WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 05:20 PM GMT

PDB ID	:	4HZB
Title	:	Crystal structure of the type VI SeMet effector-immunity complex Tae3-Tai3 from Ralstonia pickettii
Authors	:	Dong, C.; Zhang, H.; Gao, Z.Q.; Dong, Y.H.
Deposited on	:	2012-11-15
Resolution	:	2.60 Å(reported)

The following versions of software and data (see references (i)) 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 2.60 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.
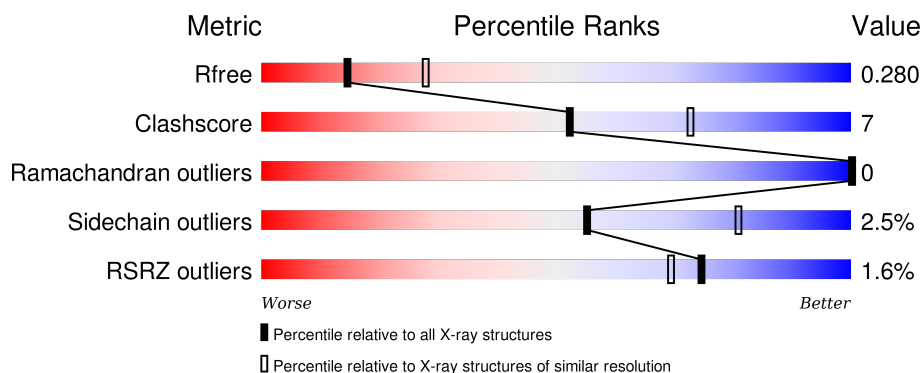
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (2.60-2.60)

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 <=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	119	% 77% 18% · ·
1	D	119	3% 79% 15% · 5%
2	B	150	3% 71% 10% 19%
2	C	150	% 65% 14% · 21%
2	E	150	% 72% 9% 19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	150	67%　　　　10%　·　23%

2   Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5816 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 cytoplasmic protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	115	Total 886	C 566	N 159	O 158	S 1	Se 2	0	0	0
1	D	113	Total 863	C 551	N 154	O 155	S 1	Se 2	0	0	0

- Molecule 2 is a protein called Putative periplasmic protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	122	Total 996	C 640	N 168	O 187	Se 1	0	0	0
2	C	119	Total 975	C 629	N 164	O 181	Se 1	0	0	0
2	E	121	Total 988	C 636	N 166	O 185	Se 1	0	0	0
2	F	116	Total 947	C 613	N 157	O 176	Se 1	0	0	0

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	2	MSE	-	EXPRESSION TAG	UNP C6BHF3
B	3	GLY	-	EXPRESSION TAG	UNP C6BHF3
B	4	SER	-	EXPRESSION TAG	UNP C6BHF3
B	5	SER	-	EXPRESSION TAG	UNP C6BHF3
B	6	HIS	-	EXPRESSION TAG	UNP C6BHF3
B	7	HIS	-	EXPRESSION TAG	UNP C6BHF3
B	8	HIS	-	EXPRESSION TAG	UNP C6BHF3
B	9	HIS	-	EXPRESSION TAG	UNP C6BHF3
B	10	HIS	-	EXPRESSION TAG	UNP C6BHF3
B	11	HIS	-	EXPRESSION TAG	UNP C6BHF3
B	12	GLU	-	EXPRESSION TAG	UNP C6BHF3
B	13	ASN	-	EXPRESSION TAG	UNP C6BHF3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	14	LEU	-	EXPRESSION TAG	UNP C6BHF3
B	15	TYR	-	EXPRESSION TAG	UNP C6BHF3
B	16	PHE	-	EXPRESSION TAG	UNP C6BHF3
B	17	GLN	-	EXPRESSION TAG	UNP C6BHF3
B	18	GLY	-	EXPRESSION TAG	UNP C6BHF3
B	19	HIS	-	EXPRESSION TAG	UNP C6BHF3
B	20	MSE	-	EXPRESSION TAG	UNP C6BHF3
B	21	GLY	-	EXPRESSION TAG	UNP C6BHF3
B	22	SER	-	EXPRESSION TAG	UNP C6BHF3
C	2	MSE	-	EXPRESSION TAG	UNP C6BHF3
C	3	GLY	-	EXPRESSION TAG	UNP C6BHF3
C	4	SER	-	EXPRESSION TAG	UNP C6BHF3
C	5	SER	-	EXPRESSION TAG	UNP C6BHF3
C	6	HIS	-	EXPRESSION TAG	UNP C6BHF3
C	7	HIS	-	EXPRESSION TAG	UNP C6BHF3
C	8	HIS	-	EXPRESSION TAG	UNP C6BHF3
C	9	HIS	-	EXPRESSION TAG	UNP C6BHF3
C	10	HIS	-	EXPRESSION TAG	UNP C6BHF3
C	11	HIS	-	EXPRESSION TAG	UNP C6BHF3
C	12	GLU	-	EXPRESSION TAG	UNP C6BHF3
C	13	ASN	-	EXPRESSION TAG	UNP C6BHF3
C	14	LEU	-	EXPRESSION TAG	UNP C6BHF3
C	15	TYR	-	EXPRESSION TAG	UNP C6BHF3
C	16	PHE	-	EXPRESSION TAG	UNP C6BHF3
C	17	GLN	-	EXPRESSION TAG	UNP C6BHF3
C	18	GLY	-	EXPRESSION TAG	UNP C6BHF3
C	19	HIS	-	EXPRESSION TAG	UNP C6BHF3
C	20	MSE	-	EXPRESSION TAG	UNP C6BHF3
C	21	GLY	-	EXPRESSION TAG	UNP C6BHF3
C	22	SER	-	EXPRESSION TAG	UNP C6BHF3
E	2	MSE	-	EXPRESSION TAG	UNP C6BHF3
E	3	GLY	-	EXPRESSION TAG	UNP C6BHF3
E	4	SER	-	EXPRESSION TAG	UNP C6BHF3
E	5	SER	-	EXPRESSION TAG	UNP C6BHF3
E	6	HIS	-	EXPRESSION TAG	UNP C6BHF3
E	7	HIS	-	EXPRESSION TAG	UNP C6BHF3
E	8	HIS	-	EXPRESSION TAG	UNP C6BHF3
E	9	HIS	-	EXPRESSION TAG	UNP C6BHF3
E	10	HIS	-	EXPRESSION TAG	UNP C6BHF3
E	11	HIS	-	EXPRESSION TAG	UNP C6BHF3
E	12	GLU	-	EXPRESSION TAG	UNP C6BHF3
E	13	ASN	-	EXPRESSION TAG	UNP C6BHF3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	14	LEU	-	EXPRESSION TAG	UNP C6BHF3
E	15	TYR	-	EXPRESSION TAG	UNP C6BHF3
E	16	PHE	-	EXPRESSION TAG	UNP C6BHF3
E	17	GLN	-	EXPRESSION TAG	UNP C6BHF3
E	18	GLY	-	EXPRESSION TAG	UNP C6BHF3
E	19	HIS	-	EXPRESSION TAG	UNP C6BHF3
E	20	MSE	-	EXPRESSION TAG	UNP C6BHF3
E	21	GLY	-	EXPRESSION TAG	UNP C6BHF3
E	22	SER	-	EXPRESSION TAG	UNP C6BHF3
F	2	MSE	-	EXPRESSION TAG	UNP C6BHF3
F	3	GLY	-	EXPRESSION TAG	UNP C6BHF3
F	4	SER	-	EXPRESSION TAG	UNP C6BHF3
F	5	SER	-	EXPRESSION TAG	UNP C6BHF3
F	6	HIS	-	EXPRESSION TAG	UNP C6BHF3
F	7	HIS	-	EXPRESSION TAG	UNP C6BHF3
F	8	HIS	-	EXPRESSION TAG	UNP C6BHF3
F	9	HIS	-	EXPRESSION TAG	UNP C6BHF3
F	10	HIS	-	EXPRESSION TAG	UNP C6BHF3
F	11	HIS	-	EXPRESSION TAG	UNP C6BHF3
F	12	GLU	-	EXPRESSION TAG	UNP C6BHF3
F	13	ASN	-	EXPRESSION TAG	UNP C6BHF3
F	14	LEU	-	EXPRESSION TAG	UNP C6BHF3
F	15	TYR	-	EXPRESSION TAG	UNP C6BHF3
F	16	PHE	-	EXPRESSION TAG	UNP C6BHF3
F	17	GLN	-	EXPRESSION TAG	UNP C6BHF3
F	18	GLY	-	EXPRESSION TAG	UNP C6BHF3
F	19	HIS	-	EXPRESSION TAG	UNP C6BHF3
F	20	MSE	-	EXPRESSION TAG	UNP C6BHF3
F	21	GLY	-	EXPRESSION TAG	UNP C6BHF3
F	22	SER	-	EXPRESSION TAG	UNP C6BHF3

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	28	Total 28	O 28	0	0
3	B	29	Total 29	O 29	0	0
3	C	22	Total 22	O 22	0	0
3	D	25	Total 25	O 25	0	0

Continued on next page...

Continued from previous page...

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	26	Total 26	O 26	0	0
3	F	31	Total 31	O 31	0	0

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 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.
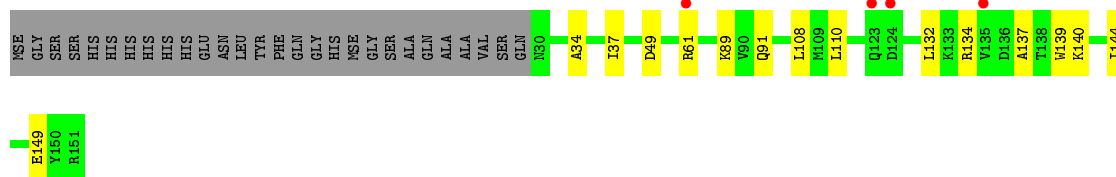
- Molecule 1: Putative cytoplasmic protein

- Molecule 1: Putative cytoplasmic protein
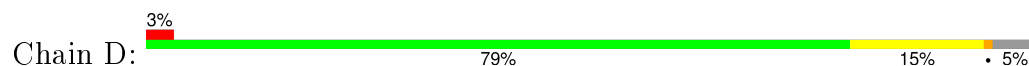
- Molecule 2: Putative periplasmic protein

- Molecule 2: Putative periplasmic protein

- Molecule 2: Putative periplasmic protein

MSE GLY SER SER HIS HIS HIS HIS HIS HIS GLU ASN LEU TYR PHE GLY GLY HIS MSE GLY SER ALA GLN ALA ALA VAL SER GLN ASN S31 D47 T51 E56 A64 V68 Q91 S95 L108 S119 V126 H127 I142 E149 Y150 R151

• Molecule 2: Putative periplasmic protein

Chain F:

67%   10% •   23%

MSE GLY SER SER HIS HIS HIS HIS HIS HIS GLU ASN LEU TYR PHE GLY GLY HIS MSE GLY SER ALA GLN ALA ALA VAL SER GLN ASN S31 F39 D47 SER ASP GLN T51 E56 P57 Y62 A78 N82 G83 V84 Y93 V104 L108 P117 V118 S119

K133 ARG VAL D136 A137 T138 R151
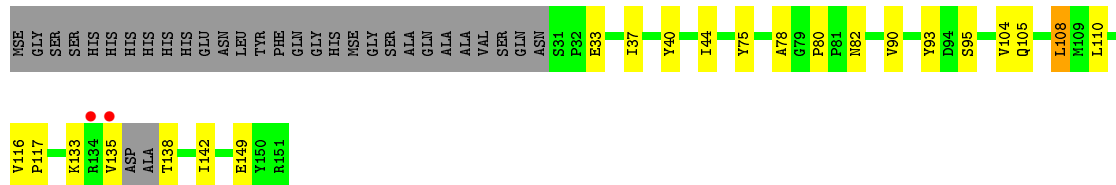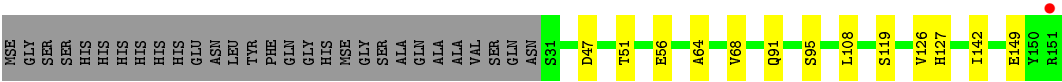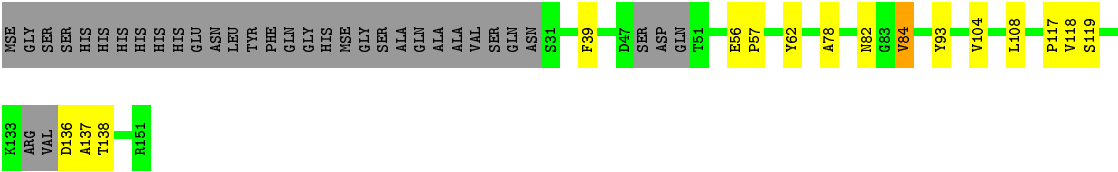
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	56.20Å 126.94Å 143.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.05 – 2.60 35.05 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.3 (35.05-2.60) 99.1 (35.05-2.60)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.84 (at 2.61Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, R_{free}	0.216 , 0.271 0.232 , 0.280	Depositor DCC
R_{free} test set	1996 reflections (6.61%)	DCC
Wilson B-factor (Å^2)	42.1	Xtriage
Anisotropy	0.550	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 32.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< |L| > = 0.46, < L^2 > = 0.29$	Xtriage
Outliers	1 of 32205 reflections (0.003%)	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	5816	wwPDB-VP
Average B, all atoms (Å^2)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 49.85 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 7.0390e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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)

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	0.47	0/914	0.60	0/1237
1	D	0.45	0/890	0.60	0/1205
2	B	0.47	0/1025	0.59	0/1401
2	C	0.45	0/1003	0.57	0/1369
2	E	0.47	0/1017	0.59	0/1390
2	F	0.45	0/974	0.59	0/1329
All	All	0.46	0/5823	0.59	0/7931

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	886	0	850	22	0
1	D	863	0	828	19	0
2	B	996	0	946	11	0
2	C	975	0	930	16	0
2	E	988	0	940	10	0
2	F	947	0	899	11	0
3	A	28	0	0	0	0
3	B	29	0	0	1	0
3	C	22	0	0	0	0
3	D	25	0	0	0	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	E	26	0	0	0	0
3	F	31	0	0	0	0
All	All	5816	0	5393	74	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 7.

The worst 5 of 74 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:D:70:VAL:HG22	1:D:82:MSE:SE	1.64	1.46
1:A:47:TYR:CB	1:A:82:MSE:HE1	1.64	1.26
1:A:64:MSE:SE	1:A:117:ARG:NH1	2.28	1.15
1:D:70:VAL:HG22	1:D:82:MSE:CE	1.87	1.04
1:A:47:TYR:HB3	1:A:82:MSE:HE1	1.37	1.03

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	113/119 (95%)	110 (97%)	3 (3%)	0	100	100
1	D	111/119 (93%)	101 (91%)	10 (9%)	0	100	100
2	B	120/150 (80%)	115 (96%)	5 (4%)	0	100	100
2	C	115/150 (77%)	112 (97%)	3 (3%)	0	100	100
2	E	119/150 (79%)	113 (95%)	6 (5%)	0	100	100
2	F	110/150 (73%)	106 (96%)	4 (4%)	0	100	100
All	All	688/838 (82%)	657 (96%)	31 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	87/88 (99%)	83 (95%)	4 (5%)	33	61
1	D	85/88 (97%)	82 (96%)	3 (4%)	43	71
2	B	107/126 (85%)	107 (100%)	0	100	100
2	C	105/126 (83%)	103 (98%)	2 (2%)	65	86
2	E	106/126 (84%)	104 (98%)	2 (2%)	65	86
2	F	101/126 (80%)	97 (96%)	4 (4%)	38	67
All	All	591/680 (87%)	576 (98%)	15 (2%)	55	81

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	5	ILE
1	D	32	GLU
2	F	93	TYR
2	C	108	LEU
2	F	84	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	18	ASN
2	E	91	GLN
2	E	103	GLN

5.3.3 RNA ⓘ

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)

There are no ligands in this entry.

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	113/119 (94%)	-0.16	1 (0%)	85	83	11, 27, 46, 53	0
1	D	111/119 (93%)	0.04	3 (2%)	58	51	14, 30, 50, 61	0
2	B	121/150 (80%)	-0.20	4 (3%)	50	43	9, 22, 50, 62	0
2	C	118/150 (78%)	-0.14	2 (1%)	73	68	13, 30, 51, 65	0
2	E	120/150 (80%)	-0.16	1 (0%)	87	85	11, 25, 47, 62	0
2	F	115/150 (76%)	-0.13	0	100	100	12, 29, 51, 71	0
All	All	698/838 (83%)	-0.13	11 (1%)	74	69	9, 27, 50, 71	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	135	VAL	2.8
1	D	31	ILE	2.6
2	C	135	VAL	2.5
1	A	116	TYR	2.5
2	B	124	ASP	2.4

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.