



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 01:36 PM GMT

PDB ID : 3UA4
Title : Crystal Structure of Protein Arginine Methyltransferase PRMT5
Authors : Sun, L.; Wang, M.; Lv, Z.; Yang, N.; Liu, Y.; Bao, S.; Gong, W.; Xu, R.M.
Deposited on : 2011-10-21
Resolution : 3.00 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
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) : trunk26865

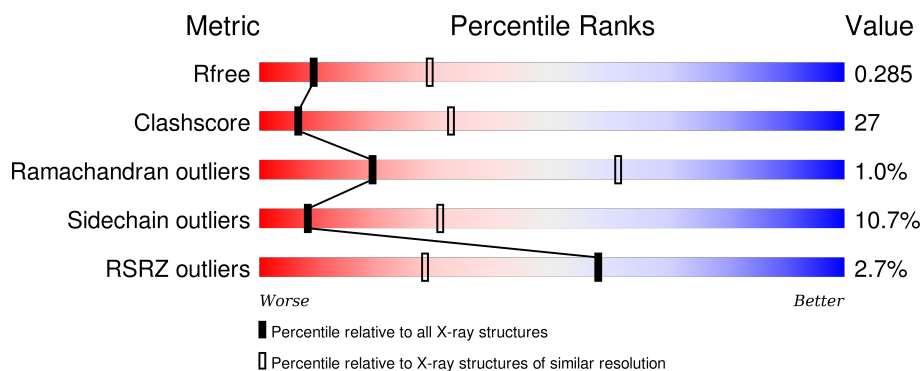
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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))
R_{free}	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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 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 $\leq 5\%$. 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	745	<div> <div>2%</div> <div>45%</div> <div>37%</div> <div>•</div> <div>14%</div> </div>
1	B	745	<div> <div>3%</div> <div>41%</div> <div>38%</div> <div>6%</div> <div>16%</div> </div>

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
2	GOL	B	743	-	-	-	X

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10345 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	640	Total	C	N	O	S	0	0	0
			5143	3291	875	957	20			
1	B	628	Total	C	N	O	S	1	0	0
			5049	3238	854	937	20			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	MET	-	EXPRESSION TAG	UNP P46580
A	-1	ALA	-	EXPRESSION TAG	UNP P46580
A	0	SER	-	EXPRESSION TAG	UNP P46580
A	735	LEU	-	EXPRESSION TAG	UNP P46580
A	736	GLU	-	EXPRESSION TAG	UNP P46580
A	737	HIS	-	EXPRESSION TAG	UNP P46580
A	738	HIS	-	EXPRESSION TAG	UNP P46580
A	739	HIS	-	EXPRESSION TAG	UNP P46580
A	740	HIS	-	EXPRESSION TAG	UNP P46580
A	741	HIS	-	EXPRESSION TAG	UNP P46580
A	742	HIS	-	EXPRESSION TAG	UNP P46580
B	-2	MET	-	EXPRESSION TAG	UNP P46580
B	-1	ALA	-	EXPRESSION TAG	UNP P46580
B	0	SER	-	EXPRESSION TAG	UNP P46580
B	735	LEU	-	EXPRESSION TAG	UNP P46580
B	736	GLU	-	EXPRESSION TAG	UNP P46580
B	737	HIS	-	EXPRESSION TAG	UNP P46580
B	738	HIS	-	EXPRESSION TAG	UNP P46580
B	739	HIS	-	EXPRESSION TAG	UNP P46580
B	740	HIS	-	EXPRESSION TAG	UNP P46580
B	741	HIS	-	EXPRESSION TAG	UNP P46580
B	742	HIS	-	EXPRESSION TAG	UNP P46580

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	74	Total	O	0	0
			74	74		
3	B	73	Total	O	0	0
			73	73		

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.

- [illegible]

- Chain B: 

V674	I495	K426	ALA	A299	C227	E151	V59	MET
I675	I496	S427	PRO	L300	G228	L152	A60	ALA
P676	V497	E428	LEU	R301	M229	S153	B61	SER
L677	V595	R429	GLN	I302		Y154	N62	MET
D679	V596	E430	PRO	V303	Y234	I155	L63	SER
	L500	Y431	LEU	L304	F235	C156		ASN
	L501	N432	SER	R305	Q236	Y157	A68	ARG
D687	D506	N433	GLU	A306	Q237	L158	T69	THR
R688	M507	T434	ASN	T307	A238	L160	F70	TYR
I689	E508	F435	LEU	T308	L239			ALA
S690	L509	R436	ASP	D309	T240			ASP
L691	S510	GLN	SER	T310		M163	R73	ASN
K692	P511	GLN	GLY	F311	L245	A164	L74	LEU
I693	E512	GLU	VAL	K312		I165		PHE
D694	C513	S441	TYR	Y313	E248	E166	F77	PRO
R695	E514	L442	ASN	N314	L249		K78	GLN
K696	H615	L443	THR	T315	T250	S171	N80	GLN
V697	P616	K443	PHE	S316	E251	K172	F81	VAL
D698	N617	N444	GLU	I317		P173	V82	ALA
K699	F618	K445	Q381	K318	L254	R174		GLU
T700	E524	L446	D382	S319	V255		P85	GLN
G701	K522	Y447	Q383	E320	D256	I178	I86	HIS
V702	R625	T448	I384	Y321	R257	L179	G87	GLU
W703	T525	V449			K258	K180	G88	GLU
Y704	I526	E450	D387	R326	K259	K181	V89	GLN
E705	S527		V388	R327		W182	V90	MET
W706	I528	T456	T389	A328	L263	I183	R91	SER
H707	P529	Y457	G390	VAL	A264	L184	A92	SER
W708	Q530	T458	E391	ARG	T265	T185	F93	GLY
E709	K531	Y461	ASN	ASN	F266	R186	N94	SER
K710	L639	Y462	VAL	VAL	T267	N187	T95	SER
K711	Y532	M463	V393	ASN	T268	S188	P96	SER
K712	V536	M464	V394	TYR	E269	R189		PRO
K713	K537	Y465	G395	ARG	S270		N97	LYS
N714	S541	R466	K398	ARG	G271	W193		SER
G715	D541	T466	D399	ARG	L272	V194	E101	ASN
D67	L544	K467	L400	PRO	F273	Q195	M102	SER
S717	I544	K468	G401	ASP	I1E	L196	P105	PRO
	L556	R469	A402	VAL	SER	P197	Y106	SER
T720	S557	R470	D403	GLY	GLY	S198	I107	ARG
F721	V651	V471	G404	GLU	ARG	A199		SER
I722	R558	T472	R405	GLY	ASN	I200	V112	ILE
Q723	T559	L473	K406	SER	GLY	E201	Q113	SER
N724	I560	T474	T407	ASN	GLU	K202	L114	VAL
	P561	E475	V408	ASP	A281		R115	GLU
G727	S562	S476	V409	SER	S282	Y206	D117	ALA
Y730	E569	N478	I410	THR	T283	D207		A44
Y731	E569	R479	Y411	HIS	P284	A208		N45
M732	E572	S480	L412	LEU	H287	F209	E120	S46
R733	D573	L481	G415	ASN	T288	T210		R47
LEU	W576	P482	G418	VAL	I289	I211	K126	
GLU	I577	T484	P419	ILE	L290	E212	I127	I50
HIS	Q578	A485	K420	GLU	L291		S128	G51
HIS	K579	K486	I420	TVR	K292	D215		N52
HIS	V585	Q492	G421	LYS	H293	I131		N53
HIS	P493	D494	T422	ASP	L294	D134		T55
HIS	A673		K423	VAL	T297			T56
			I424	LEU	D298	L143		L57
			L425	GLN				D58

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	100.56 Å 129.49 Å 149.72 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.02 – 3.00 30.02 – 3.01	Depositor EDS
% Data completeness (in resolution range)	99.4 (30.02-3.00) 99.4 (30.02-3.01)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.86 (at 3.00 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.2_432)	Depositor
R, R_{free}	0.232 , 0.288 0.227 , 0.285	Depositor DCC
R_{free} test set	1817 reflections (4.62%)	DCC
Wilson B-factor (Å ²)	43.1	Xtriage
Anisotropy	0.015	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 87.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 39351 reflections	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	10345	wwPDB-VP
Average B, all atoms (Å ²)	99.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.42% 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, 0.375 respectively for untwinned datasets, and 0.333, 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

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	0.28	0/5272	0.47	0/7161
1	B	0.27	0/5176	0.47	0/7031
All	All	0.27	0/10448	0.47	0/14192

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	5143	0	5096	272	0
1	B	5049	0	5005	285	0
2	B	6	0	8	0	0
3	A	74	0	0	5	0
3	B	73	0	0	2	0
All	All	10345	0	10109	553	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 27.

The worst 5 of 553 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:B:724:ASN:HD21	1:B:730:TYR:HB3	1.10	1.15
1:A:533:THR:HB	1:A:612:THR:HG22	1.23	1.08
1:A:635:ARG:HH11	1:A:635:ARG:HG2	1.16	1.07
1:A:74:LEU:HB3	1:A:79:TYR:HE1	1.20	1.07
1:B:198:SER:HB2	1:B:202:LYS:HD3	1.48	0.96

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	636/745 (85%)	551 (87%)	77 (12%)	8 (1%)	15	53
1	B	620/745 (83%)	538 (87%)	77 (12%)	5 (1%)	24	66
All	All	1256/1490 (84%)	1089 (87%)	154 (12%)	13 (1%)	19	61

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	617	ASN
1	A	182	TRP
1	A	284	PRO
1	B	212	GLU
1	B	509	LEU

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	567/662 (86%)	511 (90%)	56 (10%)	10	35
1	B	557/662 (84%)	493 (88%)	64 (12%)	7	28
All	All	1124/1324 (85%)	1004 (89%)	120 (11%)	8	31

5 of 120 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	714	ASN
1	B	180	LYS
1	B	659	GLU
1	A	717	SER
1	B	95	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	195	GLN
1	B	287	HIS
1	B	649	GLN
1	A	649	GLN
1	B	563	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

1 ligand 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	$\# Z > 2$
2	GOL	B	743	-	5,5,5	0.33	0	5,5,5	0.26	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	GOL	B	743	-	-	0/4/4/4	0/0/0/0

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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(Å ²)	Q<0.9
1	A	640/745 (85%)	-0.21	14 (2%) 65 35	14, 91, 179, 238	1 (0%)
1	B	628/745 (84%)	-0.08	20 (3%) 51 23	23, 101, 190, 251	15 (2%)
All	All	1268/1490 (85%)	-0.14	34 (2%) 58 28	14, 96, 183, 251	16 (1%)

The worst 5 of 34 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	313	TYR	4.9
1	A	275	SER	4.5
1	B	429	ARG	4.4
1	B	97	ASN	3.9
1	B	308	THR	3.6

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

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, 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.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	GOL	B	743	6/6	0.91	0.45	1.42	86,91,92,94	0

6.5 Other polymers ⓘ

There are no such residues in this entry.