wwPDB X-ray Structure Validation Summary Report (i)

Feb 14, 2017 – 08:27 PM EST

PDB ID	:	5LS6
Title	:	Structure of Human Polycomb Repressive Complex 2 (PRC2) with inhibitor
Authors	:	Zhang, Y.; Justin, N.; Chen, S.; Wilson, J.; Gamblin, S.
Deposited on	:	2016-08-22
Resolution	:	3.47 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

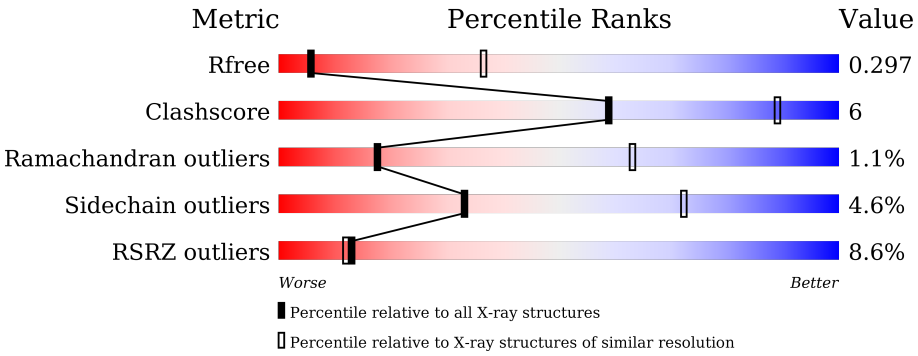
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028442
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20028442

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 3.47 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}	91344	1173 (3.60-3.36)
Clashscore	102246	1010 (3.58-3.38)
Ramachandran outliers	100387	1245 (3.60-3.36)
Sidechain outliers	100360	1246 (3.60-3.36)
RSRZ outliers	91569	1180 (3.60-3.36)

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 on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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	695	11% 69% 12% • 18%
1	D	695	7% 66% 14% • 18%
1	G	695	12% 68% 13% • 18%
1	J	695	13% 70% 11% • 18%
2	B	367	4% 79% 19% ••
2	E	367	% 78% 20% ••

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	367	2% 86% 13%
2	K	367	1% 84% 14%
3	C	129	6% 71% 21% 5%
3	F	129	5% 72% 25%
3	I	129	4% 75% 22%
3	L	129	11% 78% 18%
4	Q	11	36% 45% 9% 9%
4	R	11	9% 45% 36% 9% 9%
4	S	11	9% 45% 45% 9%
4	T	11	18% 45% 27% 18% 9%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	ZN	A	801	-	-	X	-
5	ZN	G	806	-	-	X	-
5	ZN	J	806	-	-	X	-
6	74D	A	809	X	-	-	-
6	74D	D	809	X	-	-	-
6	74D	G	809	X	-	-	-
6	74D	J	809	X	-	-	-

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 34891 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 Histone-lysine N-methyltransferase EZH2,Histone-lysine N-methyltransferase EZH2,Histone-lysine N-methyltransferase EZH2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	570	Total	C	N	O	S	0	0	0
			4588	2878	813	855	42			
1	D	570	Total	C	N	O	S	0	0	0
			4588	2878	813	855	42			
1	G	572	Total	C	N	O	S	0	0	0
			4603	2886	816	859	42			
1	J	570	Total	C	N	O	S	0	0	0
			4588	2878	813	855	42			

There are 52 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-9	GLU	-	expression tag	UNP Q15910
A	-8	THR	-	expression tag	UNP Q15910
A	-7	SER	-	expression tag	UNP Q15910
A	-6	LEU	-	expression tag	UNP Q15910
A	-5	ALA	-	expression tag	UNP Q15910
A	-4	GLU	-	expression tag	UNP Q15910
A	-3	GLU	-	expression tag	UNP Q15910
A	-2	LYS	-	expression tag	UNP Q15910
A	-1	LEU	-	expression tag	UNP Q15910
A	0	THR	-	expression tag	UNP Q15910
A	249	GLN	GLU	conflict	UNP Q15910
A	419	PRO	-	linker	UNP Q15910
A	420	GLY	-	linker	UNP Q15910
D	-9	GLU	-	expression tag	UNP Q15910
D	-8	THR	-	expression tag	UNP Q15910
D	-7	SER	-	expression tag	UNP Q15910
D	-6	LEU	-	expression tag	UNP Q15910
D	-5	ALA	-	expression tag	UNP Q15910
D	-4	GLU	-	expression tag	UNP Q15910
D	-3	GLU	-	expression tag	UNP Q15910

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-2	LYS	-	expression tag	UNP Q15910
D	-1	LEU	-	expression tag	UNP Q15910
D	0	THR	-	expression tag	UNP Q15910
D	249	GLN	GLU	conflict	UNP Q15910
D	419	PRO	-	linker	UNP Q15910
D	420	GLY	-	linker	UNP Q15910
G	-9	GLU	-	expression tag	UNP Q15910
G	-8	THR	-	expression tag	UNP Q15910
G	-7	SER	-	expression tag	UNP Q15910
G	-6	LEU	-	expression tag	UNP Q15910
G	-5	ALA	-	expression tag	UNP Q15910
G	-4	GLU	-	expression tag	UNP Q15910
G	-3	GLU	-	expression tag	UNP Q15910
G	-2	LYS	-	expression tag	UNP Q15910
G	-1	LEU	-	expression tag	UNP Q15910
G	0	THR	-	expression tag	UNP Q15910
G	249	GLN	GLU	conflict	UNP Q15910
G	419	PRO	-	linker	UNP Q15910
G	420	GLY	-	linker	UNP Q15910
J	-9	GLU	-	expression tag	UNP Q15910
J	-8	THR	-	expression tag	UNP Q15910
J	-7	SER	-	expression tag	UNP Q15910
J	-6	LEU	-	expression tag	UNP Q15910
J	-5	ALA	-	expression tag	UNP Q15910
J	-4	GLU	-	expression tag	UNP Q15910
J	-3	GLU	-	expression tag	UNP Q15910
J	-2	LYS	-	expression tag	UNP Q15910
J	-1	LEU	-	expression tag	UNP Q15910
J	0	THR	-	expression tag	UNP Q15910
J	249	GLN	GLU	conflict	UNP Q15910
J	419	PRO	-	linker	UNP Q15910
J	420	GLY	-	linker	UNP Q15910

- Molecule 2 is a protein called Polycomb protein EED.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	365	Total	C	N	O	S	0	0	0
			2959	1873	521	543	22			
2	E	365	Total	C	N	O	S	0	0	0
			2959	1873	521	543	22			
2	H	365	Total	C	N	O	S	0	0	0
			2959	1873	521	543	22			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	K	365	Total	C	N	O	S	0	0	0
			2959	1873	521	543	22			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	75	GLY	-	expression tag	UNP O75530
B	76	SER	-	expression tag	UNP O75530
E	75	GLY	-	expression tag	UNP O75530
E	76	SER	-	expression tag	UNP O75530
H	75	GLY	-	expression tag	UNP O75530
H	76	SER	-	expression tag	UNP O75530
K	75	GLY	-	expression tag	UNP O75530
K	76	SER	-	expression tag	UNP O75530

- Molecule 3 is a protein called Polycomb protein SUZ12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	125	Total	C	N	O	S	0	0	0
			1042	657	180	193	12			
3	F	125	Total	C	N	O	S	0	0	0
			1042	657	180	193	12			
3	I	125	Total	C	N	O	S	0	0	0
			1042	657	180	193	12			
3	L	125	Total	C	N	O	S	0	0	0
			1042	657	180	193	12			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	557	GLY	-	expression tag	UNP Q15022
F	557	GLY	-	expression tag	UNP Q15022
I	557	GLY	-	expression tag	UNP Q15022
L	557	GLY	-	expression tag	UNP Q15022

- Molecule 4 is a protein called Jarid2 K116me3.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	Q	10	Total	C	N	O	0	0	0
			85	54	17	14			
4	R	10	Total	C	N	O	0	0	0
			85	54	17	14			

Continued on next page...

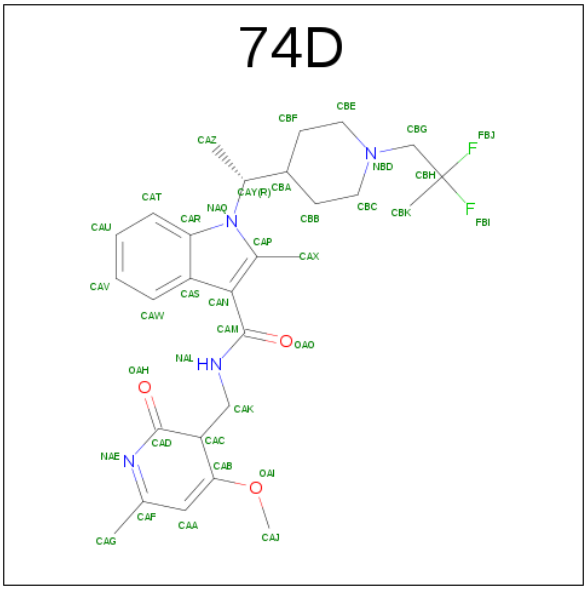
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
4	S	10	Total C N O 85 54 17 14	0	0	0
4	T	10	Total C N O 85 54 17 14	0	0	0

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	G	8	Total Zn 8 8	0	0
5	J	8	Total Zn 8 8	0	0
5	A	8	Total Zn 8 8	0	0
5	D	8	Total Zn 8 8	0	0

- Molecule 6 is 1-[(1 {R})-1-[1-[2,2-bis(fluoranyl)propyl]piperidin-4-yl]ethyl]- {N}-[(4-methoxy-6-methyl-2-oxidanylidene-3 {H}-pyridin-3-yl)methyl]-2-methyl-indole-3-carboxamide (three-letter code: 74D) (formula: $C_{28}H_{36}F_2N_4O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C F N O 37 28 2 4 3	0	0
6	D	1	Total C F N O 37 28 2 4 3	0	0

Continued on next page...

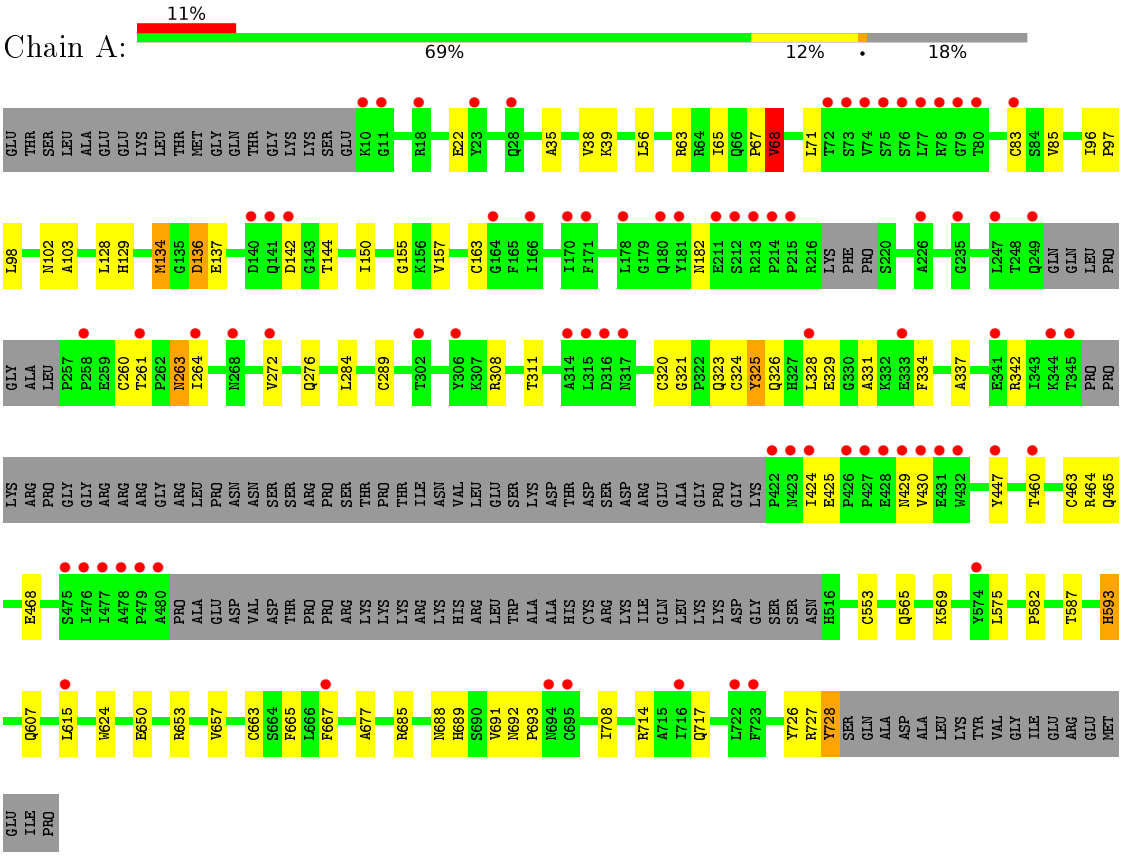
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	G	1	Total	C	F	N	O	0	0
			37	28	2	4	3		
6	J	1	Total	C	F	N	O	0	0
			37	28	2	4	3		

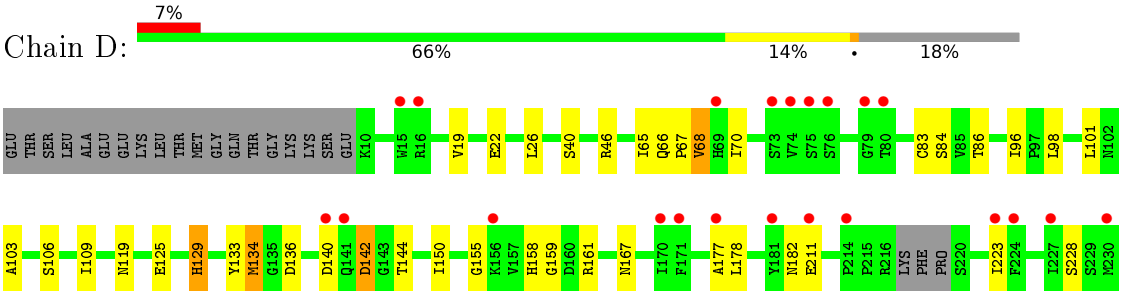
3 Residue-property plots (i)

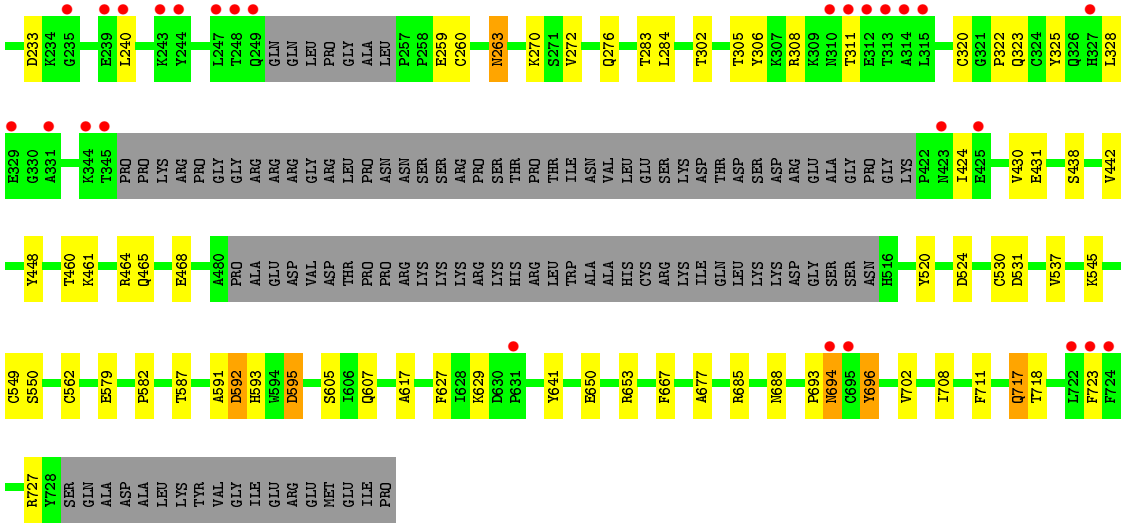
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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: Histone-lysine N-methyltransferase EZH2,Histone-lysine N-methyltransferase EZH2,Histone-lysine N-methyltransferase EZH2



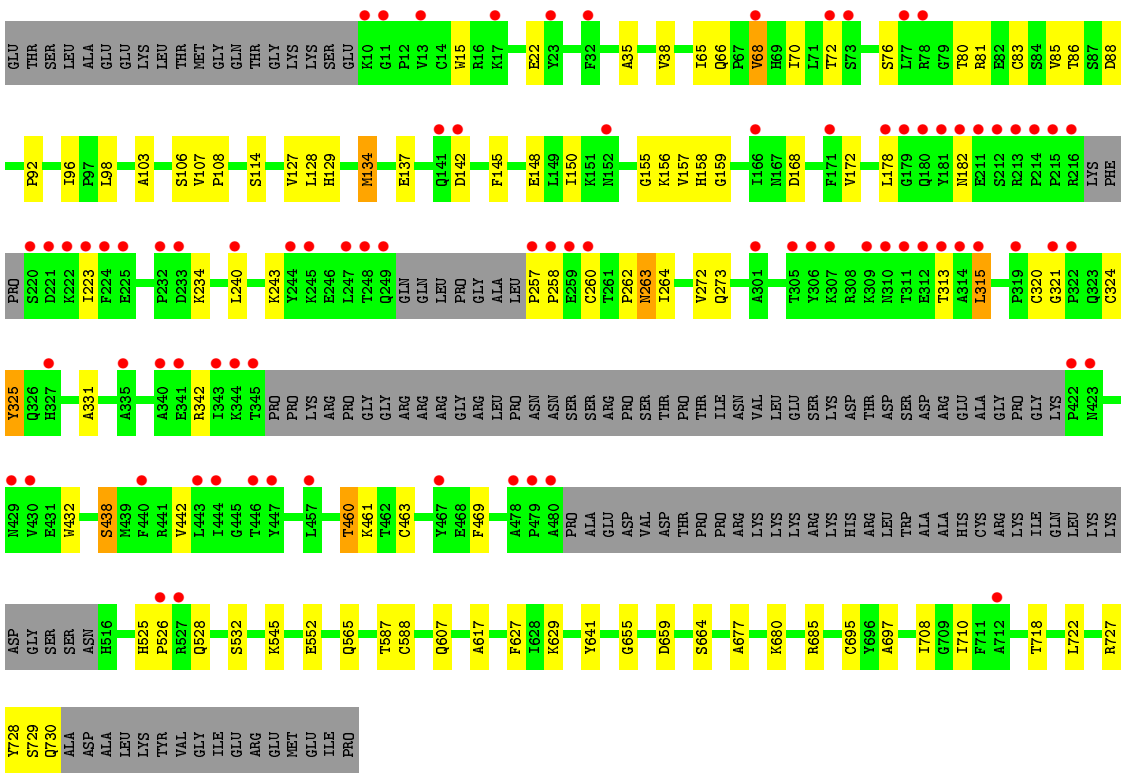
• Molecule 1: Histone-lysine N-methyltransferase EZH2,Histone-lysine N-methyltransferase EZH2,Histone-lysine N-methyltransferase EZH2





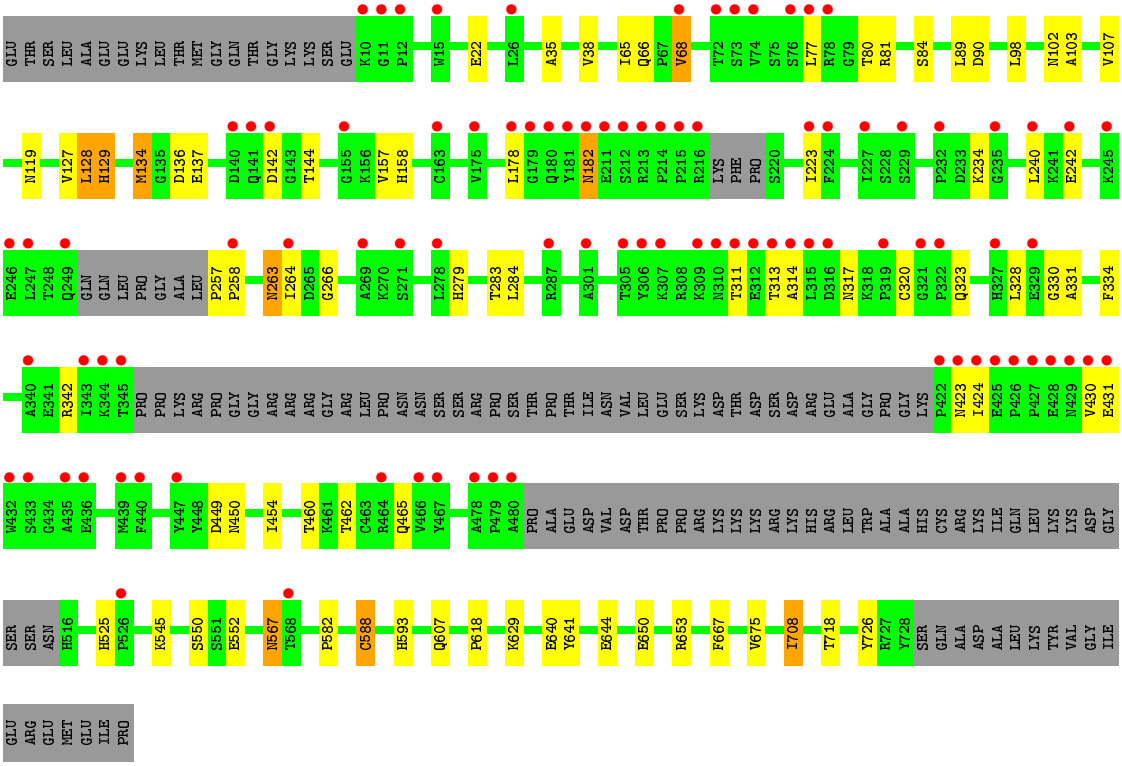
• Molecule 1: Histone-lysine N-methyltransferase EZH2,Histone-lysine N-methyltransferase EZH2,Histone-lysine N-methyltransferase EZH2

Chain G:

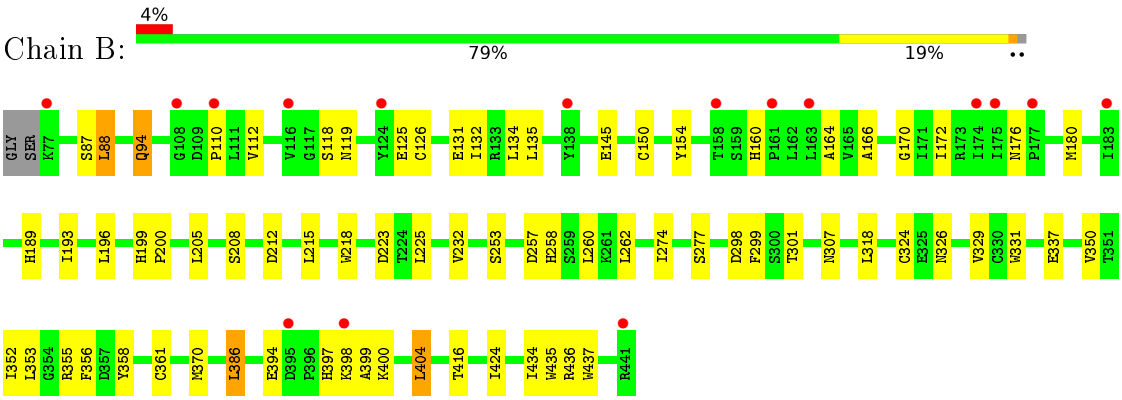


• Molecule 1: Histone-lysine N-methyltransferase EZH2,Histone-lysine N-methyltransferase EZH2,Histone-lysine N-methyltransferase EZH2

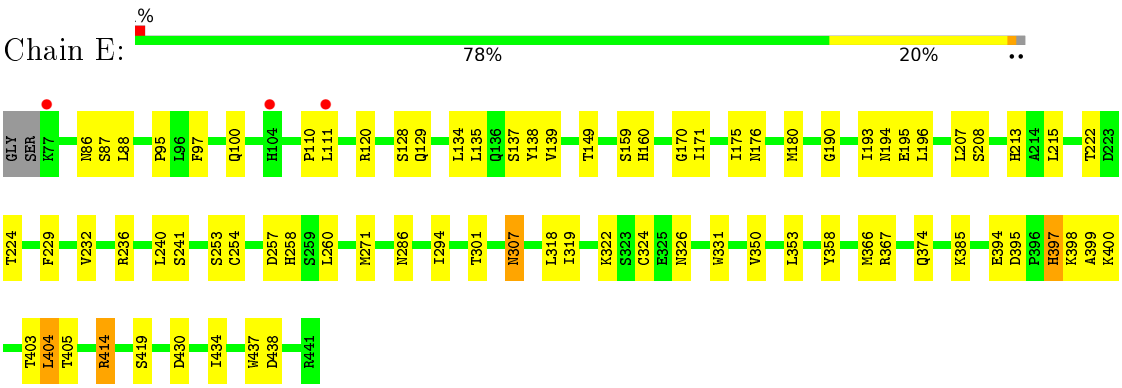
Chain J:



• Molecule 2: Polycomb protein EED

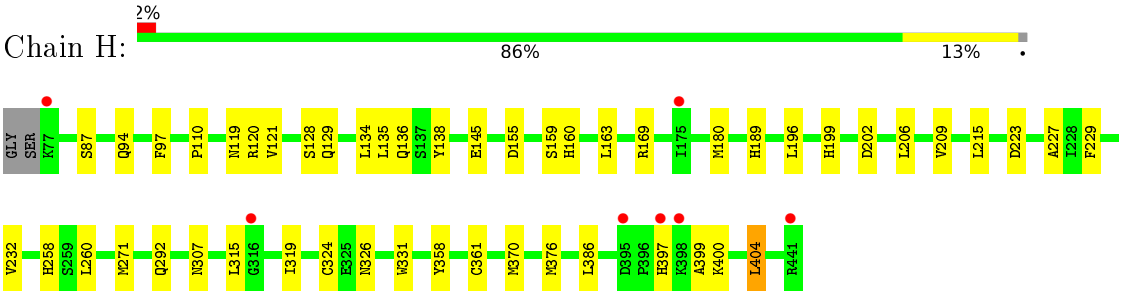


• Molecule 2: Polycomb protein EED



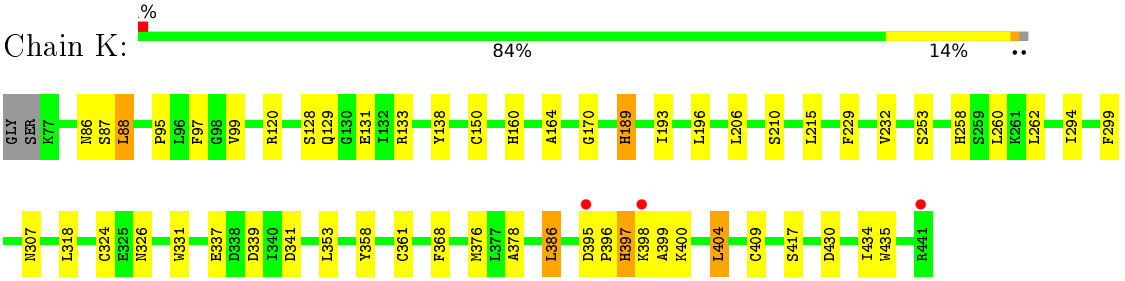
• Molecule 2: Polycomb protein EED

Chain H:



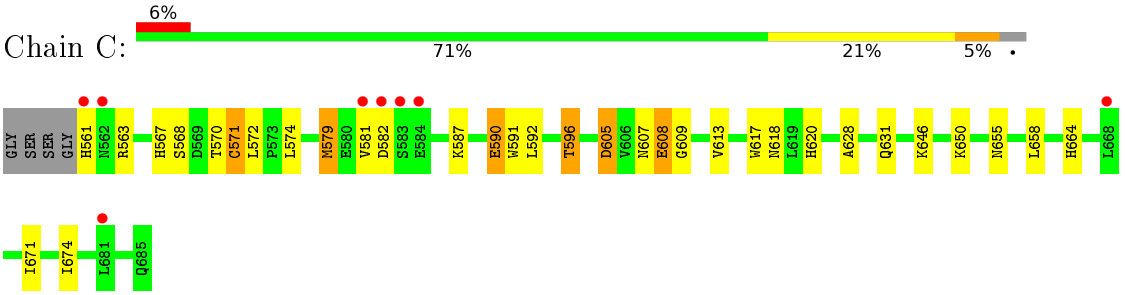
• Molecule 2: Polycomb protein EED

Chain K:



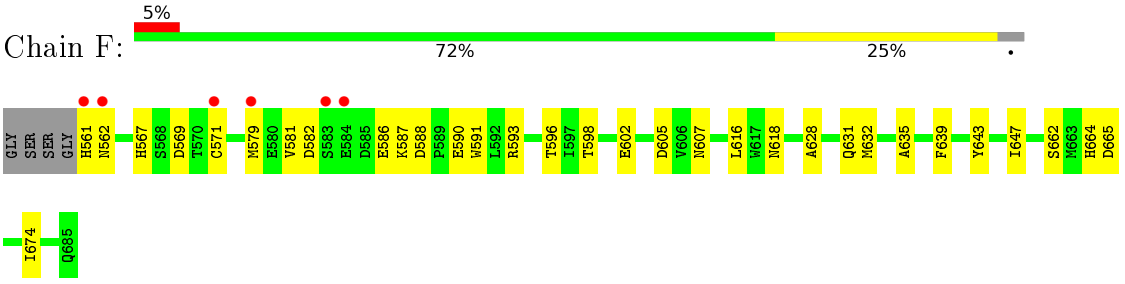
• Molecule 3: Polycomb protein SUZ12

Chain C:



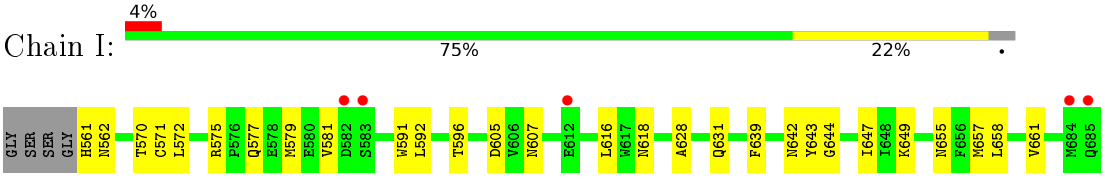
• Molecule 3: Polycomb protein SUZ12

Chain F:



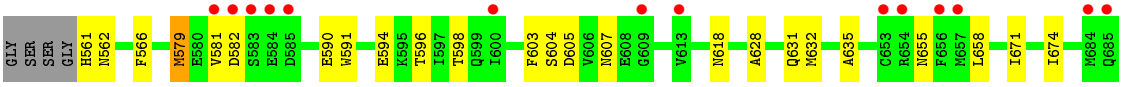
• Molecule 3: Polycomb protein SUZ12

Chain I:



• Molecule 3: Polycomb protein SUZ12

Chain L:



• Molecule 4: Jarid2 K116me3

Chain Q:

• Molecule 4: Jarid2 K116me3

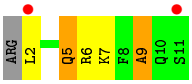
Chain R:

• Molecule 4: Jarid2 K116me3

Chain S:

• Molecule 4: Jarid2 K116me3

Chain T:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	131.31Å 170.27Å 275.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	68.84 – 3.47 68.84 – 3.47	Depositor EDS
% Data completeness (in resolution range)	99.9 (68.84-3.47) 99.9 (68.84-3.47)	Depositor EDS
R_{merge}	0.30	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.92 (at 3.49Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.229 , 0.299 0.231 , 0.297	Depositor DCC
R_{free} test set	3952 reflections (5.16%)	DCC
Wilson B-factor (Å^2)	78.0	Xtriage
Anisotropy	0.722	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 56.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.42$, $< L^2 > = 0.25$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	34891	wwPDB-VP
Average B, all atoms (Å^2)	105.0	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 34.62 % 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 6.6281e-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: ZN, M3L, 74D

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	0.50	0/4691	0.74	0/6327
1	D	0.51	0/4691	0.74	0/6327
1	G	0.51	0/4706	0.76	2/6347 (0.0%)
1	J	0.53	1/4691 (0.0%)	0.75	0/6327
2	B	0.52	0/3034	0.76	0/4107
2	E	0.50	0/3034	0.76	0/4107
2	H	0.52	0/3034	0.78	0/4107
2	K	0.50	0/3034	0.77	0/4107
3	C	0.51	0/1063	0.79	2/1427 (0.1%)
3	F	0.49	0/1063	0.77	0/1427
3	I	0.46	0/1063	0.77	0/1427
3	L	0.50	0/1063	0.77	0/1427
4	Q	0.65	0/73	1.05	0/96
4	R	0.70	0/73	0.73	0/96
4	S	0.66	0/73	1.02	0/96
4	T	0.75	0/73	1.10	0/96
All	All	0.51	1/35459 (0.0%)	0.76	4/47848 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	3
1	D	0	3
1	G	0	2
1	J	0	2
3	C	0	3
3	F	0	1
3	L	0	2

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
All	All	0	16

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	J	182	ASN	C-O	5.38	1.33	1.23

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	G	685	ARG	NE-CZ-NH2	5.97	123.29	120.30
1	G	315	LEU	CA-CB-CG	5.97	129.02	115.30
3	C	563	ARG	NE-CZ-NH1	5.20	122.90	120.30
3	C	608	GLU	N-CA-C	5.11	124.80	111.00

There are no chirality outliers.

5 of 16 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	134	MET	Peptide
1	A	311	THR	Peptide
1	A	326	GLN	Peptide
3	C	561	HIS	Peptide
3	C	579	MET	Peptide

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	4588	0	4445	74	0
1	D	4588	0	4442	55	0
1	G	4603	0	4457	51	0
1	J	4588	0	4447	37	0
2	B	2959	0	2881	49	0
2	E	2959	0	2881	44	0
2	H	2959	0	2881	30	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	K	2959	0	2881	34	0
3	C	1042	0	1021	22	0
3	F	1042	0	1021	20	0
3	I	1042	0	1021	17	0
3	L	1042	0	1021	17	0
4	Q	85	0	90	6	0
4	R	85	0	90	3	0
4	S	85	0	90	5	0
4	T	85	0	90	5	0
5	A	8	0	0	2	0
5	D	8	0	0	0	0
5	G	8	0	0	3	0
5	J	8	0	0	2	0
6	A	37	0	0	4	0
6	D	37	0	0	0	0
6	G	37	0	0	0	0
6	J	37	0	0	0	0
All	All	34891	0	33759	401	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 6.

The worst 5 of 401 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:A:667:PHE:CD2	1:A:708:ILE:HD12	1.39	1.55
1:A:667:PHE:CD2	1:A:708:ILE:CD1	2.00	1.41
1:A:667:PHE:CE2	1:A:708:ILE:CD1	2.07	1.36
1:A:667:PHE:CE2	1:A:708:ILE:HD13	1.72	1.22
1:A:667:PHE:CE2	1:A:708:ILE:HD12	1.76	1.12

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	558/695 (80%)	501 (90%)	51 (9%)	6 (1%)	17	63
1	D	558/695 (80%)	504 (90%)	44 (8%)	10 (2%)	11	52
1	G	560/695 (81%)	513 (92%)	41 (7%)	6 (1%)	17	63
1	J	558/695 (80%)	501 (90%)	50 (9%)	7 (1%)	15	59
2	B	363/367 (99%)	330 (91%)	31 (8%)	2 (1%)	30	74
2	E	363/367 (99%)	331 (91%)	27 (7%)	5 (1%)	14	57
2	H	363/367 (99%)	346 (95%)	16 (4%)	1 (0%)	46	83
2	K	363/367 (99%)	334 (92%)	27 (7%)	2 (1%)	30	74
3	C	123/129 (95%)	113 (92%)	8 (6%)	2 (2%)	12	54
3	F	123/129 (95%)	112 (91%)	10 (8%)	1 (1%)	24	69
3	I	123/129 (95%)	118 (96%)	4 (3%)	1 (1%)	24	69
3	L	123/129 (95%)	113 (92%)	10 (8%)	0	100	100
4	Q	7/11 (64%)	6 (86%)	0	1 (14%)	0	4
4	R	7/11 (64%)	7 (100%)	0	0	100	100
4	S	7/11 (64%)	7 (100%)	0	0	100	100
4	T	7/11 (64%)	4 (57%)	2 (29%)	1 (14%)	0	4
All	All	4206/4808 (88%)	3840 (91%)	321 (8%)	45 (1%)	17	63

5 of 45 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	68	VAL
1	A	103	ALA
2	B	257	ASP
1	D	103	ALA
1	J	103	ALA

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	511/618 (83%)	486 (95%)	25 (5%)	31 70
1	D	511/618 (83%)	482 (94%)	29 (6%)	25 66
1	G	513/618 (83%)	482 (94%)	31 (6%)	24 64
1	J	511/618 (83%)	482 (94%)	29 (6%)	25 66
2	B	328/329 (100%)	316 (96%)	12 (4%)	41 76
2	E	328/329 (100%)	316 (96%)	12 (4%)	41 76
2	H	328/329 (100%)	318 (97%)	10 (3%)	48 81
2	K	328/329 (100%)	321 (98%)	7 (2%)	61 86
3	C	119/121 (98%)	113 (95%)	6 (5%)	30 69
3	F	119/121 (98%)	116 (98%)	3 (2%)	55 84
3	I	119/121 (98%)	115 (97%)	4 (3%)	44 79
3	L	119/121 (98%)	118 (99%)	1 (1%)	86 95
4	Q	7/8 (88%)	5 (71%)	2 (29%)	0 3
4	R	7/8 (88%)	5 (71%)	2 (29%)	0 3
4	S	7/8 (88%)	6 (86%)	1 (14%)	4 23
4	T	7/8 (88%)	5 (71%)	2 (29%)	0 3
All	All	3862/4304 (90%)	3686 (95%)	176 (5%)	33 72

5 of 176 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	367	ARG
1	G	134	MET
2	K	160	HIS
2	E	403	THR
1	G	76	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 63 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	655	ASN
2	H	176	ASN
4	Q	10	GLN
1	G	152	ASN
1	G	607	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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	M3L	Q	7	4	9,11,12	0.66	0	12,14,16	0.70	0
4	M3L	R	7	4	9,11,12	0.77	0	12,14,16	0.74	0
4	M3L	S	7	4	9,11,12	0.65	0	12,14,16	0.82	1 (8%)
4	M3L	T	7	4	9,11,12	0.57	0	12,14,16	0.84	1 (8%)

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	M3L	Q	7	4	-	0/8/10/12	0/0/0/0
4	M3L	R	7	4	-	0/8/10/12	0/0/0/0
4	M3L	S	7	4	-	0/8/10/12	0/0/0/0
4	M3L	T	7	4	-	0/8/10/12	0/0/0/0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	T	7	M3L	O-C-CA	-2.41	119.26	125.72
4	S	7	M3L	O-C-CA	-2.18	119.87	125.72

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	Q	7	M3L	1	0
4	R	7	M3L	1	0
4	S	7	M3L	1	0
4	T	7	M3L	2	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 36 ligands modelled in this entry, 32 are monoatomic - leaving 4 for Mogul analysis.

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	74D	A	809	-	32,40,40	2.93	9 (28%)	44,59,59	2.67	9 (20%)
6	74D	D	809	-	32,40,40	2.58	8 (25%)	44,59,59	3.09	14 (31%)
6	74D	G	809	-	32,40,40	2.64	7 (21%)	44,59,59	3.40	18 (40%)
6	74D	J	809	-	32,40,40	2.82	8 (25%)	44,59,59	3.19	16 (36%)

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	74D	A	809	-	1/1/9/11	0/17/50/50	0/3/4/4
6	74D	D	809	-	1/1/9/11	0/17/50/50	0/3/4/4
6	74D	G	809	-	1/1/9/11	0/17/50/50	0/3/4/4
6	74D	J	809	-	1/1/9/11	0/17/50/50	0/3/4/4

The worst 5 of 32 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	809	74D	CAN-CAM	-11.96	1.33	1.51
6	J	809	74D	CAN-CAM	-10.75	1.34	1.51
6	D	809	74D	CAN-CAM	-9.73	1.36	1.51
6	G	809	74D	CAN-CAM	-8.99	1.37	1.51
6	D	809	74D	CAG-CAF	-6.29	1.38	1.50

The worst 5 of 57 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	J	809	74D	OAI-CAB-CAA	-8.72	119.94	126.00
6	D	809	74D	CAC-CAK-NAL	-6.64	101.57	112.03
6	G	809	74D	OAI-CAB-CAA	-6.10	121.76	126.00
6	A	809	74D	OAI-CAB-CAA	-5.71	122.03	126.00
6	J	809	74D	CBA-CAY-NAQ	-5.62	104.12	112.03

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	G	809	74D	CAC
6	D	809	74D	CAC
6	J	809	74D	CAC
6	A	809	74D	CAC

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	809	74D	4	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	G	1
1	J	1
1	A	1
1	D	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	182:ASN	C	211:GLU	N	6.31
1	G	182:ASN	C	211:GLU	N	5.74
1	D	182:ASN	C	211:GLU	N	5.71
1	J	182:ASN	C	211:GLU	N	5.36

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	570/695 (82%)	0.86	76 (13%) 4 5	51, 108, 183, 239	0
1	D	570/695 (82%)	0.55	49 (8%) 13 12	56, 100, 177, 259	0
1	G	572/695 (82%)	0.81	83 (14%) 3 3	49, 107, 192, 252	0
1	J	570/695 (82%)	0.95	93 (16%) 2 2	52, 108, 203, 254	0
2	B	365/367 (99%)	0.33	16 (4%) 38 31	54, 94, 139, 213	0
2	E	365/367 (99%)	0.24	3 (0%) 87 81	57, 93, 137, 198	0
2	H	365/367 (99%)	0.17	7 (1%) 70 62	49, 80, 120, 216	0
2	K	365/367 (99%)	0.13	3 (0%) 87 81	46, 74, 118, 208	0
3	C	125/129 (96%)	0.55	8 (6%) 23 19	67, 100, 168, 233	0
3	F	125/129 (96%)	0.36	6 (4%) 34 28	58, 92, 161, 219	0
3	I	125/129 (96%)	0.43	5 (4%) 42 34	73, 104, 163, 202	0
3	L	125/129 (96%)	0.67	14 (11%) 7 7	69, 113, 163, 229	0
4	Q	9/11 (81%)	0.27	0 100 100	77, 101, 126, 131	0
4	R	9/11 (81%)	0.60	1 (11%) 7 7	77, 95, 135, 138	0
4	S	9/11 (81%)	0.81	1 (11%) 7 7	93, 119, 137, 139	0
4	T	9/11 (81%)	1.21	2 (22%) 1 1	101, 117, 159, 165	0
All	All	4278/4808 (88%)	0.56	367 (8%) 13 12	46, 95, 181, 259	0

The worst 5 of 367 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	J	479	PRO	10.3
1	J	429	ASN	9.6
1	J	480	ALA	9.6
1	G	480	ALA	8.8
1	A	75	SER	8.7

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
4	M3L	R	7	12/13	0.96	0.25	-	50,57,84,84	0
4	M3L	T	7	12/13	0.96	0.26	-	67,74,83,87	0
4	M3L	Q	7	12/13	0.97	0.23	-	55,63,83,84	0
4	M3L	S	7	12/13	0.96	0.27	-	60,71,87,89	0

6.3 Carbohydrates (i)

There are no carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
6	74D	G	809	37/37	0.96	0.31	1.19	61,69,76,77	0
6	74D	A	809	37/37	0.94	0.29	0.79	62,78,86,91	0
6	74D	D	809	37/37	0.95	0.30	0.68	61,70,89,93	0
5	ZN	J	807	1/1	0.98	0.18	0.57	71,71,71,71	0
6	74D	J	809	37/37	0.95	0.27	0.41	65,77,87,89	0
5	ZN	G	808	1/1	1.00	0.17	-0.05	49,49,49,49	0
5	ZN	D	807	1/1	0.99	0.17	-0.15	66,66,66,66	0
5	ZN	D	808	1/1	0.99	0.17	-0.19	56,56,56,56	0
5	ZN	G	807	1/1	0.99	0.17	-0.35	59,59,59,59	0
5	ZN	J	804	1/1	0.99	0.13	-0.51	94,94,94,94	0
5	ZN	J	808	1/1	1.00	0.16	-0.55	49,49,49,49	0
5	ZN	G	806	1/1	0.99	0.16	-0.57	61,61,61,61	0
5	ZN	A	805	1/1	0.99	0.15	-0.60	66,66,66,66	0
5	ZN	G	804	1/1	0.99	0.14	-0.62	70,70,70,70	0
5	ZN	D	806	1/1	0.99	0.15	-0.63	58,58,58,58	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
5	ZN	D	801	1/1	0.99	0.15	-0.70	68,68,68,68	0
5	ZN	D	804	1/1	1.00	0.13	-0.71	61,61,61,61	0
5	ZN	J	801	1/1	0.99	0.15	-0.76	68,68,68,68	0
5	ZN	J	806	1/1	0.99	0.17	-0.77	75,75,75,75	0
5	ZN	A	808	1/1	1.00	0.16	-0.83	55,55,55,55	0
5	ZN	A	804	1/1	0.99	0.14	-0.84	56,56,56,56	0
5	ZN	G	803	1/1	0.98	0.13	-0.85	73,73,73,73	0
5	ZN	G	801	1/1	0.99	0.14	-0.86	71,71,71,71	0
5	ZN	J	803	1/1	0.99	0.12	-0.87	89,89,89,89	0
5	ZN	A	802	1/1	0.96	0.11	-0.88	132,132,132,132	0
5	ZN	D	803	1/1	1.00	0.13	-0.93	59,59,59,59	0
5	ZN	D	805	1/1	0.99	0.13	-0.94	65,65,65,65	0
5	ZN	A	803	1/1	1.00	0.13	-1.05	65,65,65,65	0
5	ZN	J	805	1/1	0.99	0.11	-1.06	80,80,80,80	0
5	ZN	G	805	1/1	0.99	0.13	-1.14	72,72,72,72	0
5	ZN	A	801	1/1	0.99	0.13	-1.20	70,70,70,70	0
5	ZN	D	802	1/1	0.99	0.11	-1.27	84,84,84,84	0
5	ZN	J	802	1/1	0.94	0.10	-1.48	147,147,147,147	0
5	ZN	A	806	1/1	0.99	0.13	-1.55	53,53,53,53	0
5	ZN	A	807	1/1	0.99	0.15	-1.59	55,55,55,55	0
5	ZN	G	802	1/1	0.98	0.12	-2.24	146,146,146,146	0

6.5 Other polymers (i)

There are no such residues in this entry.