F	52	GLY	-	expression tag	UNP P40337
F	53	SER	-	expression tag	UNP P40337
I	38	MET	-	initiating methionine	UNP P40337
I	39	HIS	-	expression tag	UNP P40337
I	40	HIS	-	expression tag	UNP P40337
I	41	HIS	-	expression tag	UNP P40337
I	42	HIS	-	expression tag	UNP P40337
I	43	HIS	-	expression tag	UNP P40337
I	44	HIS	-	expression tag	UNP P40337
I	45	GLY	-	expression tag	UNP P40337
I	46	GLU	-	expression tag	UNP P40337
I	47	ASN	-	expression tag	UNP P40337
I	48	LEU	-	expression tag	UNP P40337
I	49	TYR	-	expression tag	UNP P40337
I	50	PHE	-	expression tag	UNP P40337
I	51	GLN	-	expression tag	UNP P40337
I	52	GLY	-	expression tag	UNP P40337
I	53	SER	-	expression tag	UNP P40337

- Molecule 5 is a protein called 3-PYRIDIN-4-YL-2,4-DIHYDRO-INDENO[1,2-.C.]PYRAZOLE.

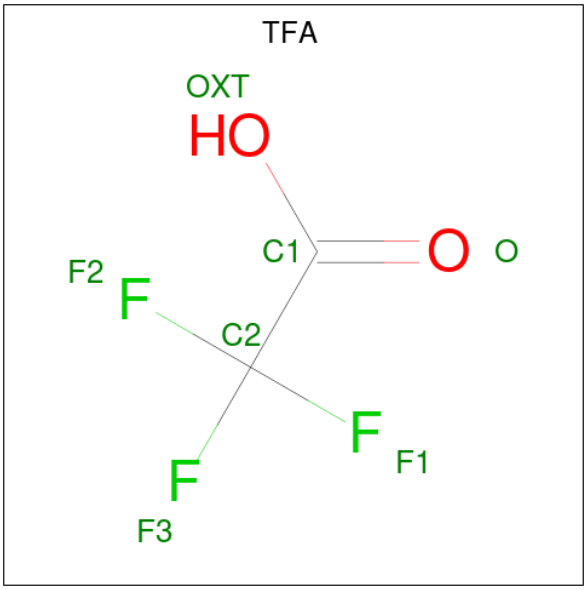
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	W	5	Total	C	N	O	0	0	1
			37	28	4	5			
5	X	5	Total	C	N	O	0	0	1
			37	28	4	5			
5	Y	5	Total	C	N	O	0	0	1
			37	28	4	5			

Continued on next page...

Continued from previous page...

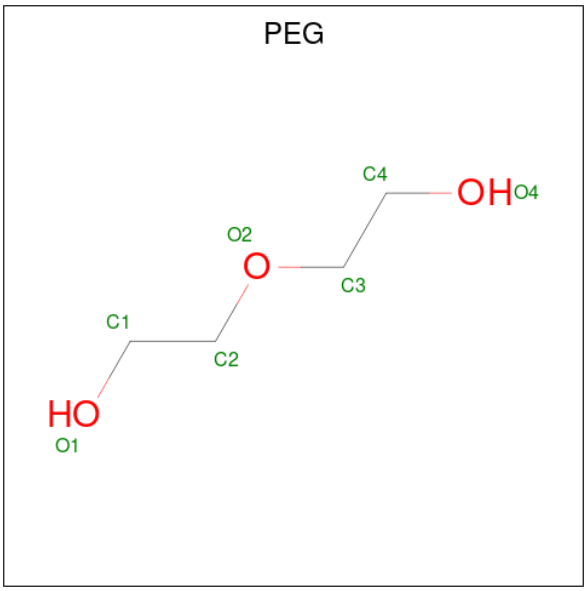
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
5	Z	5	Total C N O 37 28 4 5	0	0	1

- Molecule 6 is trifluoroacetic acid (three-letter code: TFA) (formula: $C_2HF_3O_2$).



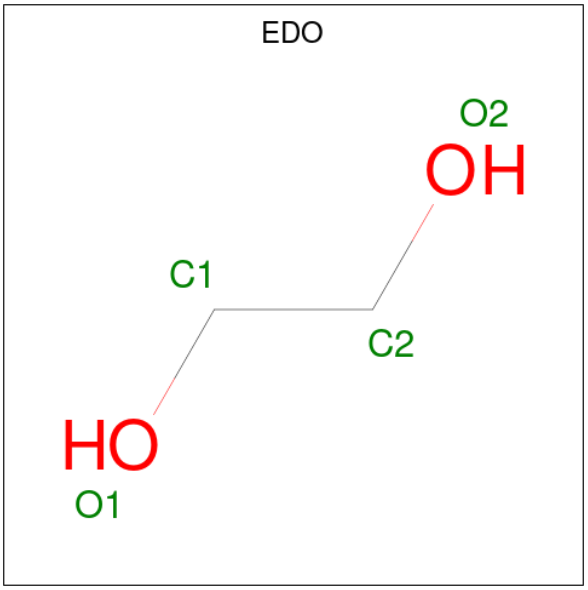
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	1	Total C F O 7 2 3 2	0	0
6	K	1	Total C F O 7 2 3 2	0	0

- Molecule 7 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



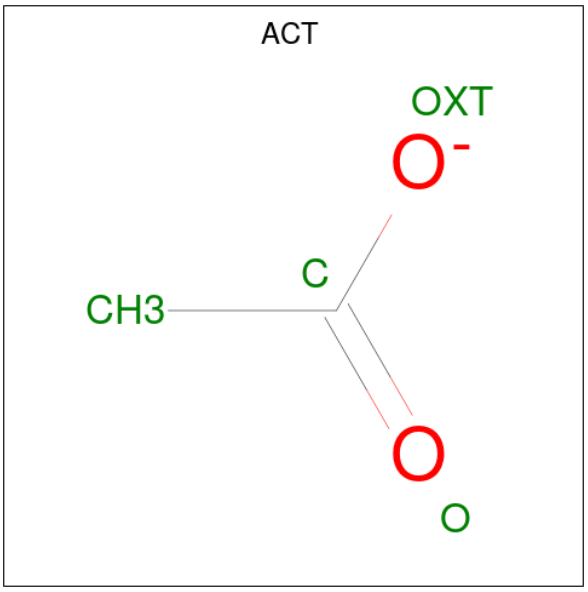
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	1	Total C O 7 4 3	0	0

- Molecule 8 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



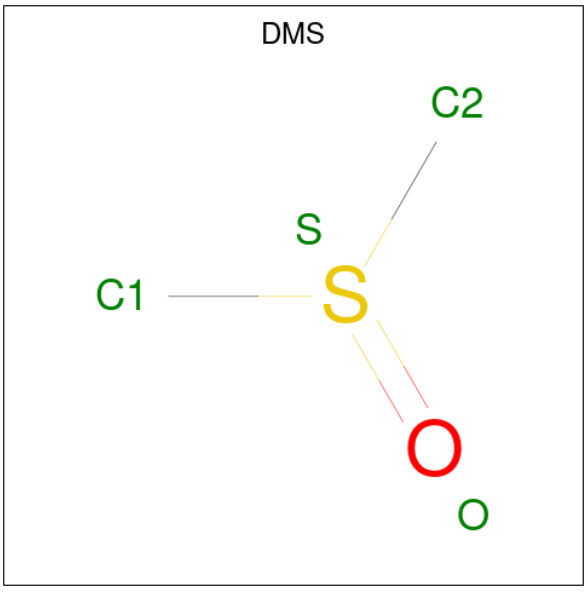
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	C	1	Total C O 4 2 2	0	0
8	K	1	Total C O 4 2 2	0	0

- Molecule 9 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	C	1	Total C O 4 2 2	0	0
9	F	1	Total C O 4 2 2	0	0
9	H	1	Total C O 4 2 2	0	0

- Molecule 10 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C_2H_6OS).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	H	1	Total C O S 4 2 1 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	I	1	Total C O S 4 2 1 1	0	0
10	L	1	Total C O S 4 2 1 1	0	0

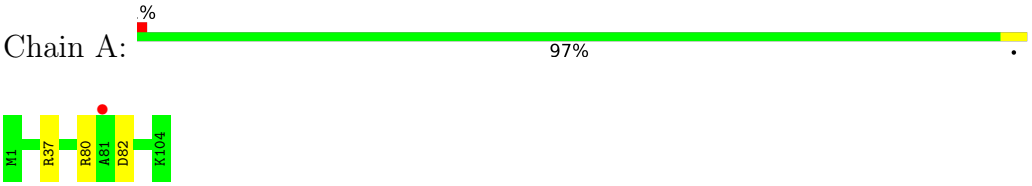
- Molecule 11 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	A	17	Total O 17 17	0	0
11	B	3	Total O 3 3	0	0
11	C	15	Total O 15 15	0	0
11	D	4	Total O 4 4	0	0
11	E	6	Total O 6 6	0	0
11	F	6	Total O 6 6	0	0
11	G	12	Total O 12 12	0	0
11	H	8	Total O 8 8	0	0
11	I	6	Total O 6 6	0	0
11	J	11	Total O 11 11	0	0
11	K	16	Total O 16 16	0	0
11	L	21	Total O 21 21	0	0
11	X	1	Total O 1 1	0	0
11	Y	1	Total O 1 1	0	0
11	Z	1	Total O 1 1	0	0

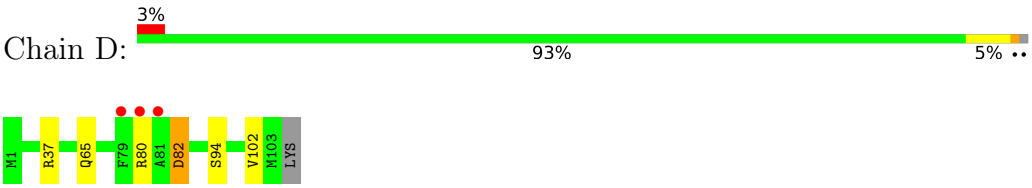
3 Residue-property plots (i)

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.

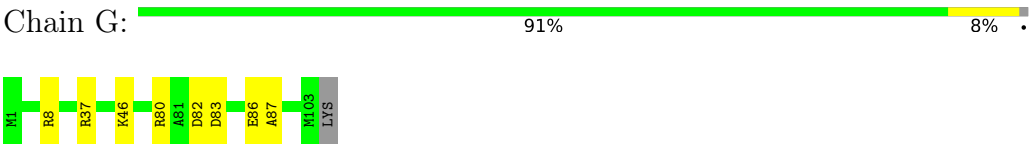
• Molecule 1: Elongin-B



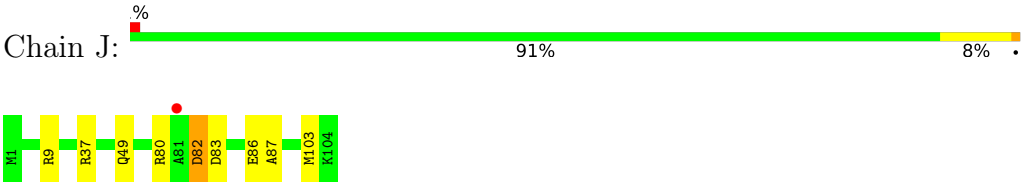
• Molecule 1: Elongin-B



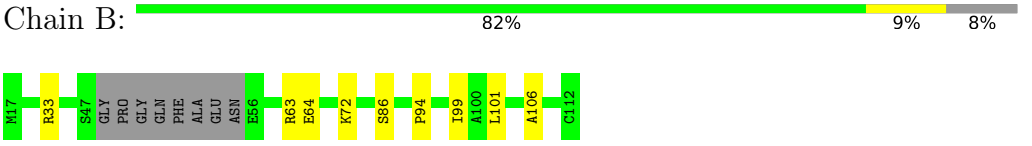
• Molecule 1: Elongin-B



• Molecule 1: Elongin-B

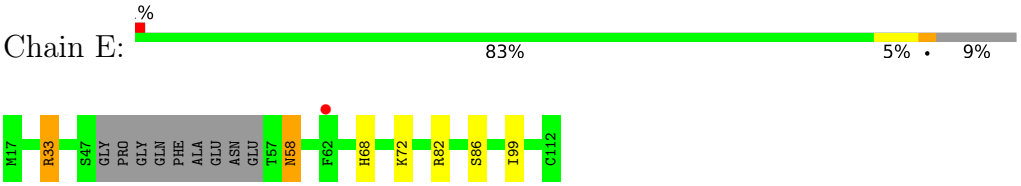


• Molecule 2: Elongin-C



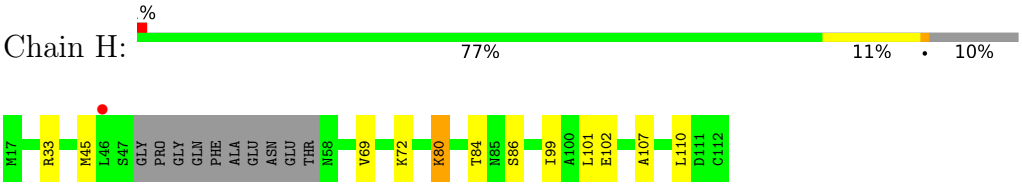
• Molecule 2: Elongin-C

Chain E:



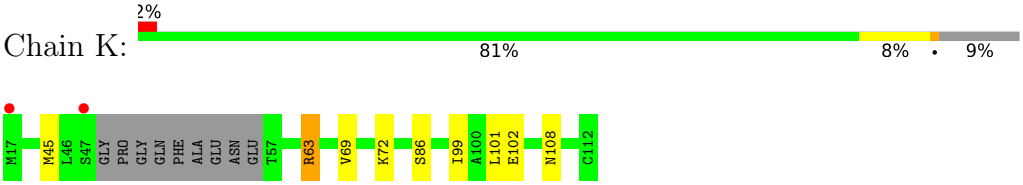
• Molecule 2: Elongin-C

Chain H:



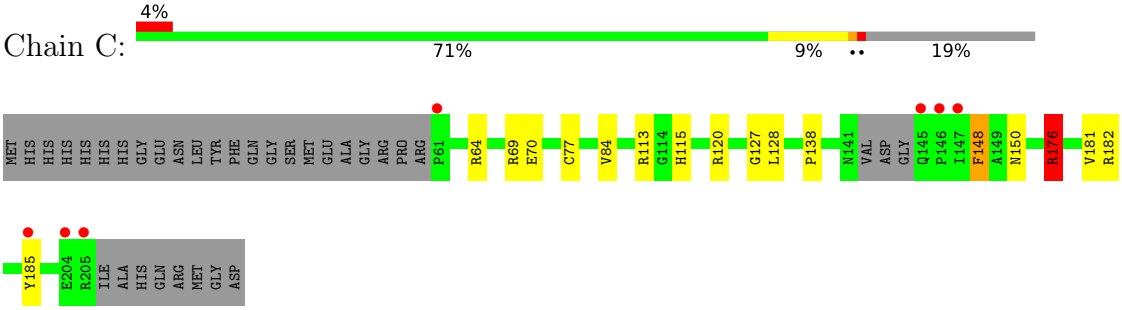
• Molecule 2: Elongin-C

Chain K:



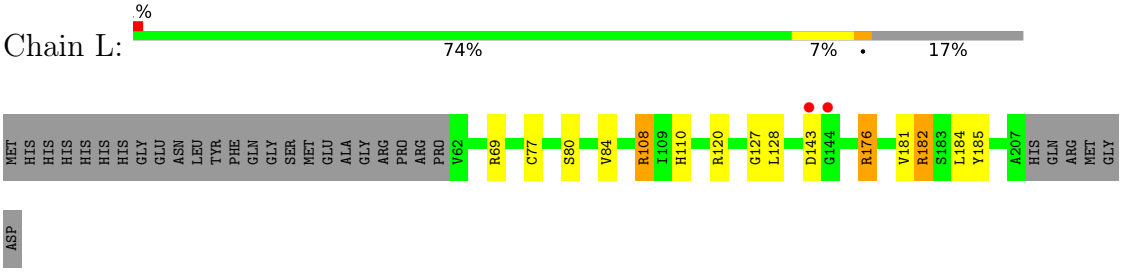
• Molecule 3: von Hippel-Lindau disease tumor suppressor

Chain C:



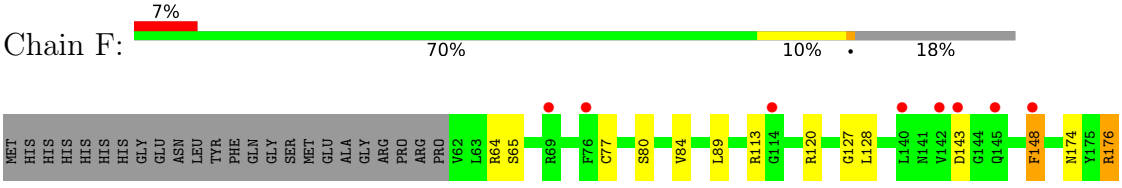
• Molecule 3: von Hippel-Lindau disease tumor suppressor

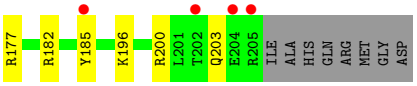
Chain L:



• Molecule 4: von Hippel-Lindau disease tumor suppressor

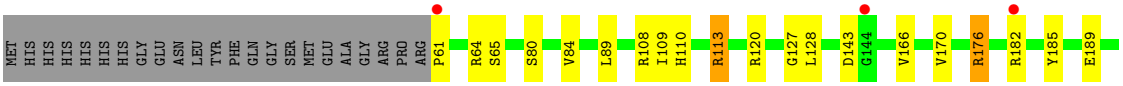
Chain F:

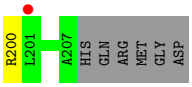




• Molecule 4: von Hippel-Lindau disease tumor suppressor

Chain I:





• Molecule 5: 3-PYRIDIN-4-YL-2,4-DIHYDRO-INDENO[1,2-.C.]PYRAZOLE

Chain W:

There are no outlier residues recorded for this chain.

• Molecule 5: 3-PYRIDIN-4-YL-2,4-DIHYDRO-INDENO[1,2-.C.]PYRAZOLE

Chain X:

• Molecule 5: 3-PYRIDIN-4-YL-2,4-DIHYDRO-INDENO[1,2-.C.]PYRAZOLE

Chain Y:

• Molecule 5: 3-PYRIDIN-4-YL-2,4-DIHYDRO-INDENO[1,2-.C.]PYRAZOLE

Chain Z:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 2 2	Depositor
Cell constants a, b, c, α, β, γ	93.30Å 93.30Å 362.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.86 – 2.47 48.81 – 2.47	Depositor EDS
% Data completeness (in resolution range)	99.6 (48.86-2.47) 99.7 (48.81-2.47)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.77 (at 2.48Å)	Xtriage
Refinement program	REFMAC 5.8.0222	Depositor
R, R_{free}	0.218 , 0.256 0.221 , 0.255	Depositor DCC
R_{free} test set	2942 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	55.9	Xtriage
Anisotropy	0.017	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 36.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	11148	wwPDB-VP
Average B, all atoms (Å^2)	69.0	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 32.86 % 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 8.7794e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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

Bond lengths and bond angles in the following residue types are not validated in this section: A1AQ4, CSX, DMS, PEG, TFA, NH2, EDO, HYP, ACT, TBG, ACE

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	0.63	0/838	0.74	0/1132
1	D	0.51	0/829	0.71	0/1121
1	G	0.50	0/829	0.71	0/1121
1	J	0.57	0/838	0.70	0/1132
2	B	0.57	0/717	0.65	0/967
2	E	0.54	0/708	0.67	0/955
2	H	0.51	0/701	0.72	0/945
2	K	0.59	0/708	0.70	0/955
3	C	0.51	0/1204	0.74	0/1640
3	L	0.60	0/1218	0.78	0/1660
4	F	0.52	0/1223	0.76	0/1669
4	I	0.53	0/1233	0.75	0/1682
All	All	0.55	0/11046	0.73	0/14979

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	G	0	1
1	J	0	1
2	B	0	1
2	E	0	2
2	H	0	1
2	K	0	1
3	C	0	2
3	L	0	3
4	F	0	3
4	I	0	3
All	All	0	18

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (18) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	33	ARG	Sidechain
3	C	176	ARG	Sidechain
3	C	69	ARG	Sidechain
2	E	33	ARG	Sidechain
2	E	82	ARG	Sidechain
4	F	113	ARG	Sidechain
4	F	176	ARG	Sidechain
4	F	177	ARG	Sidechain
1	G	8	ARG	Sidechain
2	H	33	ARG	Sidechain
4	I	108	ARG	Sidechain
4	I	176	ARG	Sidechain
4	I	200	ARG	Sidechain
1	J	9	ARG	Sidechain
2	K	63	ARG	Sidechain
3	L	108	ARG	Sidechain
3	L	176	ARG	Sidechain
3	L	69	ARG	Sidechain

5.2 Too-close contacts (i)

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	822	0	824	1	0
1	D	813	0	811	6	1
1	G	813	0	811	2	0
1	J	822	0	824	4	1
2	B	703	0	701	5	0
2	E	694	0	695	5	0
2	H	687	0	688	9	0
2	K	694	0	695	10	0
3	C	1181	0	1178	14	0
3	L	1196	0	1196	8	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	F	1192	0	1188	15	0
4	I	1202	0	1205	10	0
5	W	37	0	20	0	0
5	X	37	0	20	0	0
5	Y	37	0	20	0	0
5	Z	37	0	20	0	0
6	B	7	0	0	0	0
6	K	7	0	0	0	0
7	B	7	0	10	2	0
8	C	4	0	6	3	0
8	K	4	0	6	0	0
9	C	4	0	3	0	0
9	F	4	0	3	0	0
9	H	4	0	3	0	0
10	H	4	0	6	0	0
10	I	4	0	6	1	0
10	L	4	0	6	1	0
11	A	17	0	0	0	0
11	B	3	0	0	0	0
11	C	15	0	0	2	0
11	D	4	0	0	0	0
11	E	6	0	0	0	0
11	F	6	0	0	0	0
11	G	12	0	0	0	0
11	H	8	0	0	0	0
11	I	6	0	0	0	0
11	J	11	0	0	0	0
11	K	16	0	0	1	0
11	L	21	0	0	1	0
11	X	1	0	0	0	0
11	Y	1	0	0	0	0
11	Z	1	0	0	0	0
All	All	11148	0	10945	75	1

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 3.

All (75) 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)
4:F:148:PHE:HZ	2:H:45:MET:HG3	1.41	0.86

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:E:33:ARG:HH12	2:E:58:ASN:HD21	1.22	0.84
2:E:33:ARG:HH12	2:E:58:ASN:ND2	1.88	0.71
3:C:70:GLU:OE1	3:C:113:ARG:NH1	2.24	0.71
4:I:176:ARG:HH11	4:I:189:GLU:CD	1.93	0.71
4:F:148:PHE:HZ	2:H:45:MET:CG	2.05	0.68
4:I:113:ARG:HH11	4:I:113:ARG:HB3	1.60	0.67
2:H:107:ALA:O	2:H:110:LEU:O	2.14	0.65
2:K:108:ASN:ND2	3:L:184:LEU:HD21	2.13	0.62
3:C:148:PHE:CE1	2:K:45:MET:HG3	2.35	0.61
4:F:148:PHE:CZ	2:H:45:MET:HG3	2.28	0.61
4:I:176:ARG:NH1	4:I:189:GLU:OE1	2.33	0.61
3:L:110:HIS:H	10:L:301:DMS:C2	2.15	0.60
4:I:110:HIS:H	10:I:301:DMS:C2	2.14	0.60
1:J:103:MET:CE	2:K:101:LEU:HG	2.32	0.59
1:D:80:ARG:HD3	1:D:82:ASP:O	2.06	0.56
4:F:148:PHE:CZ	2:H:45:MET:CG	2.88	0.56
4:F:77:CYS:HG	4:F:148:PHE:HE1	1.55	0.55
3:C:182:ARG:HA	3:C:185:TYR:CD2	2.42	0.54
3:L:182:ARG:NH1	11:L:401:HOH:O	2.39	0.54
1:A:37:ARG:NH2	1:A:80:ARG:O	2.41	0.53
1:G:37:ARG:NH2	1:G:80:ARG:O	2.42	0.53
2:B:72:LYS:NZ	2:B:94:PRO:O	2.42	0.53
1:D:37:ARG:NH2	1:D:80:ARG:O	2.42	0.52
4:F:182:ARG:HA	4:F:185:TYR:CD2	2.44	0.52
4:I:182:ARG:HA	4:I:185:TYR:CD2	2.45	0.52
1:D:102:VAL:HG21	4:F:174:ASN:HB3	1.91	0.51
1:D:102:VAL:HG22	1:D:102:VAL:O	2.11	0.51
2:H:72:LYS:HG3	2:H:99:ILE:CD1	2.40	0.51
3:L:182:ARG:HA	3:L:185:TYR:CD2	2.46	0.50
1:J:37:ARG:NH2	1:J:80:ARG:O	2.43	0.50
2:K:72:LYS:HG3	2:K:99:ILE:CD1	2.42	0.50
4:F:176:ARG:HH22	4:F:185:TYR:HB3	1.77	0.50
3:C:148:PHE:CZ	2:K:45:MET:HG3	2.48	0.49
2:E:72:LYS:HG3	2:E:99:ILE:CD1	2.43	0.49
2:K:72:LYS:HE3	11:K:306:HOH:O	2.12	0.48
4:I:61:PRO:O	4:I:64:ARG:NH2	2.38	0.48
2:E:33:ARG:NH1	2:E:58:ASN:ND2	2.59	0.48
3:C:148:PHE:CZ	2:K:45:MET:CG	2.98	0.47
3:C:181:VAL:HB	8:C:301:EDO:H11	1.97	0.47
2:H:80:LYS:HD2	2:H:84:THR:OG1	2.14	0.47
3:C:182:ARG:N	8:C:301:EDO:H12	2.31	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:148:PHE:HD2	3:C:150:ASN:ND2	2.14	0.46
4:F:200:ARG:O	4:F:203:GLN:HG2	2.15	0.46
2:B:101:LEU:HD23	2:B:101:LEU:HA	1.84	0.45
2:B:106:ALA:HB2	7:B:202:PEG:H42	1.99	0.45
1:D:102:VAL:HG21	4:F:174:ASN:CB	2.46	0.45
4:F:65:SER:OG	4:F:89:LEU:O	2.34	0.45
2:K:69:VAL:HG21	2:K:102:GLU:HB3	1.99	0.44
4:I:166:VAL:O	4:I:170:VAL:HG12	2.18	0.44
4:F:84:VAL:HG22	4:F:128:LEU:CD1	2.48	0.44
4:F:120:ARG:HD3	4:F:127:GLY:HA2	1.99	0.44
2:B:64:GLU:OE1	7:B:202:PEG:H11	2.18	0.43
4:I:120:ARG:HD3	4:I:127:GLY:HA2	2.00	0.43
2:H:101:LEU:HD23	2:H:101:LEU:HA	1.91	0.43
3:C:181:VAL:HB	8:C:301:EDO:C1	2.49	0.43
3:L:84:VAL:HG22	3:L:128:LEU:CD1	2.48	0.43
3:C:115:HIS:O	3:C:138:PRO:HD2	2.19	0.42
2:K:108:ASN:HD22	3:L:184:LEU:HD21	1.84	0.42
3:C:176:ARG:HD3	11:C:403:HOH:O	2.18	0.42
4:F:176:ARG:NH2	4:F:185:TYR:HB3	2.34	0.42
3:L:120:ARG:HD3	3:L:127:GLY:HA2	2.01	0.42
3:C:84:VAL:HG22	3:C:128:LEU:CD1	2.50	0.42
4:F:77:CYS:SG	4:F:148:PHE:HE1	2.42	0.42
2:B:72:LYS:HG3	2:B:99:ILE:CD1	2.49	0.42
4:I:65:SER:OG	4:I:89:LEU:O	2.32	0.42
3:C:148:PHE:CD2	3:C:150:ASN:ND2	2.88	0.41
3:C:120:ARG:HD3	3:C:127:GLY:HA2	2.02	0.41
4:I:84:VAL:HG22	4:I:128:LEU:CD1	2.50	0.41
2:H:69:VAL:HG21	2:H:102:GLU:HB3	2.02	0.41
1:G:86:GLU:HG3	1:G:87:ALA:O	2.21	0.41
1:D:94:SER:O	2:E:68:HIS:HB3	2.20	0.40
11:C:415:HOH:O	3:L:181:VAL:HG23	2.21	0.40
1:J:103:MET:HE3	2:K:101:LEU:HG	2.01	0.40
1:J:86:GLU:HG3	1:J:87:ALA:O	2.21	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:65:GLN:O	1:J:49:GLN:NE2[8_655]	2.16	0.04

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	102/104 (98%)	97 (95%)	4 (4%)	1 (1%)	15	26
1	D	101/104 (97%)	95 (94%)	5 (5%)	1 (1%)	15	26
1	G	101/104 (97%)	95 (94%)	5 (5%)	1 (1%)	15	26
1	J	102/104 (98%)	98 (96%)	3 (3%)	1 (1%)	15	26
2	B	84/96 (88%)	82 (98%)	2 (2%)	0	100	100
2	E	83/96 (86%)	81 (98%)	1 (1%)	1 (1%)	13	22
2	H	82/96 (85%)	81 (99%)	1 (1%)	0	100	100
2	K	83/96 (86%)	81 (98%)	2 (2%)	0	100	100
3	C	138/176 (78%)	135 (98%)	3 (2%)	0	100	100
3	L	143/176 (81%)	140 (98%)	3 (2%)	0	100	100
4	F	143/176 (81%)	140 (98%)	3 (2%)	0	100	100
4	I	145/176 (82%)	140 (97%)	4 (3%)	1 (1%)	22	36
All	All	1307/1504 (87%)	1265 (97%)	36 (3%)	6 (0%)	29	46

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	82	ASP
2	E	58	ASN
1	G	82	ASP
4	I	143	ASP
1	J	82	ASP
1	D	82	ASP

5.3.2 Protein sidechains (i)

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	92/92 (100%)	92 (100%)	0	100	100
1	D	91/92 (99%)	91 (100%)	0	100	100
1	G	91/92 (99%)	89 (98%)	2 (2%)	52	75
1	J	92/92 (100%)	90 (98%)	2 (2%)	52	75
2	B	80/85 (94%)	78 (98%)	2 (2%)	47	71
2	E	79/85 (93%)	78 (99%)	1 (1%)	69	86
2	H	78/85 (92%)	76 (97%)	2 (3%)	46	70
2	K	79/85 (93%)	77 (98%)	2 (2%)	47	71
3	C	134/160 (84%)	131 (98%)	3 (2%)	52	75
3	L	135/160 (84%)	130 (96%)	5 (4%)	34	57
4	F	136/161 (84%)	131 (96%)	5 (4%)	34	57
4	I	137/161 (85%)	134 (98%)	3 (2%)	52	75
All	All	1224/1350 (91%)	1197 (98%)	27 (2%)	52	75

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	63	ARG
2	B	86	SER
3	C	64	ARG
3	C	148	PHE
3	C	176	ARG
2	E	86	SER
4	F	64	ARG
4	F	80	SER
4	F	143	ASP
4	F	148	PHE
4	F	196	LYS
1	G	46	LYS
1	G	83	ASP
2	H	80	LYS
2	H	86	SER
4	I	80	SER
4	I	109	ILE
4	I	113	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	J	82	ASP
1	J	83	ASP
2	K	63	ARG
2	K	86	SER
3	L	80	SER
3	L	108	ARG
3	L	143	ASP
3	L	176	ARG
3	L	182	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	85	ASN
3	C	150	ASN
2	E	58	ASN
2	E	85	ASN
4	F	174	ASN
4	I	110	HIS
3	L	110	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)

14 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	$\#\|Z\| > 2$
5	A1AQ4	Y	4	5	17,18,19	0.57	0	20,23,25	0.74	1 (5%)
5	HYP	Y	3	5	6,8,9	0.75	0	5,10,12	1.50	1 (20%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	A1AQ4	W	4	5	17,18,19	0.47	0	20,23,25	1.07	0
5	A1AQ4	Z	4	5	17,18,19	0.37	0	20,23,25	1.06	1 (5%)
3	CSX	C	77	3	3,6,7	0.67	0	1,6,8	3.47	1 (100%)
5	HYP	Z	3	5	6,8,9	0.58	0	5,10,12	1.39	0
5	A1AQ4	X	4	5	17,18,19	0.43	0	20,23,25	0.64	0
5	HYP	W	3	5	6,8,9	0.84	0	5,10,12	1.65	0
5	TBG	W	2	5	6,7,8	0.83	0	7,10,12	1.45	0
5	TBG	Y	2	5	6,7,8	0.52	0	7,10,12	1.16	0
5	HYP	X	3	5	6,8,9	0.74	0	5,10,12	1.27	1 (20%)
5	TBG	X	2	5	6,7,8	0.53	0	7,10,12	0.84	0
3	CSX	L	77	3	3,6,7	0.61	0	1,6,8	2.71	1 (100%)
5	TBG	Z	2	5	6,7,8	0.73	0	7,10,12	1.25	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	A1AQ4	Y	4	5	-	0/9/10/12	0/2/2/2
5	HYP	Y	3	5	-	0/0/11/13	0/1/1/1
5	A1AQ4	W	4	5	-	0/9/10/12	0/2/2/2
5	A1AQ4	Z	4	5	-	1/9/10/12	0/2/2/2
3	CSX	C	77	3	-	1/1/5/7	-
5	HYP	Z	3	5	-	0/0/11/13	0/1/1/1
5	A1AQ4	X	4	5	-	0/9/10/12	0/2/2/2
5	HYP	W	3	5	-	0/0/11/13	0/1/1/1
5	TBG	W	2	5	-	0/6/8/10	-
5	TBG	Y	2	5	-	0/6/8/10	-
5	HYP	X	3	5	-	0/0/11/13	0/1/1/1
5	TBG	X	2	5	-	0/6/8/10	-
3	CSX	L	77	3	-	1/1/5/7	-
5	TBG	Z	2	5	-	0/6/8/10	-

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	77	CSX	CA-CB-SG	-3.47	105.78	113.36
3	L	77	CSX	CA-CB-SG	-2.71	107.44	113.36
5	X	3	HYP	CB-CG-CD	-2.49	100.22	103.27

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	Y	3	HYP	CB-CG-CD	-2.49	100.22	103.27
5	Z	4	A1AQ4	CG-CB-CA	-2.45	109.14	114.10
5	Y	4	A1AQ4	CB-CA-C	-2.10	107.53	111.47

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	77	CSX	N-CA-CB-SG
3	L	77	CSX	N-CA-CB-SG
5	Z	4	A1AQ4	C-CA-CB-CG

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)

11 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	$\#\|Z\| > 2$
6	TFA	K	202	-	6,6,6	0.68	0	9,9,9	0.85	0
8	EDO	C	301	-	3,3,3	0.53	0	2,2,2	0.33	0
9	ACT	C	302	-	3,3,3	0.85	0	3,3,3	0.56	0
10	DMS	L	301	-	3,3,3	0.63	0	3,3,3	0.73	0
10	DMS	H	202	-	3,3,3	0.55	0	3,3,3	0.73	0
7	PEG	B	202	-	6,6,6	0.69	0	5,5,5	0.53	0
10	DMS	I	301	-	3,3,3	0.68	0	3,3,3	0.56	0
6	TFA	B	201	-	6,6,6	0.77	0	9,9,9	1.15	1 (11%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
9	ACT	H	201	-	3,3,3	0.78	0	3,3,3	0.66	0
9	ACT	F	301	-	3,3,3	0.81	0	3,3,3	0.78	0
8	EDO	K	201	-	3,3,3	0.72	0	2,2,2	0.28	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
6	TFA	K	202	-	-	0/6/6/6	-
8	EDO	C	301	-	-	1/1/1/1	-
7	PEG	B	202	-	-	3/4/4/4	-
6	TFA	B	201	-	-	0/6/6/6	-
8	EDO	K	201	-	-	0/1/1/1	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	B	201	TFA	F3-C2-C1	-2.10	104.04	111.85

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	B	202	PEG	O1-C1-C2-O2
7	B	202	PEG	C4-C3-O2-C2
7	B	202	PEG	C1-C2-O2-C3
8	C	301	EDO	O1-C1-C2-O2

There are no ring outliers.

4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	C	301	EDO	3	0
10	L	301	DMS	1	0
7	B	202	PEG	2	0
10	I	301	DMS	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^{th} 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	104/104 (100%)	-0.14	1 (0%) 82 84	42, 58, 88, 107	0
1	D	103/104 (99%)	0.29	3 (2%) 51 54	52, 86, 122, 133	0
1	G	103/104 (99%)	-0.02	0 100 100	50, 72, 92, 102	0
1	J	104/104 (100%)	-0.18	1 (0%) 82 84	40, 61, 97, 110	0
2	B	88/96 (91%)	-0.16	0 100 100	40, 55, 93, 118	0
2	E	87/96 (90%)	0.01	1 (1%) 80 82	44, 72, 98, 117	0
2	H	86/96 (89%)	0.01	1 (1%) 79 80	51, 68, 107, 126	0
2	K	87/96 (90%)	-0.11	2 (2%) 60 62	38, 58, 88, 102	0
3	C	141/176 (80%)	0.08	7 (4%) 28 30	41, 67, 115, 134	0
3	L	145/176 (82%)	-0.06	2 (1%) 75 77	39, 53, 89, 123	0
4	F	144/176 (81%)	0.36	12 (8%) 11 10	45, 78, 121, 148	0
4	I	147/176 (83%)	0.00	4 (2%) 54 56	47, 64, 110, 130	0
5	W	0/5	-	-	-	-
5	X	0/5	-	-	-	-
5	Y	0/5	-	-	-	-
5	Z	0/5	-	-	-	-
All	All	1339/1524 (87%)	0.02	34 (2%) 57 59	38, 66, 107, 148	0

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	F	148	PHE	5.0
3	C	145	GLN	4.5
4	F	205	ARG	4.4
1	D	81	ALA	4.3
4	F	145	GLN	3.9
1	J	81	ALA	3.8
4	F	185	TYR	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	C	205	ARG	3.4
4	F	142	VAL	3.4
3	C	61	PRO	3.3
4	F	143	ASP	3.2
4	I	61	PRO	3.1
4	I	182	ARG	3.1
4	I	201	LEU	2.9
1	D	79	PHE	2.9
4	F	140	LEU	2.8
4	F	204	GLU	2.8
3	C	185	TYR	2.8
4	F	76	PHE	2.7
3	C	204	GLU	2.6
3	L	143	ASP	2.5
4	I	144	GLY	2.5
4	F	202	THR	2.4
2	K	17	MET	2.4
1	A	81	ALA	2.3
2	H	46	LEU	2.3
3	L	144	GLY	2.3
2	E	62	PHE	2.3
3	C	146	PRO	2.2
2	K	47	SER	2.2
1	D	80	ARG	2.2
3	C	147	ILE	2.2
4	F	69	ARG	2.0
4	F	114	GLY	2.0

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^{th} 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^2)	Q<0.9
3	CSX	C	77	7/8	0.88	0.12	63,68,90,92	0
5	A1AQ4	Y	4	17/18	0.89	0.31	76,82,90,90	0
5	HYP	Y	3	8/9	0.90	0.18	83,89,92,95	0
5	TBG	Y	2	8/9	0.90	0.30	87,91,95,95	0
5	TBG	Z	2	8/9	0.92	0.24	54,60,69,70	0
5	HYP	Z	3	8/9	0.93	0.21	62,67,68,75	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	A1AQ4	W	4	17/18	0.95	0.14	42,45,50,52	0
3	CSX	L	77	7/8	0.95	0.11	47,50,69,76	0
5	A1AQ4	X	4	17/18	0.96	0.18	52,57,60,60	0
5	TBG	W	2	8/9	0.96	0.17	48,51,53,54	0
5	A1AQ4	Z	4	17/18	0.96	0.28	62,67,75,75	0
5	TBG	X	2	8/9	0.97	0.15	51,53,56,57	0
5	HYP	W	3	8/9	0.97	0.13	41,46,50,53	0
5	HYP	X	3	8/9	0.98	0.13	43,52,55,57	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^{th} 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^2)	Q<0.9
8	EDO	K	201	4/4	0.65	0.23	72,75,75,78	0
10	DMS	L	301	4/4	0.84	0.21	73,87,89,97	0
9	ACT	F	301	4/4	0.86	0.15	83,86,87,89	0
9	ACT	H	201	4/4	0.88	0.16	68,72,77,78	0
7	PEG	B	202	7/7	0.88	0.23	61,74,76,77	0
10	DMS	I	301	4/4	0.90	0.18	72,89,89,91	0
10	DMS	H	202	4/4	0.91	0.14	88,90,93,102	0
8	EDO	C	301	4/4	0.91	0.17	41,47,54,59	0
6	TFA	K	202	7/7	0.91	0.12	82,85,98,105	0
9	ACT	C	302	4/4	0.94	0.10	74,76,81,88	0
6	TFA	B	201	7/7	0.94	0.08	60,79,84,86	0

6.5 Other polymers (i)

There are no such residues in this entry.