wwPDB EM Map/Model Validation Report (i)

Aug 8, 2016 – 02:49 PM EDT

PDB ID	:	5KGF
EMDB ID:	:	EMD-8246
Title	:	Structural model of 53BP1 bound to a ubiquitylated and methylated nucleosome, at 4.5 A resolution
Authors	:	Wilson, M.D.; Benlekbir, S.; Sicheri, F.; Rubinstein, J.L.; Durocher, D.
Deposited on	:	2016-06-13
Resolution	:	4.54 Å(reported)

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could
stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp

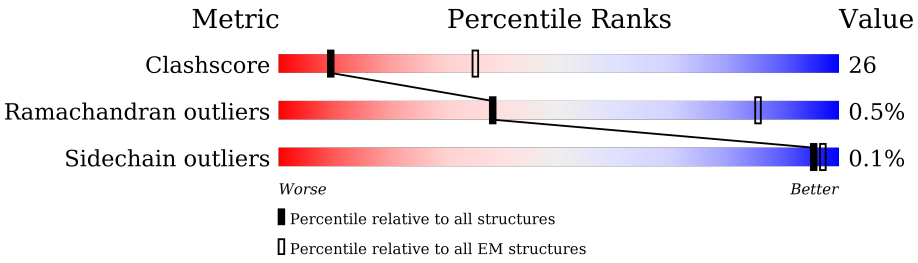
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027939

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 4.54 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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%

Mol	Chain	Length	Quality of chain
1	A	136	43% 32% 26%
1	E	136	43% 30% • 26%
2	B	103	45% 36% • 18%
2	F	103	49% 33% 18%
3	C	130	46% 40% • 13%
3	G	130	45% 39% • 13%
4	D	126	43% 33% • 21%
4	H	126	49% 26% • 23%
5	I	145	42% 41% 17%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	J	145	39% 49% 12%
7	K	21	43% 57%
7	L	21	48% 52%
8	M	76	34% 66%
8	O	76	36% 64%

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 13782 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Histone H3.2.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	101	Total	C	N	O	S	0	0
			832	525	161	143	3		
1	E	100	Total	C	N	O	S	0	0
			825	520	160	142	3		

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	84	Total	C	N	O	S	0	0
			675	425	133	115	2		
2	F	84	Total	C	N	O	S	0	0
			675	425	133	115	2		

- Molecule 3 is a protein called Histone H2A type 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
3	C	113	Total	C	N	O	0	0
			874	549	174	151		
3	G	113	Total	C	N	O	0	0
			874	549	174	151		

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	13	ARG	LYS	engineered mutation	UNP P0C0S8
C	16	SER	THR	engineered mutation	UNP P0C0S8
C	36	ARG	LYS	engineered mutation	UNP P0C0S8
G	13	ARG	LYS	engineered mutation	UNP P0C0S8
G	16	SER	THR	engineered mutation	UNP P0C0S8
G	36	ARG	LYS	engineered mutation	UNP P0C0S8

- Molecule 4 is a protein called Histone H2B type 1-C/E/F/G/I.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	100	Total	C	N	O	S	0	0
			788	493	147	146	2		
4	H	97	Total	C	N	O	S	0	0
			766	479	142	143	2		

- Molecule 5 is a DNA chain called DNA (145-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
5	I	145	Total	C	N	O	P	0	0
			2952	1404	537	867	144		

- Molecule 6 is a DNA chain called DNA (145-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
6	J	145	Total	C	N	O	P	0	0
			2987	1416	558	869	144		

- Molecule 7 is a protein called Tumor suppressor p53-binding protein 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
7	L	21	Total	C	N	O	0	0
			166	100	34	32		
7	K	21	Total	C	N	O	0	0
			166	100	34	32		

- Molecule 8 is a protein called Ubiquitin.

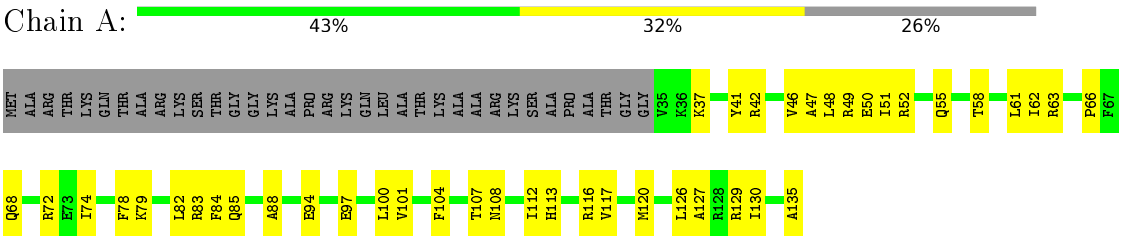
Mol	Chain	Residues	Atoms					AltConf	Trace
8	O	76	Total	C	N	O	S	0	0
			601	378	105	117	1		
8	M	76	Total	C	N	O	S	0	0
			601	378	105	117	1		

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. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

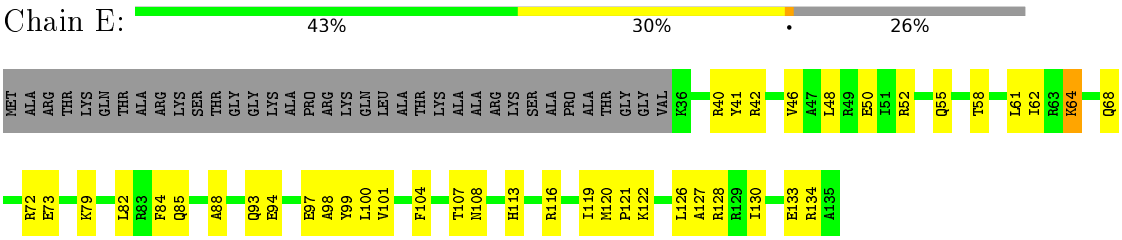
- Molecule 1: Histone H3.2

Chain A:



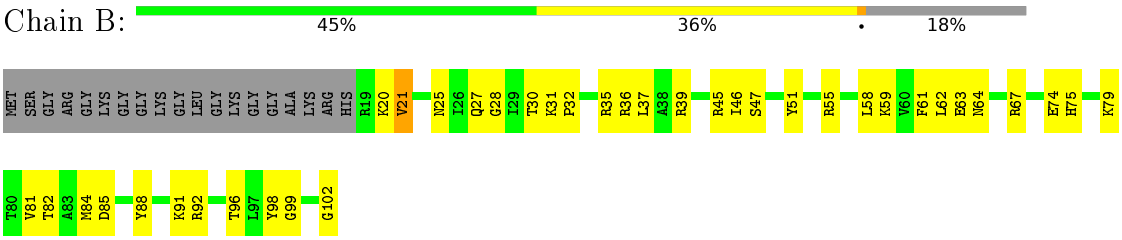
- Molecule 1: Histone H3.2

Chain E:



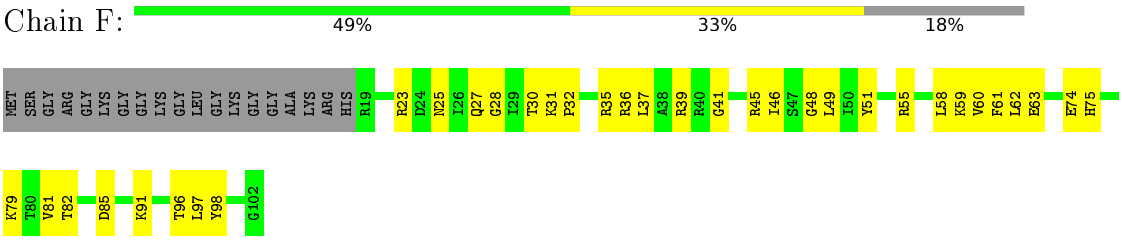
- Molecule 2: Histone H4

Chain B:



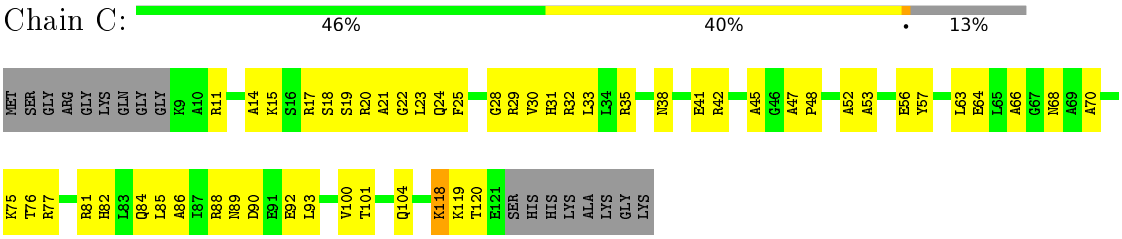
- Molecule 2: Histone H4

Chain F:



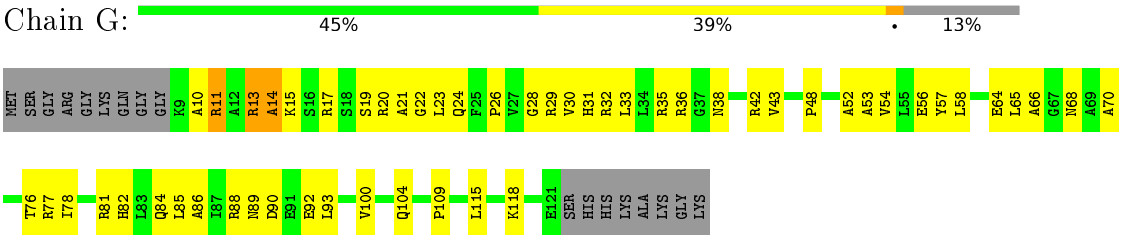
- Molecule 3: Histone H2A type 1

Chain C:



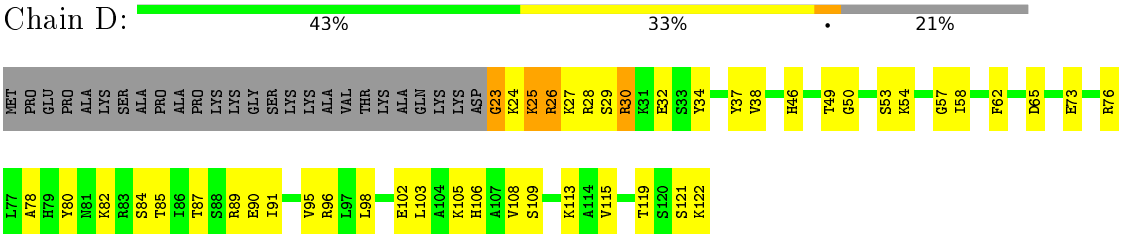
• Molecule 3: Histone H2A type 1

Chain G:



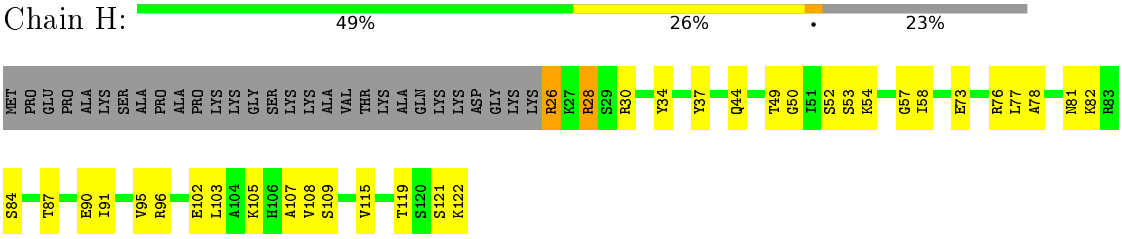
• Molecule 4: Histone H2B type 1-C/E/F/G/I

Chain D:



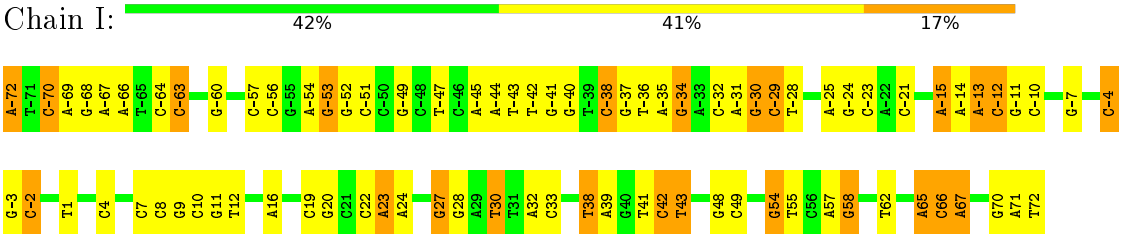
• Molecule 4: Histone H2B type 1-C/E/F/G/I

Chain H:



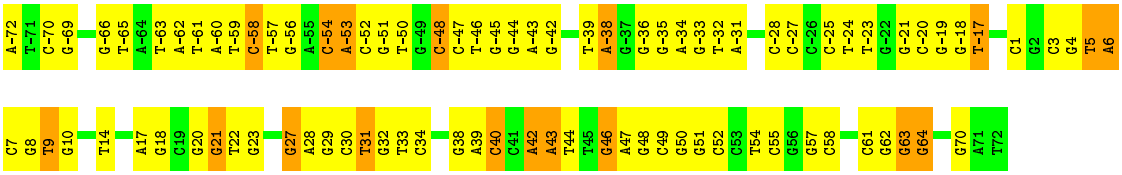
• Molecule 5: DNA (145-MER)

Chain I:



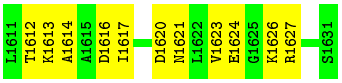
• Molecule 6: DNA (145-MER)

Chain J:



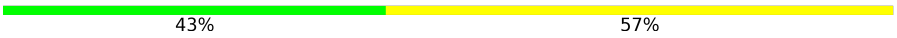
• Molecule 7: Tumor suppressor p53-binding protein 1

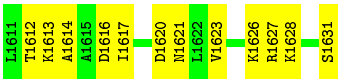
Chain L:



• Molecule 7: Tumor suppressor p53-binding protein 1

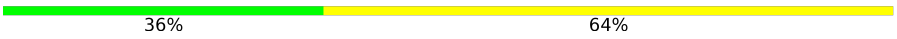
Chain K:

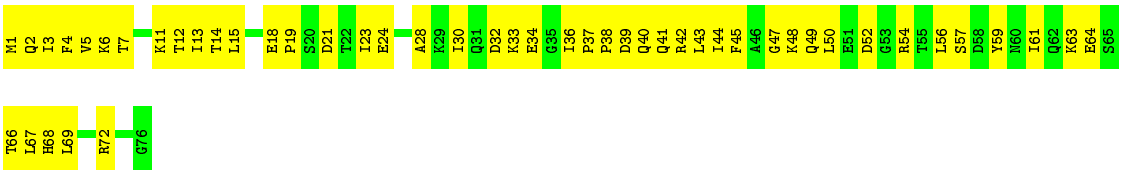




• Molecule 8: Ubiquitin

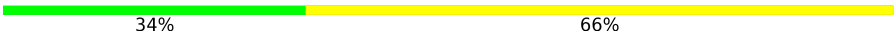
Chain O:

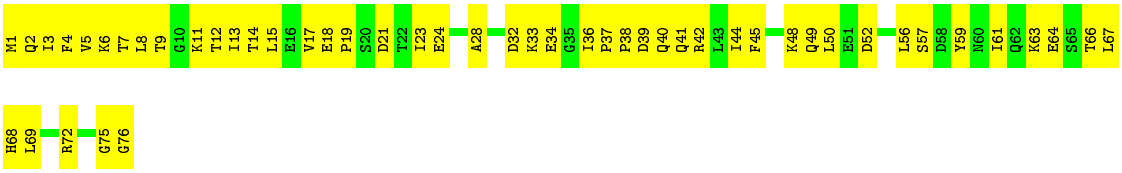




• Molecule 8: Ubiquitin

Chain M:





4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	45361	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	25000	Depositor
Image detector	Not provided	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: M2L

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	#$\|Z\|$ >2	RMSZ	#$\|Z\|$ >2
1	A	0.46	0/844	0.65	0/1130
1	E	0.46	0/837	0.62	0/1120
2	B	0.48	0/670	0.63	0/894
2	F	0.48	0/670	0.58	0/894
3	C	0.45	0/884	0.74	0/1190
3	G	0.44	0/884	0.69	0/1190
4	D	0.48	0/799	0.62	0/1067
4	H	0.48	0/777	0.56	0/1040
5	I	0.74	1/3308 (0.0%)	1.43	43/5099 (0.8%)
6	J	0.73	0/3354	1.45	48/5180 (0.9%)
7	K	0.33	0/165	0.98	1/216 (0.5%)
7	L	0.35	0/165	0.95	1/216 (0.5%)
8	M	0.32	0/607	0.54	0/816
8	O	0.32	0/607	0.54	0/816
All	All	0.60	1/14571 (0.0%)	1.11	93/20868 (0.4%)

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
2	F	0	1
3	C	0	1
3	G	0	3
4	D	0	3
4	H	0	2
7	K	0	1
7	L	0	1
All	All	0	12

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	I	-72	DA	O5'-C5'	9.26	1.65	1.42

The worst 5 of 93 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	J	20	DG	O3'-P-O5'	-11.97	81.26	104.00
6	J	20	DG	OP2-P-O3'	-10.03	83.14	105.20
6	J	14	DT	O4'-C1'-N1	8.79	114.15	108.00
6	J	21	DG	O4'-C1'-N9	8.60	114.02	108.00
6	J	34	DC	P-O3'-C3'	8.10	129.41	119.70

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	118	LYS	Peptide
4	D	23	GLY	Peptide
4	D	25	LYS	Peptide
4	D	26	ARG	Mainchain
2	F	23	ARG	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	832	0	878	67	0
1	E	825	0	867	46	0
2	B	675	0	722	58	0
2	F	675	0	722	37	0
3	C	874	0	939	100	0
3	G	874	0	939	95	0
4	D	788	0	824	84	0
4	H	766	0	795	62	0
5	I	2952	0	1629	119	0
6	J	2987	0	1630	176	0
7	K	166	0	180	26	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	L	166	0	180	14	0
8	M	601	0	629	49	0
8	O	601	0	629	51	0
All	All	13782	0	11563	665	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 26.

The worst 5 of 665 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)
5:I:-72:DA:O5'	5:I:-72:DA:C5'	1.65	1.41
4:D:30:ARG:CB	6:J:49:DC:H4'	1.46	1.41
4:D:30:ARG:HB2	6:J:49:DC:C4'	1.54	1.36
1:E:41:TYR:OH	5:I:-66:DA:H5'	1.19	1.33
4:D:29:SER:CB	4:D:30:ARG:HH21	1.43	1.31

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 PDB entries followed by that with respect to all EM entries.

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	99/136 (73%)	91 (92%)	8 (8%)	0	100	100
1	E	98/136 (72%)	92 (94%)	5 (5%)	1 (1%)	19	65
2	B	81/103 (79%)	72 (89%)	8 (10%)	1 (1%)	16	62
2	F	81/103 (79%)	74 (91%)	7 (9%)	0	100	100
3	C	111/130 (85%)	100 (90%)	10 (9%)	1 (1%)	21	67
3	G	111/130 (85%)	104 (94%)	5 (4%)	2 (2%)	11	55
4	D	98/126 (78%)	93 (95%)	5 (5%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
4	H	95/126 (75%)	91 (96%)	4 (4%)	0	100 100
7	K	19/21 (90%)	15 (79%)	4 (21%)	0	100 100
7	L	19/21 (90%)	15 (79%)	4 (21%)	0	100 100
8	M	74/76 (97%)	73 (99%)	1 (1%)	0	100 100
8	O	74/76 (97%)	73 (99%)	1 (1%)	0	100 100
All	All	960/1184 (81%)	893 (93%)	62 (6%)	5 (0%)	38 77

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	G	14	ALA
3	C	14	ALA
3	G	13	ARG
2	B	21	VAL
1	E	64	LYS

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 PDB entries followed by that with respect to all EM entries.

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	88/111 (79%)	88 (100%)	0	100 100
1	E	87/111 (78%)	87 (100%)	0	100 100
2	B	68/78 (87%)	68 (100%)	0	100 100
2	F	68/78 (87%)	68 (100%)	0	100 100
3	C	88/99 (89%)	88 (100%)	0	100 100
3	G	88/99 (89%)	88 (100%)	0	100 100
4	D	86/106 (81%)	85 (99%)	1 (1%)	78 90
4	H	84/106 (79%)	84 (100%)	0	100 100
7	K	18/18 (100%)	18 (100%)	0	100 100
7	L	18/18 (100%)	18 (100%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
8	M	68/68 (100%)	68 (100%)	0	100 100
8	O	68/68 (100%)	68 (100%)	0	100 100
All	All	829/960 (86%)	828 (100%)	1 (0%)	95 97

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	30	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	113	HIS
2	F	75	HIS
3	G	84	GLN
1	E	68	GLN
3	G	31	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)

2 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	$\#\|Z\| > 2$
2	M2L	B	20	2	8,10,11	5.78	2 (25%)	8,11,13	2.25	4 (50%)
2	M2L	F	20	2	8,10,11	5.75	2 (25%)	8,11,13	2.22	4 (50%)

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
2	M2L	B	20	2	-	0/7/9/11	0/0/0/0
2	M2L	F	20	2	-	0/7/9/11	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	20	M2L	CB-SG	-14.58	1.52	1.81
2	F	20	M2L	CB-SG	-14.48	1.53	1.81
2	F	20	M2L	CD-SG	-7.27	1.52	1.81
2	B	20	M2L	CD-SG	-7.24	1.53	1.81

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	20	M2L	CM1-NZ-CE	2.30	119.92	110.77
2	F	20	M2L	CM2-NZ-CE	2.31	119.97	110.77
2	B	20	M2L	CM1-NZ-CE	2.32	119.98	110.77
2	B	20	M2L	CB-SG-CD	2.32	109.47	102.42
2	B	20	M2L	CM2-NZ-CE	2.33	120.04	110.77

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	20	M2L	1	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.