wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 10, 2022 – 12:54 PM EST

PDB ID	:	1WYZ
Title	:	X-Ray structure of the putative methyltransferase from Bacteroides thetaiotaomicron VPI-5482 at the resolution 2.5 A. Norteast Structural Genomics Consortium target Btr28
Authors	:	Kuzin, A.P.; Chen, Y.; Forouhar, F.; Vorobiev, S.M.; Acton, T.; Ma, L.-C.; Xiao, R.; Montelione, G.T.; Hunt, J.F.; Tong, L.; Northeast Structural Genomics Consortium (NESG)
Deposited on	:	2005-02-21
Resolution	:	2.50 Å(reported)

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

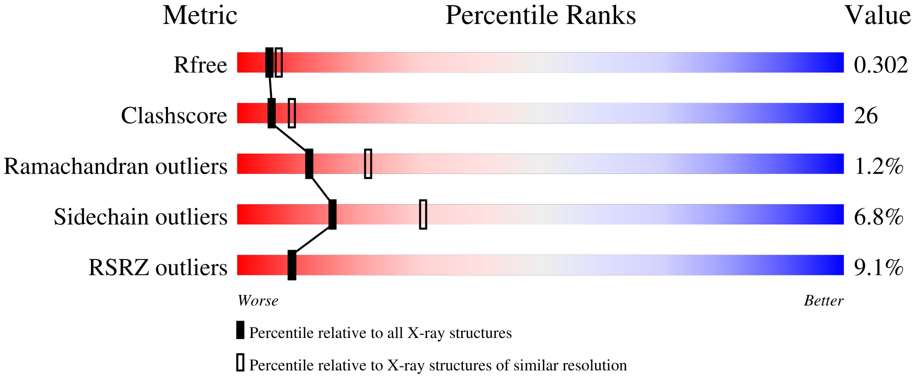
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)

1 Overall quality at a glance ⓘ

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

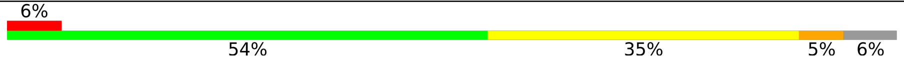
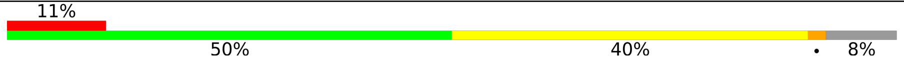
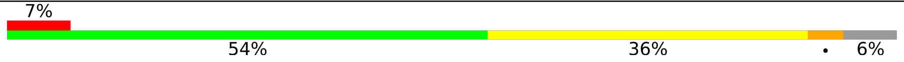
The reported resolution of this entry is 2.50 Å.

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}	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	242	
1	B	242	
1	C	242	
1	D	242	

Validation Pipeline (wwPDB-VP) : 2.31.2

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 7245 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 putative S-adenosylmethionine-dependent methyltransferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	223	Total 1765	C 1134	N 301	O 323	S 4	Se 3	0	0	0
1	B	227	Total 1796	C 1154	N 305	O 330	S 4	Se 3	0	0	0
1	C	223	Total 1762	C 1134	N 297	O 324	S 4	Se 3	0	0	0
1	D	227	Total 1796	C 1154	N 305	O 330	S 4	Se 3	0	0	0

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	modified residue	UNP Q8A031
A	86	MSE	MET	modified residue	UNP Q8A031
A	129	MSE	MET	modified residue	UNP Q8A031
A	181	MSE	MET	modified residue	UNP Q8A031
A	235	LEU	-	expression tag	UNP Q8A031
A	236	GLU	-	expression tag	UNP Q8A031
A	237	HIS	-	expression tag	UNP Q8A031
A	238	HIS	-	expression tag	UNP Q8A031
A	239	HIS	-	expression tag	UNP Q8A031
A	240	HIS	-	expression tag	UNP Q8A031
A	241	HIS	-	expression tag	UNP Q8A031
A	242	HIS	-	expression tag	UNP Q8A031
B	1	MSE	MET	modified residue	UNP Q8A031
B	86	MSE	MET	modified residue	UNP Q8A031
B	129	MSE	MET	modified residue	UNP Q8A031
B	181	MSE	MET	modified residue	UNP Q8A031
B	235	LEU	-	expression tag	UNP Q8A031
B	236	GLU	-	expression tag	UNP Q8A031
B	237	HIS	-	expression tag	UNP Q8A031
B	238	HIS	-	expression tag	UNP Q8A031
B	239	HIS	-	expression tag	UNP Q8A031

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	240	HIS	-	expression tag	UNP Q8A031
B	241	HIS	-	expression tag	UNP Q8A031
B	242	HIS	-	expression tag	UNP Q8A031
C	1	MSE	MET	modified residue	UNP Q8A031
C	86	MSE	MET	modified residue	UNP Q8A031
C	129	MSE	MET	modified residue	UNP Q8A031
C	181	MSE	MET	modified residue	UNP Q8A031
C	235	LEU	-	expression tag	UNP Q8A031
C	236	GLU	-	expression tag	UNP Q8A031
C	237	HIS	-	expression tag	UNP Q8A031
C	238	HIS	-	expression tag	UNP Q8A031
C	239	HIS	-	expression tag	UNP Q8A031
C	240	HIS	-	expression tag	UNP Q8A031
C	241	HIS	-	expression tag	UNP Q8A031
C	242	HIS	-	expression tag	UNP Q8A031
D	1	MSE	MET	modified residue	UNP Q8A031
D	86	MSE	MET	modified residue	UNP Q8A031
D	129	MSE	MET	modified residue	UNP Q8A031
D	181	MSE	MET	modified residue	UNP Q8A031
D	235	LEU	-	expression tag	UNP Q8A031
D	236	GLU	-	expression tag	UNP Q8A031
D	237	HIS	-	expression tag	UNP Q8A031
D	238	HIS	-	expression tag	UNP Q8A031
D	239	HIS	-	expression tag	UNP Q8A031
D	240	HIS	-	expression tag	UNP Q8A031
D	241	HIS	-	expression tag	UNP Q8A031
D	242	HIS	-	expression tag	UNP Q8A031

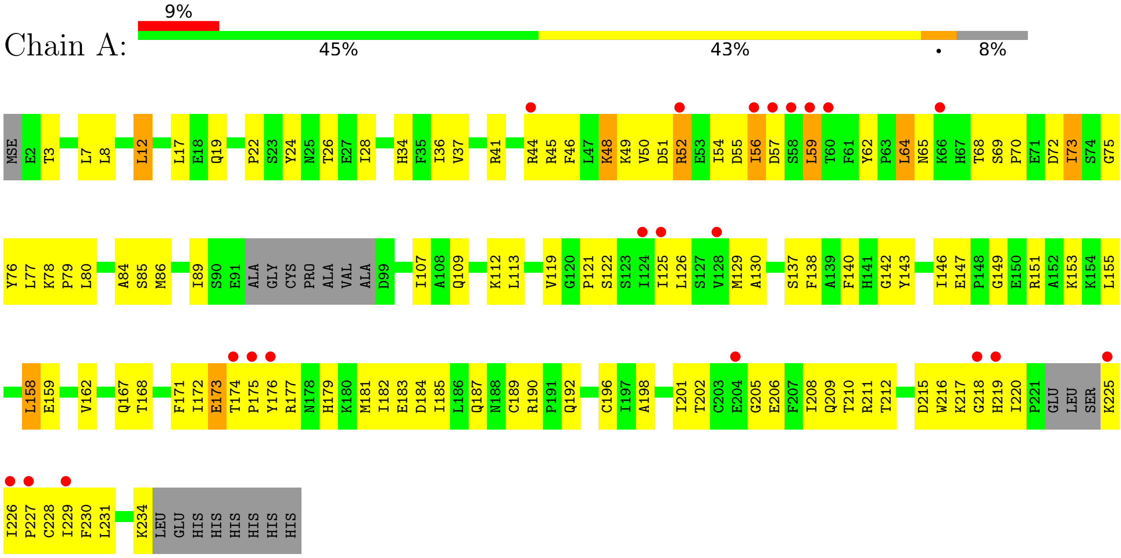
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	34	Total O 34 34	0	0
2	B	38	Total O 38 38	0	0
2	C	28	Total O 28 28	0	0
2	D	26	Total O 26 26	0	0

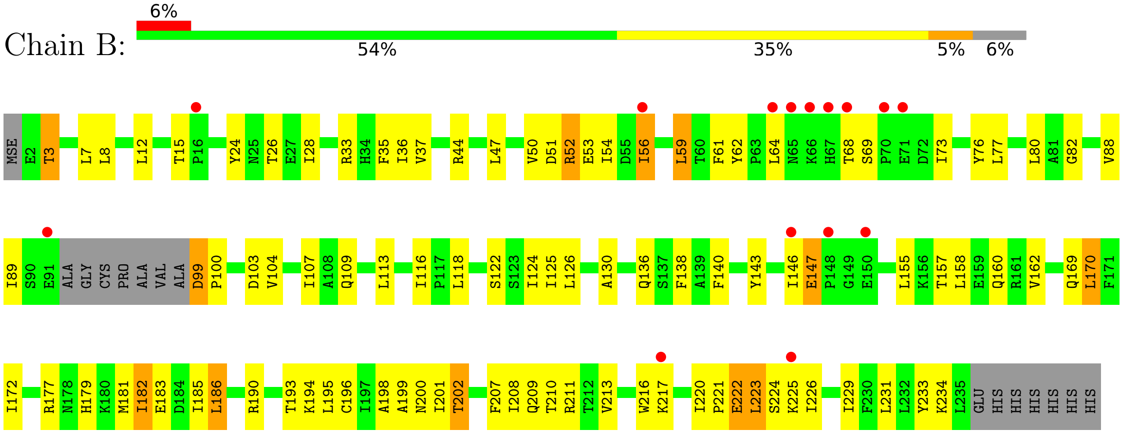
3 Residue-property plots (i)

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 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.

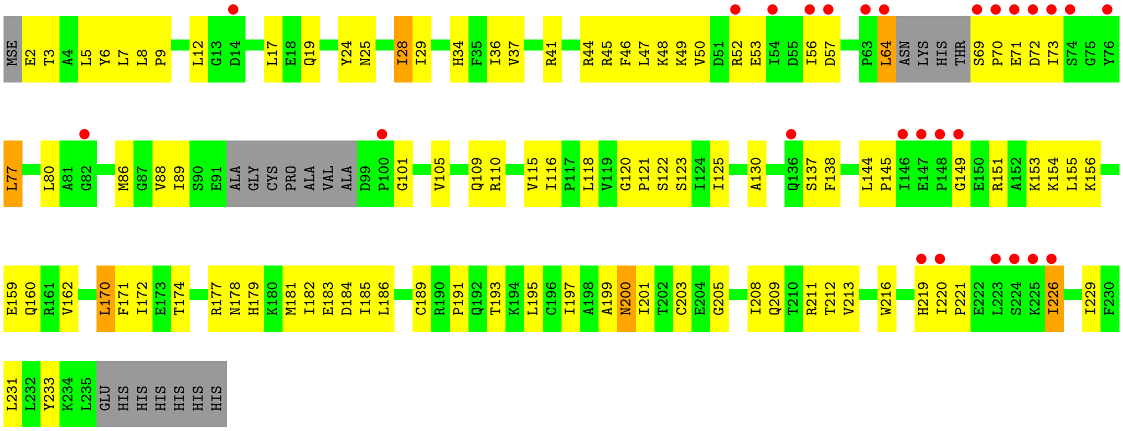
- Molecule 1: putative S-adenosylmethionine-dependent methyltransferase



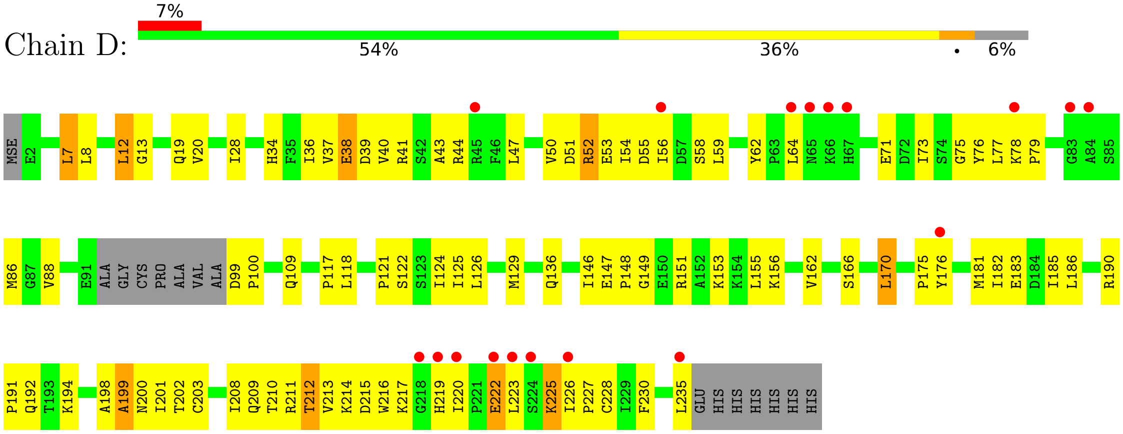
- Molecule 1: putative S-adenosylmethionine-dependent methyltransferase



- Molecule 1: putative S-adenosylmethionine-dependent methyltransferase



- Molecule 1: putative S-adenosylmethionine-dependent methyltransferase



4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	85.78Å 101.54Å 108.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.74 – 2.50 28.74 – 2.49	Depositor EDS
% Data completeness (in resolution range)	88.3 (28.74-2.50) 97.1 (28.74-2.49)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.89 (at 2.48Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.226 , 0.293 0.238 , 0.302	Depositor DCC
R_{free} test set	3077 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å^2)	40.3	Xtriage
Anisotropy	0.511	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 71.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	7245	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.40% 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 ⓘ

5.1 Standard geometry ⓘ

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	0.39	0/1798	0.66	2/2428 (0.1%)
1	B	0.40	0/1830	0.65	0/2473
1	C	0.41	0/1794	0.68	1/2423 (0.0%)
1	D	0.39	0/1830	0.64	0/2473
All	All	0.40	0/7252	0.66	3/9797 (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	205	GLY	N-CA-C	-7.15	95.22	113.10
1	C	205	GLY	N-CA-C	-7.15	95.22	113.10
1	A	234	LYS	N-CA-C	6.44	128.38	111.00

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	1765	0	1808	121	0
1	B	1796	0	1842	99	0
1	C	1762	0	1808	89	0
1	D	1796	0	1842	86	0
2	A	34	0	0	1	0
2	B	38	0	0	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	28	0	0	1	0
2	D	26	0	0	1	0
All	All	7245	0	7300	369	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 26.

The worst 5 of 369 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:C:44:ARG:HG3	1:C:56:ILE:CD1	1.54	1.37
1:C:44:ARG:CG	1:C:56:ILE:HD11	1.64	1.28
1:A:36:ILE:HD13	1:A:62:TYR:HB2	1.37	1.05
1:A:73:ILE:HB	1:A:107:ILE:HD12	1.44	1.00
1:B:36:ILE:HD13	1:B:62:TYR:HB2	1.44	0.98

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	217/242 (90%)	197 (91%)	17 (8%)	3 (1%)	11	20
1	B	223/242 (92%)	201 (90%)	18 (8%)	4 (2%)	8	14
1	C	217/242 (90%)	196 (90%)	19 (9%)	2 (1%)	17	31
1	D	223/242 (92%)	201 (90%)	20 (9%)	2 (1%)	17	31
All	All	880/968 (91%)	795 (90%)	74 (8%)	11 (1%)	12	21

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	59	LEU
1	B	82	GLY
1	C	101	GLY
1	C	200	ASN
1	A	56	ILE

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	196/207 (95%)	185 (94%)	11 (6%)	21	40
1	B	200/207 (97%)	183 (92%)	17 (8%)	10	21
1	C	196/207 (95%)	185 (94%)	11 (6%)	21	40
1	D	200/207 (97%)	185 (92%)	15 (8%)	13	26
All	All	792/828 (96%)	738 (93%)	54 (7%)	16	30

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	3	THR
1	C	77	LEU
1	D	212	THR
1	C	28	ILE
1	C	57	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	34	HIS
1	D	136	GLN
1	D	111	GLN
1	D	179	HIS
1	B	136	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

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($Å^2$)	Q<0.9
1	A	220/242 (90%)	0.52	21 (9%)	8	8	21, 44, 99, 126	0
1	B	224/242 (92%)	0.35	15 (6%)	17	18	15, 41, 84, 153	0
1	C	220/242 (90%)	0.64	27 (12%)	4	3	16, 45, 85, 123	0
1	D	224/242 (92%)	0.33	18 (8%)	12	12	16, 42, 82, 116	0
All	All	888/968 (91%)	0.46	81 (9%)	9	9	15, 43, 88, 153	0

The worst 5 of 81 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	67	HIS	9.9
1	A	226	ILE	9.4
1	A	176	TYR	7.1
1	B	66	LYS	6.4
1	A	59	LEU	6.0

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

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.