wwPDB EM Validation Summary Report

Oct 2, 2024 – 12:09 PM JST

PDB ID	:	8HXZ
EMDB ID	:	EMD-35083
Title	:	Cryo-EM structure of Eaf3 CHD in complex with nucleosome
Authors	:	Cui, H.; Wang, H.
Deposited on	:	2023-01-05
Resolution	:	3.40 Å(reported)

This is a wwPDB EM Validation Summary 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/EMValidationReportHelp
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:

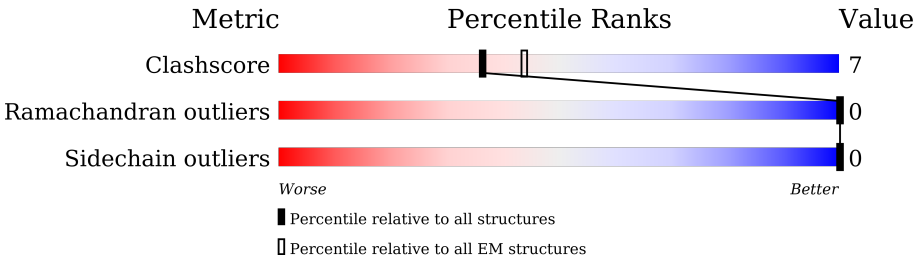
EMDB validation analysis	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance

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

The reported resolution of this entry is 3.40 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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%

Mol	Chain	Length	Quality of chain
1	A	135	56% 19% 24%
1	E	135	54% 18% • 27%
2	B	102	61% 20% 20%
2	F	102	58% 21% 22%
3	C	129	65% 19% 16%
3	G	129	60% 22% 18%
4	D	122	57% 21% 21%
4	H	122	60% 17% • 22%
5	I	352	39% 8% 53%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	J	352	
7	M	401	

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 13740 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.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	102	Total 837	C 529	N 162	O 143	S 3	0	0
1	E	98	Total 810	C 512	N 157	O 139	S 2	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	110	ALA	CYS	engineered mutation	UNP A0A310TTQ1
E	110	ALA	CYS	engineered mutation	UNP A0A310TTQ1

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	82	Total 653	C 412	N 127	O 113	S 1	0	0
2	F	80	Total 638	C 401	N 125	O 111	S 1	0	0

- Molecule 3 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms				AltConf	Trace
3	C	109	Total 843	C 531	N 167	O 145	0	0
3	G	106	Total 818	C 516	N 160	O 142	0	0

- Molecule 4 is a protein called Histone H2B.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	96	Total 757	C 475	N 140	O 140	S 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
4	H	95	Total	C	N	O	S	0	0
			745	469	134	140	2		

- Molecule 5 is a DNA chain called DNA (352-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
5	I	164	Total	C	N	O	P	0	0
			3345	1591	602	988	164		

- Molecule 6 is a DNA chain called DNA (352-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
6	J	164	Total	C	N	O	P	0	0
			3379	1601	634	980	164		

- Molecule 7 is a protein called Chromatin modification-related protein EAF3.

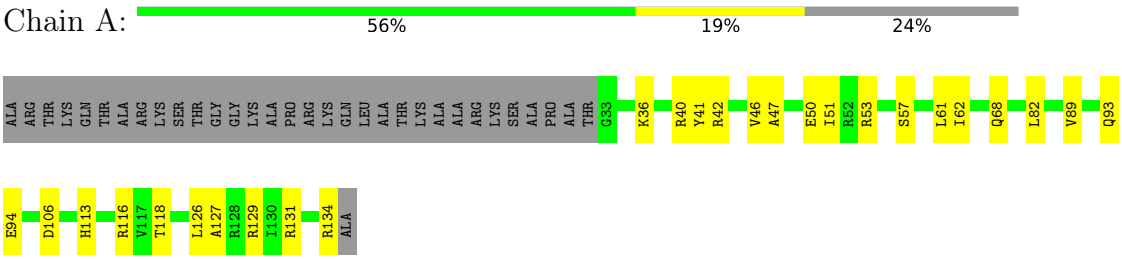
Mol	Chain	Residues	Atoms					AltConf	Trace
7	M	111	Total	C	N	O	S	0	0
			915	591	155	164	5		

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. 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.

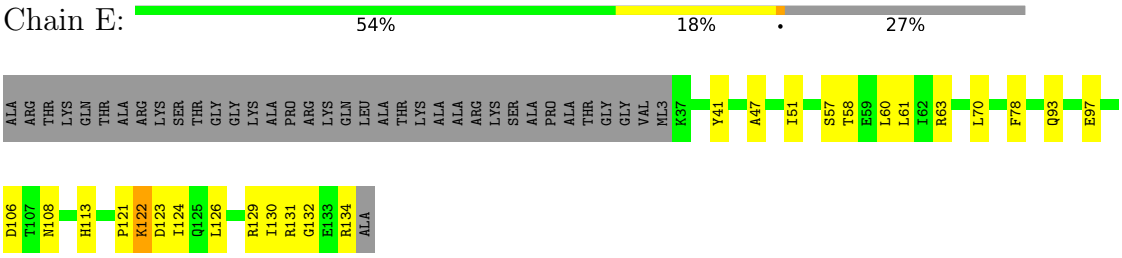
• Molecule 1: Histone H3

Chain A:



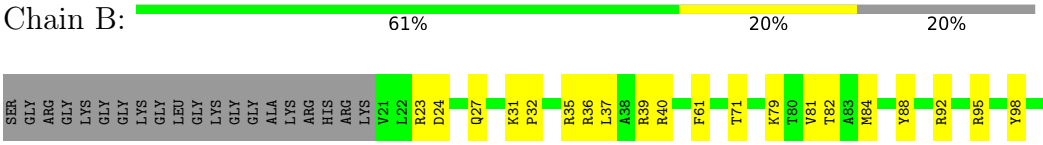
• Molecule 1: Histone H3

Chain E:



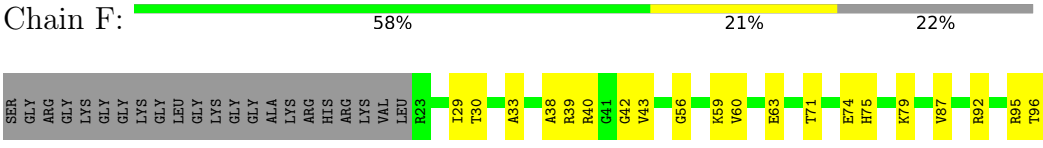
• Molecule 2: Histone H4

Chain B:



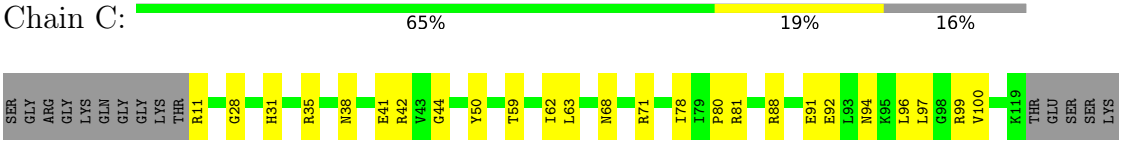
• Molecule 2: Histone H4

Chain F:



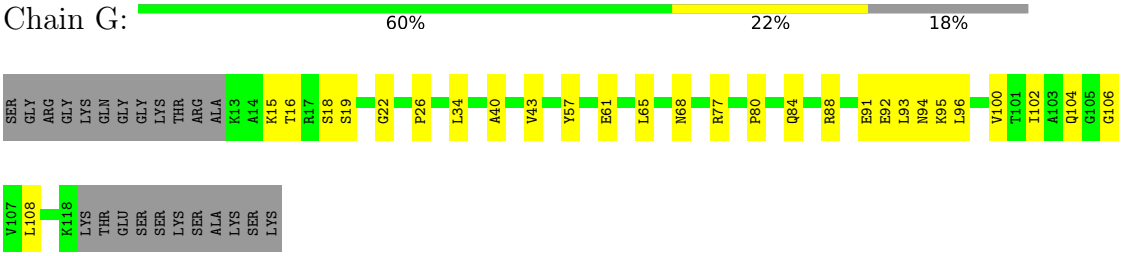
• Molecule 3: Histone H2A

Chain C:



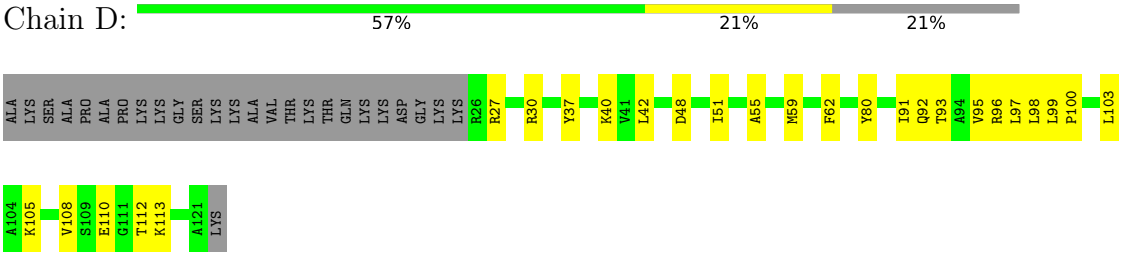
• Molecule 3: Histone H2A

Chain G:



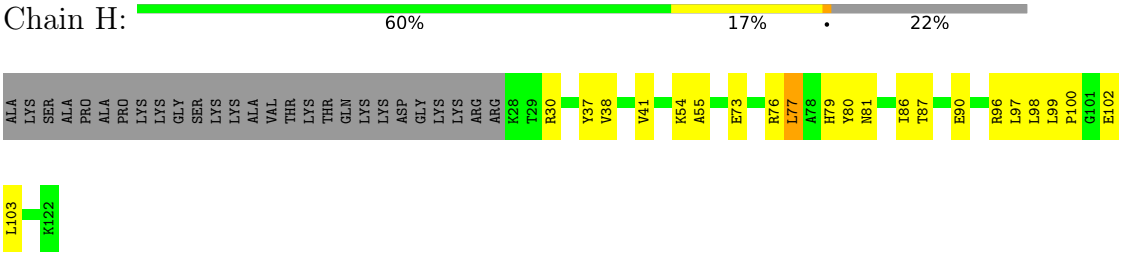
• Molecule 4: Histone H2B

Chain D:



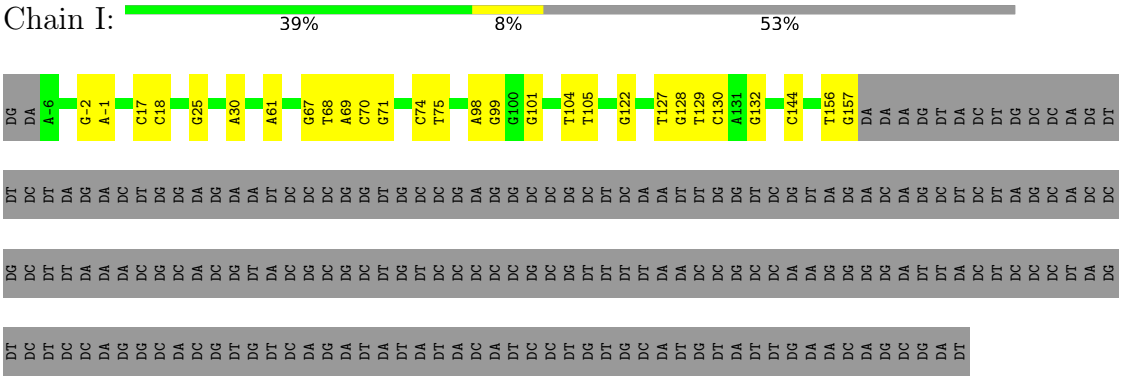
• Molecule 4: Histone H2B

Chain H:



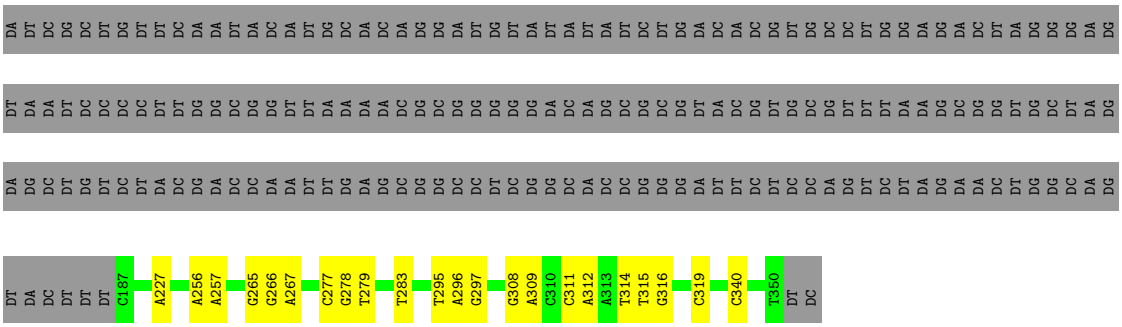
• Molecule 5: DNA (352-MER)

Chain I:



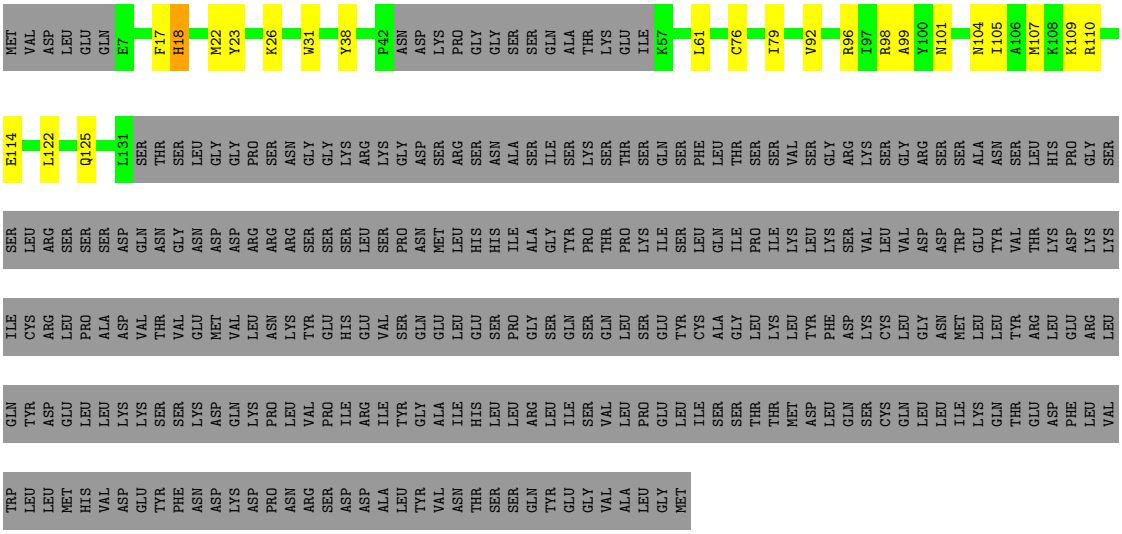
• Molecule 6: DNA (352-MER)

Chain J:



• Molecule 7: Chromatin modification-related protein EAF3

Chain M:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	42076	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	44	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ML3

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	0.33	0/836	0.80	0/1120
1	E	0.30	0/822	0.75	1/1103 (0.1%)
2	B	0.34	0/660	0.80	0/883
2	F	0.39	0/645	0.83	0/862
3	C	0.33	0/853	0.72	0/1149
3	G	0.33	0/828	0.68	0/1117
4	D	0.36	0/768	0.69	0/1032
4	H	0.36	0/756	0.69	2/1015 (0.2%)
5	I	0.59	0/3747	0.99	0/5777
6	J	0.53	0/3795	0.93	0/5859
7	M	0.28	0/937	0.56	1/1253 (0.1%)
All	All	0.46	0/14647	0.86	4/21170 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	H	97	LEU	CA-CB-CG	6.60	130.49	115.30
4	H	77	LEU	CA-CB-CG	5.65	128.29	115.30
7	M	18	HIS	N-CA-C	-5.53	96.08	111.00
1	E	122	LYS	CA-CB-CG	5.10	124.62	113.40

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen

atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	837	0	885	27	0
1	E	810	0	853	21	0
2	B	653	0	696	25	0
2	F	638	0	676	23	0
3	C	843	0	908	26	0
3	G	818	0	877	26	0
4	D	757	0	786	33	0
4	H	745	0	773	26	0
5	I	3345	0	1845	19	0
6	J	3379	0	1841	22	0
7	M	915	0	919	22	0
All	All	13740	0	11059	185	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 7.

The worst 5 of 185 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)
3:C:92:GLU:HG3	4:D:100:PRO:HG2	1.31	1.06
2:B:92:ARG:HH12	4:D:97:LEU:HB3	1.21	1.03
3:C:96:LEU:HD21	4:D:100:PRO:HD3	1.58	0.82
7:M:18:HIS:HB3	7:M:23:TYR:CE2	2.18	0.79
7:M:17:PHE:HD1	7:M:22:MET:HE1	1.49	0.78

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	99/135 (73%)	96 (97%)	3 (3%)	0	100	100
1	E	96/135 (71%)	94 (98%)	2 (2%)	0	100	100
2	B	80/102 (78%)	78 (98%)	2 (2%)	0	100	100
2	F	78/102 (76%)	76 (97%)	2 (3%)	0	100	100
3	C	107/129 (83%)	105 (98%)	2 (2%)	0	100	100
3	G	104/129 (81%)	101 (97%)	3 (3%)	0	100	100
4	D	94/122 (77%)	92 (98%)	2 (2%)	0	100	100
4	H	93/122 (76%)	84 (90%)	9 (10%)	0	100	100
7	M	107/401 (27%)	106 (99%)	1 (1%)	0	100	100
All	All	858/1377 (62%)	832 (97%)	26 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	86/108 (80%)	86 (100%)	0	100	100
1	E	85/108 (79%)	85 (100%)	0	100	100
2	B	67/78 (86%)	67 (100%)	0	100	100
2	F	65/78 (83%)	65 (100%)	0	100	100
3	C	86/101 (85%)	86 (100%)	0	100	100
3	G	84/101 (83%)	84 (100%)	0	100	100
4	D	82/102 (80%)	82 (100%)	0	100	100
4	H	81/102 (79%)	81 (100%)	0	100	100
7	M	96/359 (27%)	96 (100%)	0	100	100
All	All	732/1137 (64%)	732 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such

sidechains are listed below:

Mol	Chain	Res	Type
2	B	93	GLN
3	C	38	ASN
1	E	108	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	$\#\|Z\| > 2$
1	ML3	A	36	1	10,11,12	0.77	0	10,14,16	0.82	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
1	ML3	A	36	1	-	5/8/10/12	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	36	ML3	SG-CD-CE-NZ

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
1	A	36	ML3	CD-CE-NZ-CM1
1	A	36	ML3	CD-CE-NZ-CM2
1	A	36	ML3	CD-CE-NZ-CM3
1	A	36	ML3	CA-CB-SG-CD

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	36	ML3	4	0

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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.