



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 12:48 am 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](mailto: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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

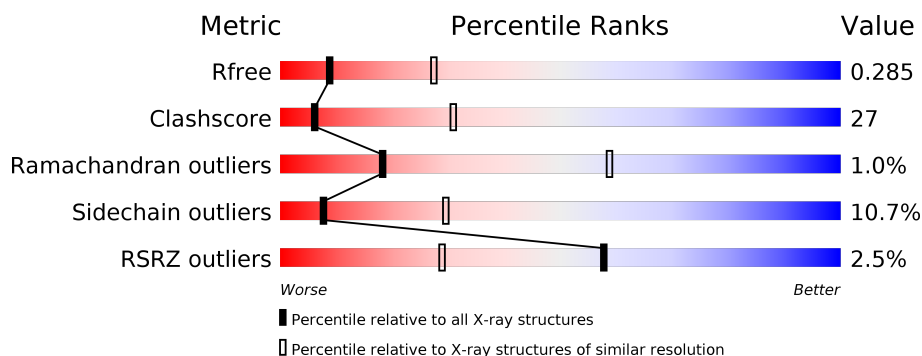
# 1 Overall quality at a glance

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))
$R_{free}$	100719	1692 (3.00-3.00)
Clashscore	112137	2037 (3.00-3.00)
Ramachandran outliers	110173	1973 (3.00-3.00)
Sidechain outliers	110143	1976 (3.00-3.00)
RSRZ outliers	101464	1716 (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  $\geq 3$ , 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> <div></div> <div>45%</div> <div>37%</div> <div>•</div> <div>14%</div> </div> </div>
1	B	745	<div> <div>3%</div> <div> <div></div> <div>41%</div> <div>38%</div> <div>6%</div> <div>16%</div> </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: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

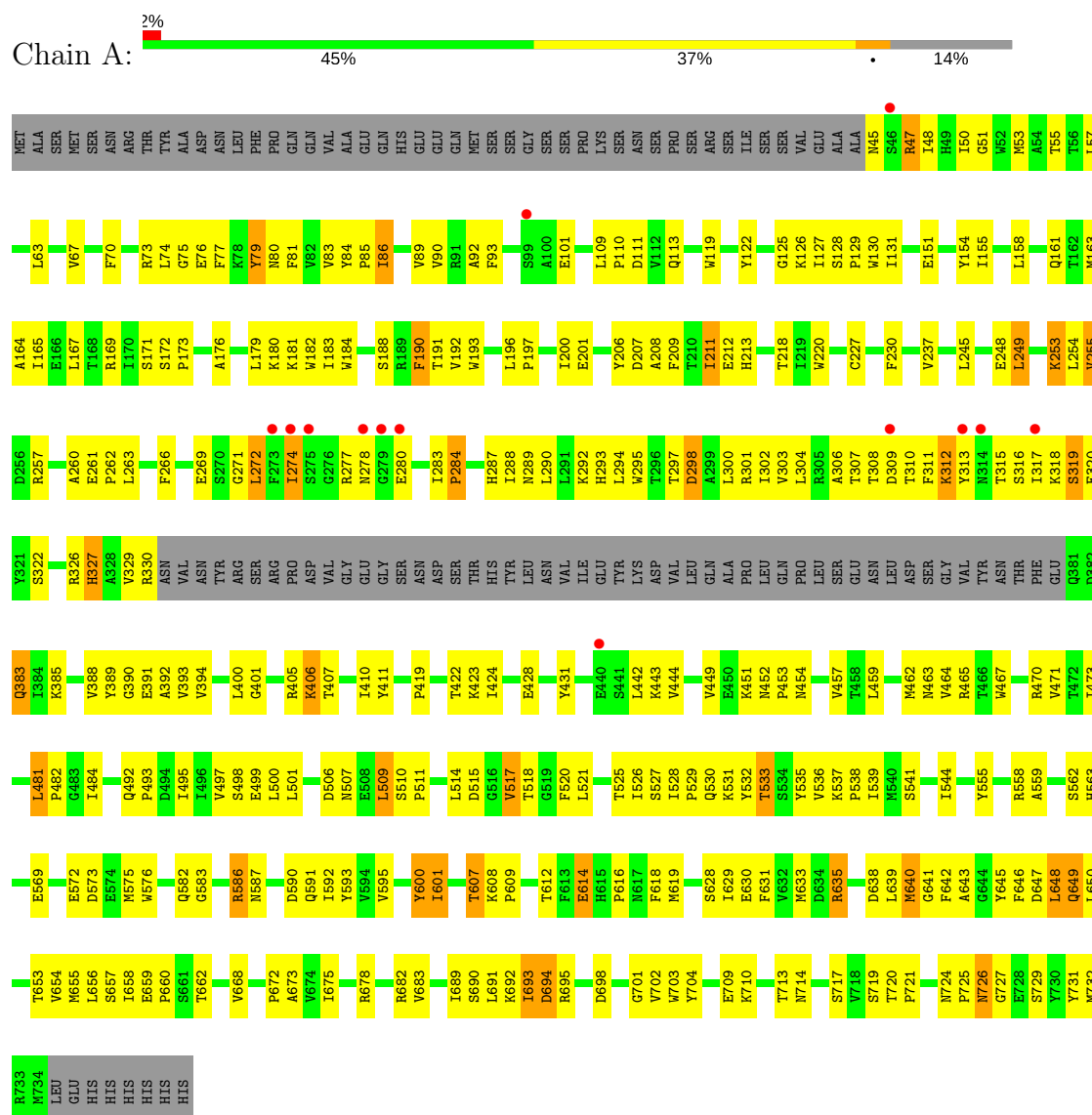
- Molecule 3 is water.

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

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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



V674	Y593	I495	K426	ALA	A299	C227	E151	V59	MET
I675	V594	I496	S427	PRO	L500	G228	L152	A60	ALA
P676	V595	V497	E428	LEU	R301	M229	S153	B61	SER
L677	V596	L500	R429	GLN	T302		Y154	N62	MET
D678	V596	L500	E430	PRO	V303	Y234	I155	L63	SER
D679	Y600	L501	Y431	LEU	L394	F235	C156		ASN
D687	I601	D506	N432	SER	R305	Q236	Y157	A68	ARG
R688	P602	N507	N433	GLU	A306	V237	L158	T69	THR
I689		E508	T434	ASN	T307	A238	G159	F70	TYR
S690	T607	E509	F435	LEU	T308	L239	L160		ALA
L691		S510	R436	ASP	D309	T240		R73	ASP
K692	F611	P511	GLN	SER	T310		M163	L74	ASN
I693	T612	E512	GLN	GLY	F311	L245	A164		LEU
D694	F613	C513	GLU	VAL	K312		I165	F77	PHE
R695	E614	S441	GLU	TYR	Y313	E248	E166		PRO
K696	H615	S441	GLU	ASN	N314	L249		K78	GLN
P697	P616	V517	L442	ASN	T315		S171	Y79	GLN
D698	N617	T518	K443	THR	T316	E251	S172	N80	GLN
R699	F618	G519	V444	PHE	S316		P173	F81	VAL
T700		F520	K445	GLU	I317	L254	R174	V82	ALA
G701	E624	L521	L446	Q381	K318	V255		P85	GLU
V702	R625	K522	Y447	D382	S319	D256	I178	I86	GLN
W703		T525	I448	Q383	E320	R257	L179	G87	GLU
W704	I629	T526	V449	I384	Y321	W258	K180	G88	GLU
E705		S527	E450	D387	B326	K259	W182	V90	GLN
W706	R635	F528	T456	V388	H327		I183	R91	MET
H707	N636	P529	V457	Y389	A328	L263	W184	A92	SER
V708	A637	Q530	T458	G390	VAL	A265	T185	F93	SER
E709	D638	K531		ARG	ARG	F266	R186	N94	GLY
K710	L639	Y532	Y461	ASN	ASN	T267	N187	T95	SER
K711	M640		M462	ASN	ASN	T268	S188	P96	SER
K712	G641	V536	M463	TYR	TYR	E269	R189	N97	PRO
T713	F642	K537	V464	ARG	ARG	S270	W193		LYS
N714		S541	R465	K398	SER	G271	V194	E101	ASN
F646	F646		T466	D399	ARG	L272	Q195	N102	ASN
D647	L648	I544	N467	L400	PRO	F273	Q196		SER
S717	Q649		K468	G401	ASP	I1E	L196	P105	SER
T720	L650	L556	R469	A402	VAL	SER	P197	V106	ARG
P721	Y651	S557	V471	D403	GLY	GLY	S198	I107	SER
I722	K652	R558	T472	G404	GLU	ARG	A199	V112	ILE
Q723	T653	A559	I473	R405	GLY	ASN	I200	Q113	SER
N724	V654	I560	T474	T407	SER	GLY	E201	L114	SER
	M655	P561	E475	K406	ASN	GLU	K202	R115	VAL
G727	L656	S562	S476	V408	ASP	A281		D117	GLU
Y730	S657	E569	D477	I410	SER	S282	Y206	N116	ALA
Y731	I658		N478	Y411	THR	T283	D207	D117	ALA
W732	E659	E572	R479	L412	HIS	P284	F209	E120	N45
R733	P660	D573	S480	G415	LEU	H287	T210		N46
LEU	H663	D573	L481	G418	ASN	N288	I211	K126	R47
GLU	T664	W576	P482	G418	VAL	I289	E212	K126	I50
HIS	M667	I577	I484	P419	ILE	L290		S128	G51
HIS	V668	Q578	A485	I420	GLU	L291	D215		W52
HIS	S669	K579	K486	I420	TYR	K292	T218	I131	M53
HIS	P672	V585	Q492	T422	LYS	H293	I219	D134	T55
HIS	A673	M589	D494	K423	ASP	L294	A221	L143	T56
				I424	VAL	T297			L57
				L425	GLN	D298			D58

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.56Å 129.49Å 149.72Å 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
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.86 (at 3.00Å)	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 <sup>2</sup> )	43.1	Xtriage
Anisotropy	0.015	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 87.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	10345	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	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.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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

### 5.1 Standard geometry

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

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%)	14	51
1	B	620/745 (83%)	538 (87%)	77 (12%)	5 (1%)	22	64
All	All	1256/1490 (84%)	1089 (87%)	154 (12%)	13 (1%)	18	59

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%)	9	34
1	B	557/662 (84%)	493 (88%)	64 (12%)	6	27
All	All	1124/1324 (85%)	1004 (89%)	120 (11%)	8	29

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	0.28	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	640/745 (85%)	-0.23	13 (2%) 65 36	14, 91, 179, 238	1 (0%)
1	B	628/745 (84%)	-0.10	19 (3%) 51 23	23, 101, 190, 251	15 (2%)
All	All	1268/1490 (85%)	-0.16	32 (2%) 58 29	14, 96, 183, 251	16 (1%)

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	313	TYR	4.7
1	A	275	SER	4.3
1	B	429	ARG	4.2
1	B	97	ASN	3.8
1	B	308	THR	3.5

### 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	GOL	B	743	6/6	0.91	0.45	1.42	86,91,92,94	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.