WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 10, 2022 – 04:47 pm BST

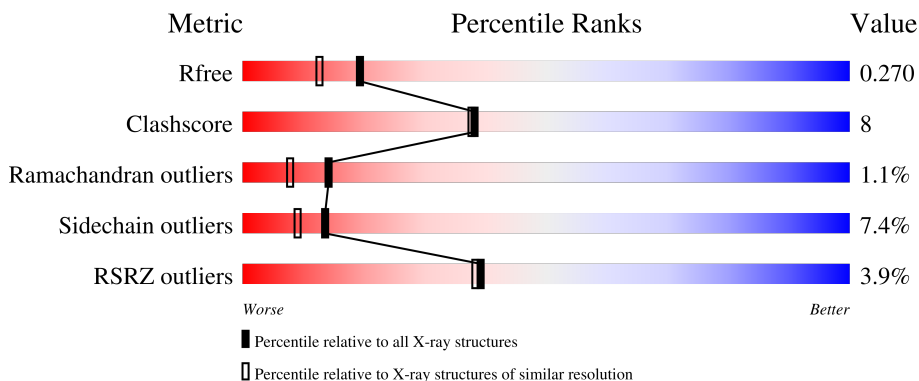|		
PDB ID	:	7ZUP
Title	:	Human PRMT5:MEP50 structure with Fragment (Example 18) and MTA Bound
Authors	:	Ahmad, M.U.; Koelmel, W.; Arkhipova, V.; Lawson, J.D.; Smith, C.R.; Gunn, R.J.
Deposited on	:	2022-05-12
Resolution	:	2.01 Å(reported)

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
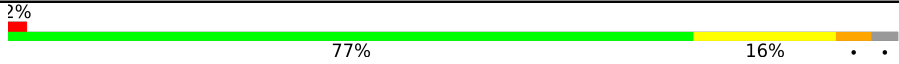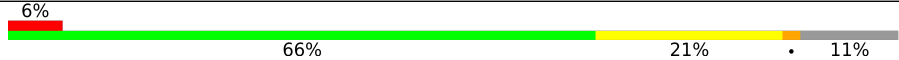*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.01 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	Percentile Ranks	Value
Rfree		0.270
Clashscore		8
Ramachandran outliers		1.1%
Sidechain outliers		7.4%
RSRZ outliers		3.9%

Worse *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (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 <=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	645	2% 77% 16% · ·
2	B	350	6% 66% 21% · 11%

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 7691 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 Protein arginine N-methyltransferase 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	625	Total 5085	C 3251	N 873	O 936	S 25	0	3	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	initiating methionine	UNP O14744
A	-6	ASP	-	expression tag	UNP O14744
A	-5	TYR	-	expression tag	UNP O14744
A	-4	LYS	-	expression tag	UNP O14744
A	-3	ASP	-	expression tag	UNP O14744
A	-2	ASP	-	expression tag	UNP O14744
A	-1	ASP	-	expression tag	UNP O14744
A	0	ASP	-	expression tag	UNP O14744
A	1	LYS	-	expression tag	UNP O14744

- Molecule 2 is a protein called Methylosome protein 50.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	310	Total 2345	C 1471	N 401	O 458	S 15	0	1	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-7	MET	-	initiating methionine	UNP Q9BQA1
B	-6	HIS	-	expression tag	UNP Q9BQA1
B	-5	HIS	-	expression tag	UNP Q9BQA1
B	-4	HIS	-	expression tag	UNP Q9BQA1
B	-3	HIS	-	expression tag	UNP Q9BQA1
B	-2	HIS	-	expression tag	UNP Q9BQA1
B	-1	HIS	-	expression tag	UNP Q9BQA1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	HIS	-	expression tag	UNP Q9BQA1
B	1	HIS	-	expression tag	UNP Q9BQA1

- Molecule 3 is 5'-DEOXY-5'-METHYLTHIOADENOSINE (three-letter code: MTA) (formula: $C_{11}H_{15}N_5O_3S$) (labeled as "Ligand of Interest" by depositor).
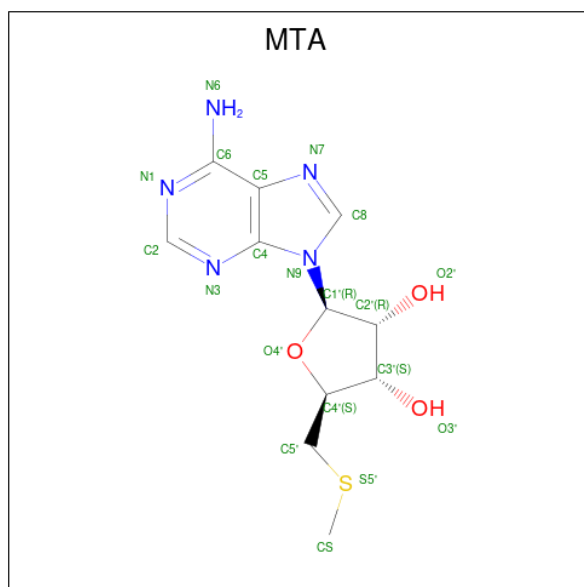
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
3	A	1	20	11	5	3	1	0	0

- Molecule 4 is 3-ethylimidazo[4,5-b]pyridin-2-amine (three-letter code: JYX) (formula: $C_8H_{10}N_4$) (labeled as "Ligand of Interest" by depositor).

JYX
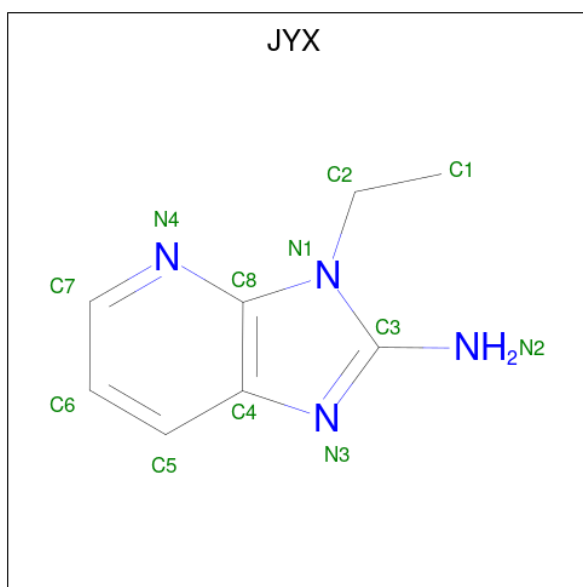
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total C N 12 8 4			0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

GOL

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

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	182	Total 182	O 182	0	0
6	B	35	Total 35	O 35	0	0

3 Residue-property plots ⓘ

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: Protein arginine N-methyltransferase 5
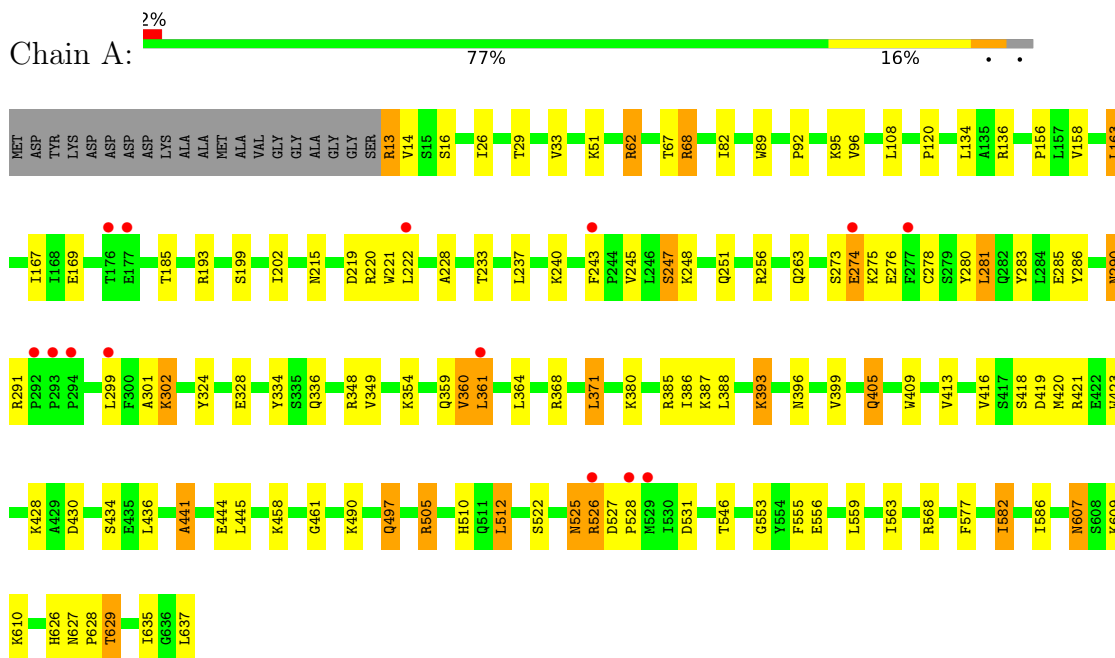
• Molecule 2: Methylosome protein 50
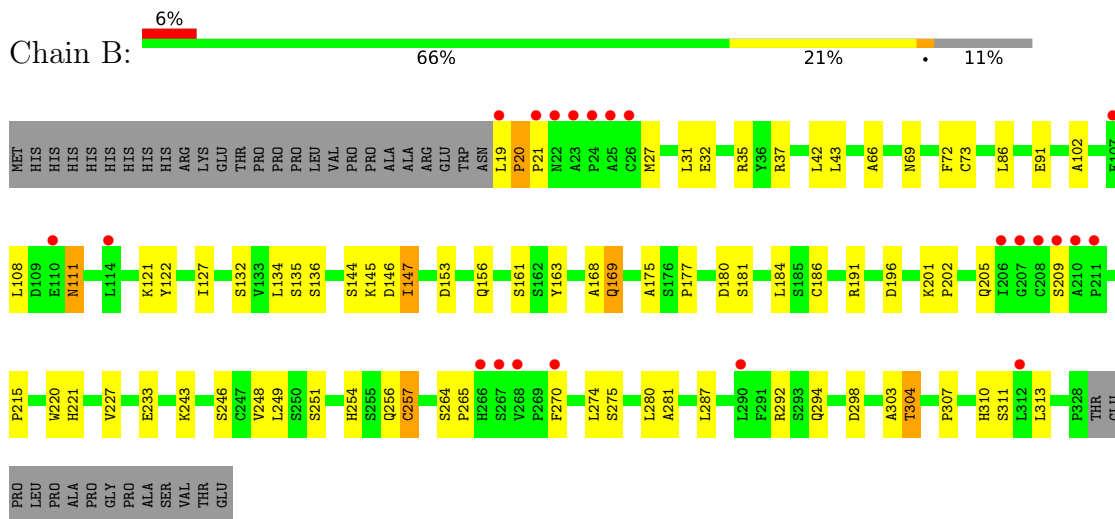
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	104.28Å 138.65Å 178.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	109.59 – 2.01 109.59 – 2.01	Depositor EDS
% Data completeness (in resolution range)	46.2 (109.59-2.01) 46.2 (109.59-2.01)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.07 (at 2.02Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.219 , 0.271 0.223 , 0.270	Depositor DCC
R_{free} test set	1957 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å^2)	33.0	Xtriage
Anisotropy	0.120	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	7691	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.23% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: MTA, JYX, GOL

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	0.65	0/5227	0.84	3/7111 (0.0%)
2	B	0.72	0/2402	0.83	0/3283
All	All	0.67	0/7629	0.84	3/10394 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	607	ASN	CB-CA-C	-5.88	98.65	110.40
1	A	497	GLN	CB-CA-C	-5.55	99.30	110.40
1	A	62	ARG	NE-CZ-NH1	-5.31	117.64	120.30

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	5085	0	4970	78	0
2	B	2345	0	2260	37	0
3	A	20	0	15	0	0
4	A	12	0	0	1	0
5	A	6	0	8	0	0
5	B	6	0	8	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	182	0	0	3	0
6	B	35	0	0	3	0
All	All	7691	0	7261	112	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 8.

The worst 5 of 112 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)
2:B:35:ARG:HA	2:B:304:THR:HG21	1.66	0.77
2:B:135:SER:HB3	2:B:177:PRO:O	1.88	0.73
1:A:219:ASP:O	1:A:222:LEU:HG	1.89	0.73
1:A:324:TYR:CB	1:A:368:ARG:HD2	2.20	0.71
1:A:276:GLU:HB3	1:A:278:CYS:SG	2.32	0.70

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	626/645 (97%)	579 (92%)	43 (7%)	4 (1%)	25	19
2	B	309/350 (88%)	292 (94%)	11 (4%)	6 (2%)	8	3
All	All	935/995 (94%)	871 (93%)	54 (6%)	10 (1%)	14	8

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	147	ILE
1	A	274	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	20	PRO
2	B	243	LYS
2	B	248	VAL

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	561/570 (98%)	518 (92%)	43 (8%)	13 8
2	B	264/298 (89%)	245 (93%)	19 (7%)	14 9
All	All	825/868 (95%)	763 (92%)	62 (8%)	13 9

5 of 62 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	387	LYS
2	B	256	GLN
1	A	512	LEU
2	B	251	SER
2	B	294	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	310	HIS
2	B	169	GLN
1	A	525	ASN
1	A	359	GLN
1	A	607	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

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

5.6 Ligand geometry ⓘ

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

5.7 Other polymers ⓘ

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

5.8 Polymer linkage issues ⓘ

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, 95th 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(Å²)	Q<0.9
1	A	625/645 (96%)	0.26	14 (2%)	62	60	17, 26, 59, 90	0
2	B	310/350 (88%)	0.49	22 (7%)	16	15	23, 38, 63, 102	0
All	All	935/995 (93%)	0.34	36 (3%)	39	38	17, 30, 60, 102	0

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	209	SER	5.3
2	B	210	ALA	5.3
1	A	274	GLU	5.1
2	B	19	LEU	4.9
2	B	21	PRO	4.8

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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, 95th 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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	GOL	A	703	6/6	0.91	0.19	45,46,48,50	0
5	GOL	B	401	6/6	0.91	0.24	65,67,68,69	0
3	MTA	A	701	20/20	0.97	0.15	18,20,21,21	0
4	JYX	A	702	12/12	0.97	0.13	17,18,18,18	0

6.5 Other polymers ⓘ

There are no such residues in this entry.