



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 24, 2016 – 11:35 AM EDT

PDB ID : 5JRQ  
Title : BRAFV600E Kinase Domain In Complex with Chemically Linked Vemurafenib Inhibitor VEM-6-VEM  
Authors : Grasso, M.J.; Marmorstein, R.  
Deposited on : 2016-05-06  
Resolution : 2.29 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027939  
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) : rb-20027939

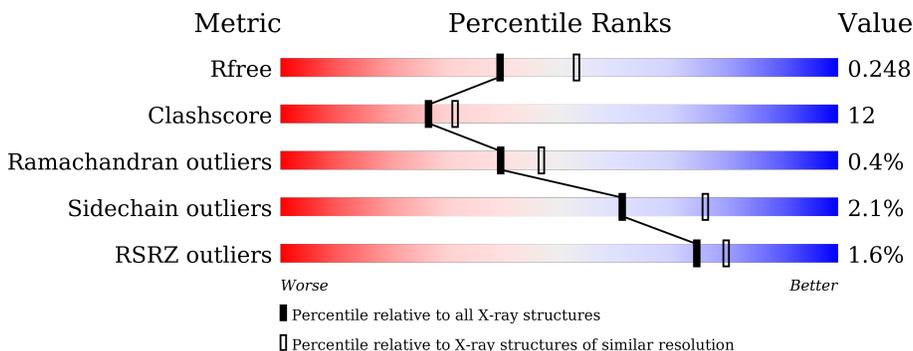
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.29 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	5193 (2.30-2.26)
Clashscore	102246	5929 (2.30-2.26)
Ramachandran outliers	100387	5851 (2.30-2.26)
Sidechain outliers	100360	5850 (2.30-2.26)
RSRZ outliers	91569	5204 (2.30-2.26)

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	280	
1	B	280	

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
3	DMS	A	802	-	-	-	X
4	GOL	A	803	-	-	-	X
5	TMA	A	804	-	-	-	X
5	TMA	B	803	-	-	-	X

## 2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 4177 atoms, of which 82 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 Serine/threonine-protein kinase B-raf.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	258	1954	1245	342	354	13	0	0	0
1	B	246	1874	1196	323	342	13	0	0	0

There are 42 discrepancies between the modelled and reference sequences:

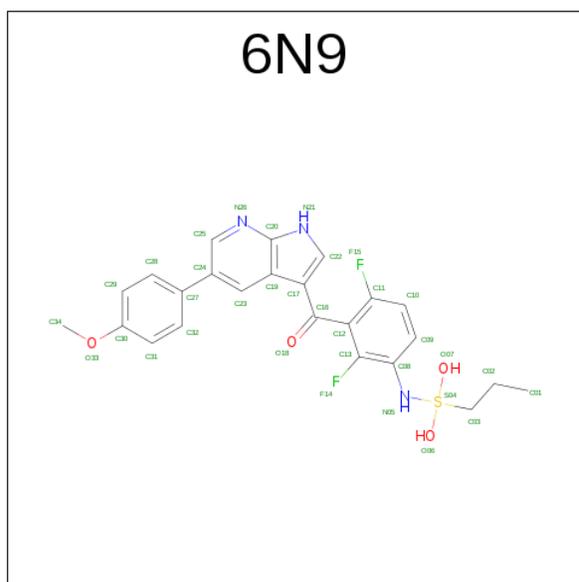
Chain	Residue	Modelled	Actual	Comment	Reference
A	444	GLY	-	expression tag	UNP P15056
A	445	SER	-	expression tag	UNP P15056
A	446	GLU	-	expression tag	UNP P15056
A	447	PHE	-	expression tag	UNP P15056
A	543	ALA	ILE	engineered mutation	UNP P15056
A	544	SER	ILE	engineered mutation	UNP P15056
A	551	LYS	ILE	engineered mutation	UNP P15056
A	562	ARG	GLN	engineered mutation	UNP P15056
A	588	ASN	LEU	engineered mutation	UNP P15056
A	600	GLU	VAL	engineered mutation	UNP P15056
A	630	SER	LYS	engineered mutation	UNP P15056
A	667	GLU	PHE	engineered mutation	UNP P15056
A	673	SER	TYR	engineered mutation	UNP P15056
A	688	ARG	ALA	engineered mutation	UNP P15056
A	706	SER	LEU	engineered mutation	UNP P15056
A	709	ARG	GLN	engineered mutation	UNP P15056
A	713	GLU	SER	engineered mutation	UNP P15056
A	716	GLU	LEU	engineered mutation	UNP P15056
A	720	GLU	SER	engineered mutation	UNP P15056
A	722	SER	PRO	engineered mutation	UNP P15056
A	723	GLY	LYS	engineered mutation	UNP P15056
B	444	GLY	-	expression tag	UNP P15056
B	445	SER	-	expression tag	UNP P15056
B	446	GLU	-	expression tag	UNP P15056
B	447	PHE	-	expression tag	UNP P15056

*Continued on next page...*

Continued from previous page...

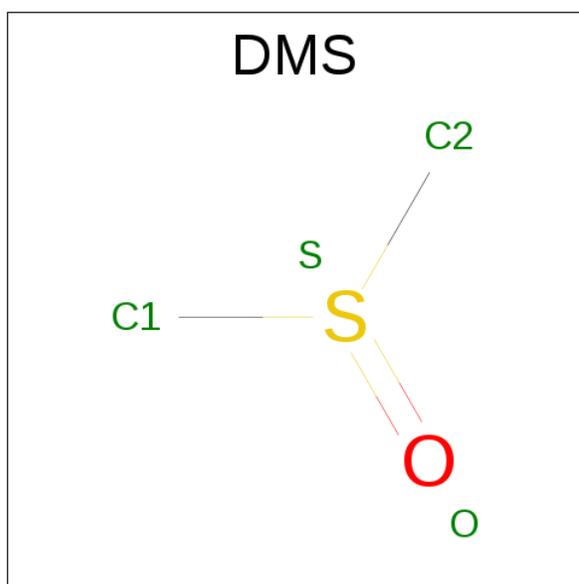
Chain	Residue	Modelled	Actual	Comment	Reference
B	543	ALA	ILE	engineered mutation	UNP P15056
B	544	SER	ILE	engineered mutation	UNP P15056
B	551	LYS	ILE	engineered mutation	UNP P15056
B	562	ARG	GLN	engineered mutation	UNP P15056
B	588	ASN	LEU	engineered mutation	UNP P15056
B	600	GLU	VAL	engineered mutation	UNP P15056
B	630	SER	LYS	engineered mutation	UNP P15056
B	667	GLU	PHE	engineered mutation	UNP P15056
B	673	SER	TYR	engineered mutation	UNP P15056
B	688	ARG	ALA	engineered mutation	UNP P15056
B	706	SER	LEU	engineered mutation	UNP P15056
B	709	ARG	GLN	engineered mutation	UNP P15056
B	713	GLU	SER	engineered mutation	UNP P15056
B	716	GLU	LEU	engineered mutation	UNP P15056
B	720	GLU	SER	engineered mutation	UNP P15056
B	722	SER	PRO	engineered mutation	UNP P15056
B	723	GLY	LYS	engineered mutation	UNP P15056

- Molecule 2 is (3-{[dihydroxy(propyl)-lambda 4 -sulfanyl]amino}-2,6-difluorophenyl)[5-(4-methoxyphenyl)-1H-pyrrolo[2,3-b]pyridin-3-yl]methanone (three-letter code: 6N9) (formula: C<sub>24</sub>H<sub>23</sub>F<sub>2</sub>N<sub>3</sub>O<sub>4</sub>S).



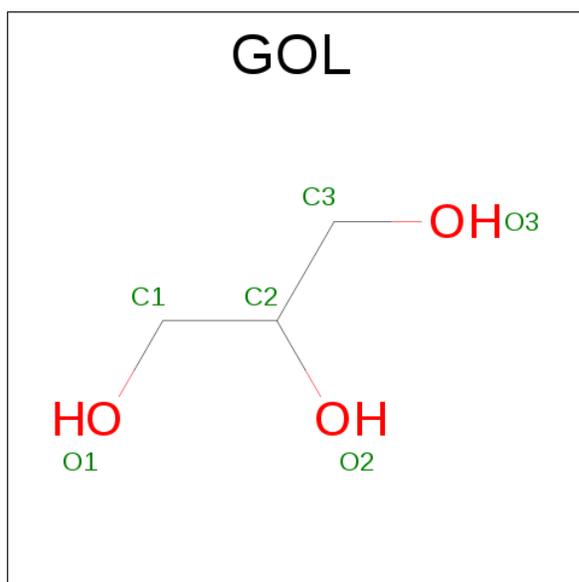
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	
			Total	C	F	H	N	O			S
2	A	1	Total	C	F	H	N	O	S	0	0
			53	24	2	19	3	4	1		
2	B	1	Total	C	F	H	N	O	S	0	0
			53	24	2	19	3	4	1		

- Molecule 3 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C<sub>2</sub>H<sub>6</sub>OS).



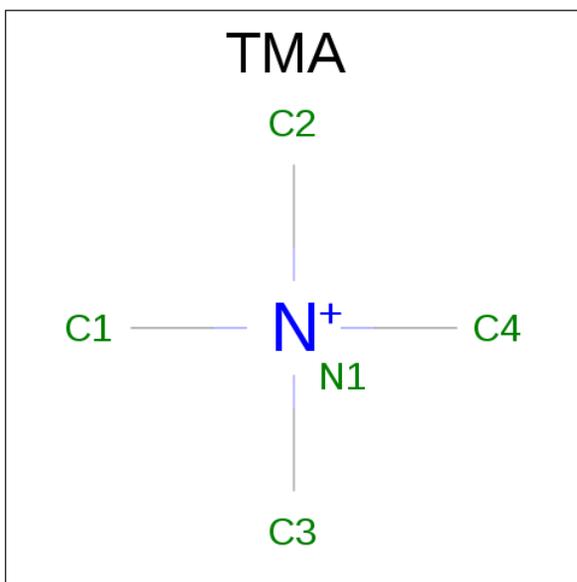
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	H	O			S
3	A	1	10	2	6	1	1	0	0
3	B	1	10	2	6	1	1	0	0

- Molecule 4 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
			Total	C	H	O		
4	A	1	14	3	8	3	0	0

- Molecule 5 is TETRAMETHYLAMMONIUM ION (three-letter code: TMA) (formula:  $C_4H_{12}N$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	H	N	0	0
			17	4	12	1		
5	B	1	Total	C	H	N	0	0
			17	4	12	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	92	Total	O	0	0
			92	92		
6	B	83	Total	O	0	0
			83	83		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	56.84Å 67.75Å 67.88Å 90.00° 90.32° 90.00°	Depositor
Resolution (Å)	29.21 – 2.29 29.21 – 2.29	Depositor EDS
% Data completeness (in resolution range)	95.3 (29.21-2.29) 87.9 (29.21-2.29)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.83 (at 2.29Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.195 , 0.254 0.190 , 0.248	Depositor DCC
$R_{free}$ test set	1878 reflections (8.99%)	DCC
Wilson B-factor (Å <sup>2</sup> )	30.6	Xtriage
Anisotropy	0.219	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 24.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.000 for -h,l,k 0.018 for -h,-l,-k 0.419 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4177	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.85% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, DMS, 6N9, TMA

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	0.40	0/1994	0.56	0/2701
1	B	0.41	0/1912	0.59	0/2587
All	All	0.41	0/3906	0.57	0/5288

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	1954	0	1863	41	0
1	B	1874	0	1792	52	0
2	A	34	19	0	0	0
2	B	34	19	0	1	0
3	A	4	6	6	0	0
3	B	4	6	6	0	0
4	A	6	8	8	0	0
5	A	5	12	12	3	0
5	B	5	12	12	3	0
6	A	92	0	0	5	0
6	B	83	0	0	2	0
All	All	4095	82	3699	93	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 12.

The worst 5 of 93 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:501:GLU:HG3	1:A:597:LEU:HD13	1.44	0.99
1:A:454:ASP:HA	1:A:457:ILE:HD11	1.47	0.95
1:B:501:GLU:HG3	1:B:597:LEU:HD13	1.48	0.93
1:B:509:ARG:HB2	5:B:803:TMA:H22	1.60	0.83
1:B:548:PHE:HB2	1:B:553:LEU:CD1	2.16	0.76

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	252/280 (90%)	237 (94%)	15 (6%)	0	100	100
1	B	236/280 (84%)	227 (96%)	7 (3%)	2 (1%)	24	26
All	All	488/560 (87%)	464 (95%)	22 (4%)	2 (0%)	39	47

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	448	ASP
1	B	599	THR

### 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	194/245 (79%)	192 (99%)	2 (1%)	82	91
1	B	188/245 (77%)	182 (97%)	6 (3%)	46	61
All	All	382/490 (78%)	374 (98%)	8 (2%)	61	76

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	493	GLN
1	B	715	GLU
1	B	580	ASN
1	B	454	ASP
1	B	573	ILE

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

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

7 ligands 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	6N9	A	801	-	32,37,37	1.89	7 (21%)	35,54,54	2.31	9 (25%)
3	DMS	A	802	-	3,3,3	0.69	0	3,3,3	0.62	0
4	GOL	A	803	-	5,5,5	0.33	0	5,5,5	0.26	0
5	TMA	A	804	-	4,4,4	1.21	0	6,6,6	0.56	0
2	6N9	B	801	-	32,37,37	2.04	6 (18%)	35,54,54	2.42	10 (28%)
3	DMS	B	802	-	3,3,3	0.72	0	3,3,3	0.49	0
5	TMA	B	803	-	4,4,4	1.19	0	6,6,6	0.32	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	6N9	A	801	-	-	0/15/23/23	0/4/4/4
3	DMS	A	802	-	-	0/0/0/0	0/0/0/0
4	GOL	A	803	-	-	0/4/4/4	0/0/0/0
5	TMA	A	804	-	-	0/0/0/0	0/0/0/0
2	6N9	B	801	-	-	0/15/23/23	0/4/4/4
3	DMS	B	802	-	-	0/0/0/0	0/0/0/0
5	TMA	B	803	-	-	0/0/0/0	0/0/0/0

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	801	6N9	C29-C30	-4.99	1.28	1.38
2	B	801	6N9	C29-C30	-4.46	1.29	1.38
2	A	801	6N9	C31-C30	-3.79	1.30	1.38
2	B	801	6N9	C31-C30	-3.72	1.31	1.38
2	B	801	6N9	C23-C19	-3.28	1.35	1.42

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	801	6N9	C13-C08-N05	-7.38	113.22	120.79
2	A	801	6N9	C13-C08-N05	-6.56	114.05	120.79
2	B	801	6N9	C24-C25-N26	-6.20	119.00	125.49
2	A	801	6N9	C24-C25-N26	-5.25	119.99	125.49
2	A	801	6N9	C32-C27-C28	-4.16	109.44	117.56

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	804	TMA	3	0
2	B	801	6N9	1	0
5	B	803	TMA	3	0

## 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	258/280 (92%)	-0.12	0 <b>100</b> <b>100</b>	13, 27, 43, 63	0
1	B	246/280 (87%)	-0.10	8 (3%) 50 58	16, 26, 40, 51	0
All	All	504/560 (90%)	-0.11	8 (1%) 74 80	13, 26, 42, 63	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	599	THR	3.7
1	B	626	ARG	3.2
1	B	455	GLY	2.6
1	B	447	PHE	2.4
1	B	598	ALA	2.3

### 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
5	TMA	A	804	5/5	0.92	0.30	8.22	29,35,41,41	0
3	DMS	A	802	4/4	0.89	0.22	2.88	27,32,39,61	0
5	TMA	B	803	5/5	0.90	0.22	2.67	25,32,39,39	0
4	GOL	A	803	6/6	0.91	0.15	2.49	23,37,45,48	0
3	DMS	B	802	4/4	0.95	0.16	1.70	29,36,59,61	0
2	6N9	B	801	34/34	0.95	0.12	-0.25	14,21,37,53	0
2	6N9	A	801	34/34	0.97	0.12	-0.39	14,21,31,46	0

## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.