



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 18, 2025 – 11:27 PM EDT

PDB ID : 3MET
Title : Crystal structure of SGF29 in complex with H3K4me2
Authors : Bian, C.B.; Xu, C.; Lam, R.; Bountra, C.; Arrowsmith, C.H.; Weigelt, J.;
Edwards, A.M.; Bochkarev, A.; Min, J.; Structural Genomics Consortium
(SGC)
Deposited on : 2010-03-31
Resolution : 2.00 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	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.21
EDS	:	FAILED
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41.4

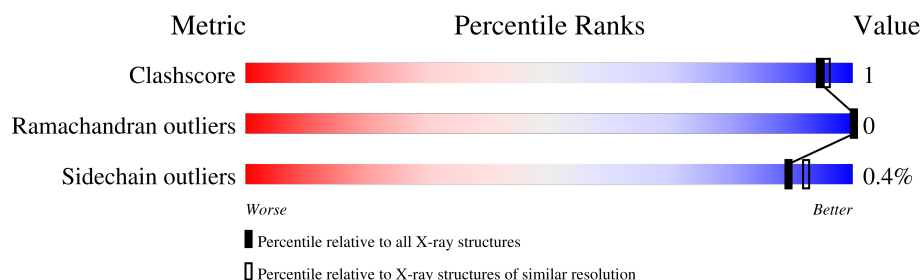
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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))
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)

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 $\leq 5\%$.

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	180	94% . .
1	B	180	94% . . .
2	C	11	73% 27%
2	D	11	36% 64%

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 3078 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 SAGA-associated factor 29 homolog.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	174	Total	C	N	O	S	Se	0	7	0
			1397	894	239	259	3	2			
1	B	174	Total	C	N	O	S	Se	0	5	0
			1375	878	236	255	3	3			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	114	GLY	-	insertion	UNP Q96ES7
B	114	GLY	-	insertion	UNP Q96ES7

- Molecule 2 is a protein called Histone H3.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	8	Total	C	N	O	0	0	0
			62	37	15	10			
2	D	4	Total	C	N	O	0	0	0
			28	18	5	5			

- Molecule 3 is UNKNOWN ATOM OR ION (three-letter code: UNX) (formula: X).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	X	0	0
			2	2		
3	B	1	Total	X	0	0
			1	1		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is SULFATE ION (three-letter code: SO₄) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	O	S	0	0
			5	4	1		
5	C	1	Total	O	S	0	0
			5	4	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	109	Total 109	O 109	0	0
6	B	84	Total 84	O 84	0	0
6	C	3	Total 3	O 3	0	0
6	D	1	Total 1	O 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. 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.

Note EDS failed to run properly.

- Molecule 1: SAGA-associated factor 29 homolog

Chain A:  94%



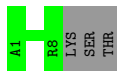
- Molecule 1: SAGA-associated factor 29 homolog

Chain B:  94%



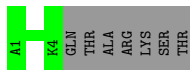
- Molecule 2: Histone H3

Chain C:  73% 27%



- Molecule 2: Histone H3

Chain D:  36% 64%



4 Data and refinement statistics

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	50.16Å 65.41Å 105.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.00	Depositor
% Data completeness (in resolution range)	99.9 (20.00-2.00)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	7.65 (at 2.01Å)	Xtriage
Refinement program	REFMAC 5.5.0072	Depositor
R, R_{free}	0.196 , 0.243	Depositor
Wilson B-factor (Å ²)	19.8	Xtriage
Anisotropy	0.005	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3078	wwPDB-VP
Average B, all atoms (Å ²)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.39% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, UNX, SO4, MLY

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	0.75	0/1449	0.77	0/1976
1	B	0.84	0/1420	0.73	0/1941
2	C	0.83	0/50	1.06	0/66
2	D	0.57	0/16	0.78	0/21
All	All	0.80	0/2935	0.75	0/4004

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	1397	0	1366	3	0
1	B	1375	0	1333	4	0
2	C	62	0	70	0	0
2	D	28	0	32	0	0
3	A	2	0	0	0	0
3	B	1	0	0	0	0
4	B	6	0	8	0	0
5	B	5	0	0	0	0
5	C	5	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	109	0	0	1	0
6	B	84	0	0	0	0
6	C	3	0	0	0	0
6	D	1	0	0	0	0
All	All	3078	0	2809	7	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 1.

All (7) 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:130:LEU:HD23	1:A:217[A]:LYS:HG2	1.91	0.53
1:B:128[A]:MSE:HE2	1:B:128[A]:MSE:HA	1.96	0.47
1:A:134:ILE:HD13	6:A:58:HOH:O	2.17	0.44
1:A:132:LEU:HD23	1:A:134:ILE:HD11	2.00	0.42
1:B:144:LEU:HD21	1:B:150:ALA:HB2	2.01	0.42
1:B:205:LEU:N	1:B:205:LEU:CD2	2.84	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	179/180 (99%)	176 (98%)	3 (2%)	0	100	100
1	B	177/180 (98%)	170 (96%)	7 (4%)	0	100	100
2	C	5/11 (46%)	5 (100%)	0	0	100	100
2	D	2/11 (18%)	2 (100%)	0	0	100	100
All	All	363/382 (95%)	353 (97%)	10 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	146/151 (97%)	146 (100%)	0	100	100
1	B	143/151 (95%)	142 (99%)	1 (1%)	81	86
2	C	4/8 (50%)	4 (100%)	0	100	100
2	D	1/8 (12%)	1 (100%)	0	100	100
All	All	294/318 (92%)	293 (100%)	1 (0%)	89	94

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

Mol	Chain	Res	Type
1	B	205	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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	MLY	D	4	2	9,10,11	0.47	0	6,11,13	0.45	0
2	MLY	C	4	2	9,10,11	0.55	0	6,11,13	0.29	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
2	MLY	D	4	2	-	3/8/9/11	-
2	MLY	C	4	2	-	0/8/9/11	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	4	MLY	CD-CE-NZ-CH1
2	D	4	MLY	CG-CD-CE-NZ
2	D	4	MLY	CE-CD-CG-CB

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 3 are unknown - leaving 3 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
4	GOL	B	1	-	5,5,5	0.73	0	5,5,5	0.85	0
5	SO4	C	12	-	4,4,4	0.28	0	6,6,6	0.26	0
5	SO4	B	2	-	4,4,4	0.28	0	6,6,6	0.31	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
4	GOL	B	1	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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

EDS failed to run properly - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

6.4 Ligands

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers

EDS failed to run properly - this section is therefore empty.