



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 07:33 AM GMT

PDB ID : 2HWJ  
Title : Crystal structure of protein Atu1540 from Agrobacterium tumefaciens  
Authors : Chang, C.; Xu, X.; Savchenko, A.; Edwards, A.M.; Joachimiak, A.; Midwest  
Center for Structural Genomics (MCSG)  
Deposited on : 2006-08-01  
Resolution : 2.61 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

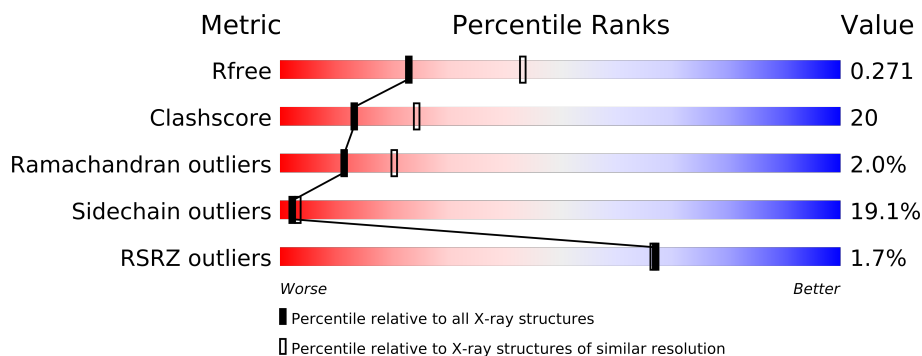
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.61 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))
$R_{free}$	66092	1940 (2.64-2.60)
Clashscore	79885	2404 (2.64-2.60)
Ramachandran outliers	78287	2360 (2.64-2.60)
Sidechain outliers	78261	2360 (2.64-2.60)
RSRZ outliers	66119	1939 (2.64-2.60)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	205	
1	B	205	
1	C	205	
1	D	205	
1	E	205	
1	F	205	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 9380 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Hypothetical protein Atu1540.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	196	Total	C	N	O	S	Se	0	0	0
			1574	1000	289	281	1	3			
1	B	189	Total	C	N	O	S	Se	0	1	0
			1530	972	284	270	1	3			
1	C	195	Total	C	N	O	S	Se	0	2	0
			1579	1005	290	280	1	3			
1	D	193	Total	C	N	O	S	Se	0	1	0
			1555	987	286	278	1	3			
1	E	188	Total	C	N	O	S	Se	0	0	0
			1519	965	281	269	1	3			
1	F	190	Total	C	N	O	S	Se	0	0	0
			1528	970	283	271	1	3			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
A	103	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
A	145	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
A	186	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
B	103	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
B	145	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
B	186	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
C	1	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
C	103	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
C	145	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
C	186	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
D	1	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
D	103	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
D	145	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
D	186	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
E	1	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	103	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
E	145	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
E	186	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
F	1	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
F	103	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
F	145	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59
F	186	MSE	MET	MODIFIED RESIDUE	UNP Q8UF59

- Molecule 2 is water.

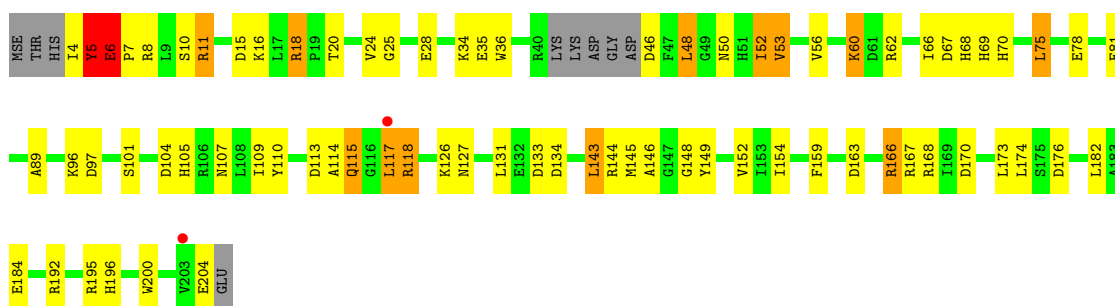
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	17	Total O 17 17	0	0
2	B	18	Total O 18 18	0	0
2	C	22	Total O 22 22	0	0
2	D	15	Total O 15 15	0	0
2	E	13	Total O 13 13	0	0
2	F	10	Total O 10 10	0	0

### 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.

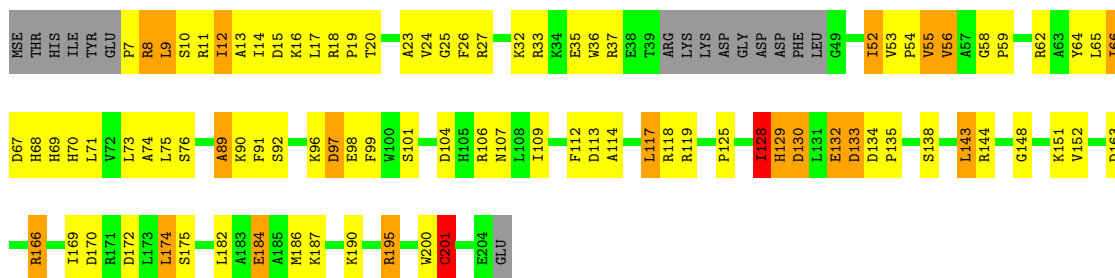
- Molecule 1: Hypothetical protein Atu1540

Chain A: 



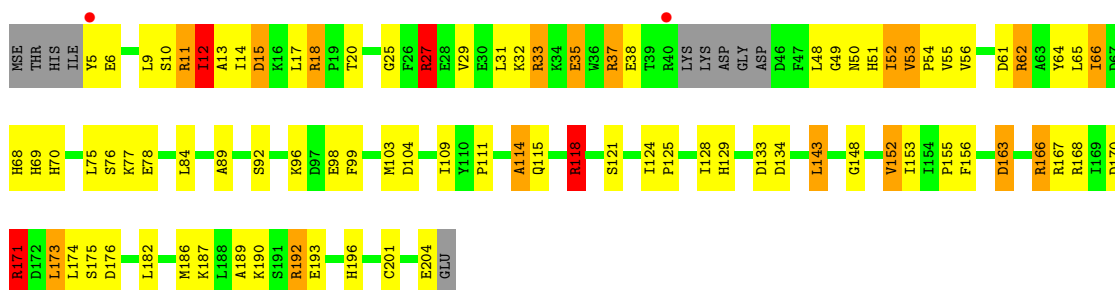
- Molecule 1: Hypothetical protein Atu1540

Chain B: 



- Molecule 1: Hypothetical protein Atu1540

Chain C: 



- Molecule 1: Hypothetical protein Atu1540

I153	I154	P155	F156	S157	D163	R166	R167	R168	I169	D170	R171	D172	L173	L174	S175	D176	S177	D180	A181	L182	K190	S191	R192	R195	I196	L197	C201	G202	V203	E204	GLU	MSE	THR	HR	ILR	TVR	E6	L9	S10	I11	I12	A13	I14	D15	R16	L17	R18	P19	T20	G25	P26	R27	E28	V29	R33	K34	R37	E38	T39	R40	LVS	LVS	ASP	GLY	ASP	ASP	F47	L48	G49	N50	H51	I52	V53	P54	V55	P59	K60	L65	I66	D67	H68	H69	H70	L71	H72
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	----	-----	-----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- Chain E: 

L169	D170	R171	D172	L173	L174	S177	D180	A181	L182	K187	K190	E193	A194	R195	H196	L197	C201	G202	V203	E204	GLU	MSE	THR	HIS	ILE	TVR	GLU	P7	R8	L9	S10	R11	I12	A13	I14	D15	R18	P19	T20	V24	R27	E28	R33	K34	E35	V36	R37	E38	T39	ARG	LYS	LYS	ASP	GLY	ASP	ASP	PHE	LEU	LEU	GLY	N50	H51	I52	V53	P54	V55	P59	K60	D61	R62	L65	I66	D67	H68	H69	L70
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- Chain F: 

I169	L173	L174	F178	D179	D180	A181	L182	A183	A184	A185	M186	K187	R182	L187	P198	C201	G202	V202	E204	GLU	MSE	THR	HIS	ILE	TVR	P7	R8	L9	S10	R11	I12	A13	I14	K16	L17	R18	P19	T20	Q21	I22	G25	F26	R27	E28	V29	E30	L31	K32	R33	K36	R37	E38	T39	ARG	LYS	LYS	ASP	GLY	ASP	ASP	PHE	LEU	G49	N50	H51	F52	V53	P54	V55	Y64	D67	H68	H69	H70
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	132.08Å 173.75Å 142.08Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.61 49.69 – 2.61	Depositor EDS
% Data completeness (in resolution range)	97.8 (50.00-2.61) 97.9 (49.69-2.61)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.78 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.198 , 0.273 0.197 , 0.271	Depositor DCC
$R_{free}$ test set	2486 reflections (5.34%)	DCC
Wilson B-factor (Å <sup>2</sup> )	67.6	Xtriage
Anisotropy	0.464	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 74.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 49011 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9380	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.37% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 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	1.26	2/1608 (0.1%)	1.28	13/2164 (0.6%)
1	B	1.26	7/1567 (0.4%)	1.20	9/2107 (0.4%)
1	C	1.41	6/1620 (0.4%)	1.30	14/2180 (0.6%)
1	D	1.30	5/1591 (0.3%)	1.07	3/2141 (0.1%)
1	E	1.24	5/1552 (0.3%)	1.17	5/2087 (0.2%)
1	F	1.15	10/1561 (0.6%)	1.03	3/2100 (0.1%)
All	All	1.27	35/9499 (0.4%)	1.18	47/12779 (0.4%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	3
1	C	0	1
1	E	0	1
1	F	0	1
All	All	0	6

The worst 5 of 35 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	175	SER	CB-OG	23.15	1.72	1.42
1	F	179	ASP	CG-OD1	14.28	1.58	1.25
1	D	40	ARG	C-O	12.39	1.46	1.23
1	E	201	CYS	CB-SG	-10.64	1.64	1.82
1	F	126	LYS	CE-NZ	9.47	1.72	1.49

The worst 5 of 47 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	179	ASP	CB-CG-OD1	-9.58	109.67	118.30

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	176	ASP	CB-CG-OD1	9.15	126.54	118.30
1	A	118	ARG	NE-CZ-NH1	8.63	124.61	120.30
1	B	133	ASP	CB-CG-OD2	-8.20	110.92	118.30
1	B	16	LYS	CD-CE-NZ	-7.85	93.65	111.70

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	200	TRP	Peptide
1	A	4	ILE	Peptide
1	A	6	GLU	Peptide
1	C	12	ILE	Peptide
1	E	156	PHE	Peptide

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1574	0	1549	60	0
1	B	1530	0	1521	68	0
1	C	1579	0	1558	65	0
1	D	1555	0	1538	91	0
1	E	1519	0	1512	63	0
1	F	1528	0	1516	63	0
2	A	17	0	0	2	0
2	B	18	0	0	1	0
2	C	22	0	0	0	0
2	D	15	0	0	2	0
2	E	13	0	0	2	0
2	F	10	0	0	2	0
All	All	9380	0	9194	378	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 20.

The worst 5 of 378 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:F:126:LYS:CE	1:F:126:LYS:NZ	1.72	1.50
1:A:5:TYR:CB	1:C:32:LYS:HE2	1.43	1.44
1:D:175:SER:CB	1:D:175:SER:OG	1.72	1.36
1:A:11:ARG:HG2	1:A:11:ARG:HH11	1.05	1.09
1:F:128:ILE:HD13	1:F:128:ILE:H	1.11	1.09

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	192/205 (94%)	179 (93%)	11 (6%)	2 (1%)	22	44
1	B	186/205 (91%)	170 (91%)	14 (8%)	2 (1%)	21	41
1	C	193/205 (94%)	181 (94%)	8 (4%)	4 (2%)	11	19
1	D	190/205 (93%)	166 (87%)	17 (9%)	7 (4%)	5	7
1	E	184/205 (90%)	162 (88%)	16 (9%)	6 (3%)	6	9
1	F	186/205 (91%)	162 (87%)	22 (12%)	2 (1%)	21	41
All	All	1131/1230 (92%)	1020 (90%)	88 (8%)	23 (2%)	11	20

5 of 23 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	5	TYR
1	A	6	GLU
1	D	19	PRO
1	D	48	LEU
1	D	177	SER

### 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	163/170 (96%)	138 (85%)	25 (15%)	4	6
1	B	160/170 (94%)	131 (82%)	29 (18%)	2	4
1	C	164/170 (96%)	132 (80%)	32 (20%)	2	3
1	D	162/170 (95%)	129 (80%)	33 (20%)	2	3
1	E	159/170 (94%)	122 (77%)	37 (23%)	1	2
1	F	159/170 (94%)	131 (82%)	28 (18%)	3	4
All	All	967/1020 (95%)	783 (81%)	184 (19%)	2	3

5 of 184 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	187	LYS
1	D	101	SER
1	F	75	LEU
1	C	193	GLU
1	D	37	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	196	HIS
1	D	82	HIS
1	F	105	HIS
1	D	68	HIS
1	D	69	HIS

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

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, 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	196/205 (95%)	-0.22	2 (1%) 79 80	48, 60, 76, 82	0
1	B	189/205 (92%)	-0.12	0 100 100	52, 71, 84, 112	0
1	C	195/205 (95%)	-0.24	2 (1%) 79 80	40, 59, 81, 94	0
1	D	193/205 (94%)	-0.08	3 (1%) 68 69	56, 71, 85, 93	0
1	E	188/205 (91%)	-0.26	2 (1%) 77 78	49, 65, 78, 96	0
1	F	190/205 (92%)	0.12	10 (5%) 25 22	57, 77, 88, 94	0
All	All	1151/1230 (93%)	-0.13	19 (1%) 67 66	40, 68, 84, 112	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	8	ARG	4.6
1	F	178	PHE	4.5
1	C	5	TYR	3.7
1	F	9	LEU	3.5
1	E	36	TRP	3.2

### 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 carbohydrates in this entry.

### 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.