



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 07:55 am BST

PDB ID : 2WV5  
Title : Crystal structure of foot-and-mouth disease virus 3C protease in complex with a decameric peptide corresponding to the VP1-2A cleavage junction with a GLN to Glu substitution at P1  
Authors : Zunszain, P.A.; Knox, S.R.; Sweeney, T.R.; Yang, J.; Roque-Rosell, N.; Belsham, G.J.; Leatherbarrow, R.J.; Curry, S.  
Deposited on : 2009-10-13  
Resolution : 2.70 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

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
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)  
Validation Pipeline (wwPDB-VP) : 2.11

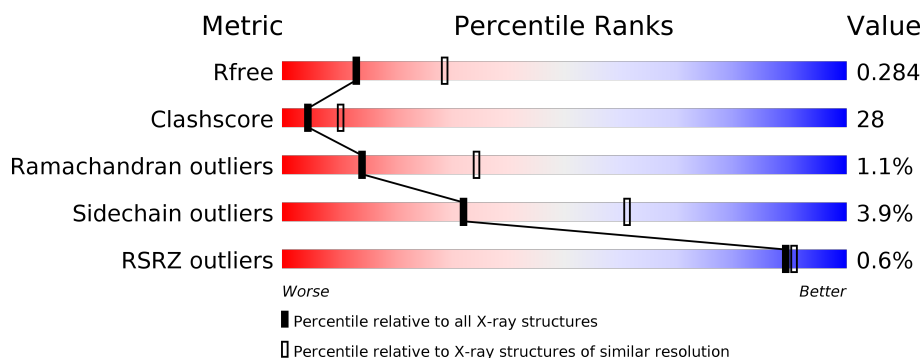
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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.




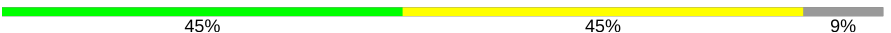
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 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  $\leq 5\%$ . 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	214	<div> <div>53%</div> <div>37%</div> <div>• 7%</div> </div>
1	B	214	<div> <div>57%</div> <div>36%</div> <div>7%</div> </div>
1	C	214	<div> <div>%</div> <div>53%</div> <div>38%</div> <div>• 7%</div> </div>
1	D	214	<div> <div>59%</div> <div>31%</div> <div>• 7%</div> </div>
2	E	11	<div> <div>55%</div> <div>36%</div> <div>9%</div> </div>
2	F	11	<div> <div>55%</div> <div>36%</div> <div>9%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	G	11	 55% 36% 9%
2	H	11	 45% 45% 9%

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6096 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 PICORNAIN 3C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	200	Total	C	N	O	S	0	0	0
			1435	907	241	273	14			
1	B	200	Total	C	N	O	S	0	0	0
			1440	911	242	273	14			
1	C	200	Total	C	N	O	S	0	0	0
			1435	907	241	273	14			
1	D	200	Total	C	N	O	S	0	0	0
			1442	912	243	273	14			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	22	ILE	ASN	SEE REMARK 999	UNP P03306
A	88	MET	ILE	SEE REMARK 999	UNP P03306
A	95	LYS	CYS	engineered mutation	UNP P03306
A	142	LEU	CYS	engineered mutation	UNP P03306
A	163	ALA	CYS	engineered mutation	UNP P03306
B	22	ILE	ASN	SEE REMARK 999	UNP P03306
B	88	MET	ILE	SEE REMARK 999	UNP P03306
B	95	LYS	CYS	engineered mutation	UNP P03306
B	142	LEU	CYS	engineered mutation	UNP P03306
B	163	ALA	CYS	engineered mutation	UNP P03306
C	22	ILE	ASN	SEE REMARK 999	UNP P03306
C	88	MET	ILE	SEE REMARK 999	UNP P03306
C	95	LYS	CYS	engineered mutation	UNP P03306
C	142	LEU	CYS	engineered mutation	UNP P03306
C	163	ALA	CYS	engineered mutation	UNP P03306
D	22	ILE	ASN	SEE REMARK 999	UNP P03306
D	88	MET	ILE	SEE REMARK 999	UNP P03306
D	95	LYS	CYS	engineered mutation	UNP P03306
D	142	LEU	CYS	engineered mutation	UNP P03306
D	163	ALA	CYS	engineered mutation	UNP P03306

- Molecule 2 is a protein called FOOT AND MOUTH DISEASE VIRUS (SEROTYPE A) VARIANT VP1 CAPSID PROTEIN.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	E	10	Total	C	N	O	0	0	0
			73	49	11	13			
2	F	10	Total	C	N	O	0	0	0
			73	49	11	13			
2	G	10	Total	C	N	O	0	0	0
			73	49	11	13			
2	H	10	Total	C	N	O	0	0	0
			73	49	11	13			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	6	GLU	GLN	engineered mutation	UNP Q65050
F	6	GLU	GLN	engineered mutation	UNP Q65050
G	6	GLU	GLN	engineered mutation	UNP Q65050
H	6	GLU	GLN	engineered mutation	UNP Q65050

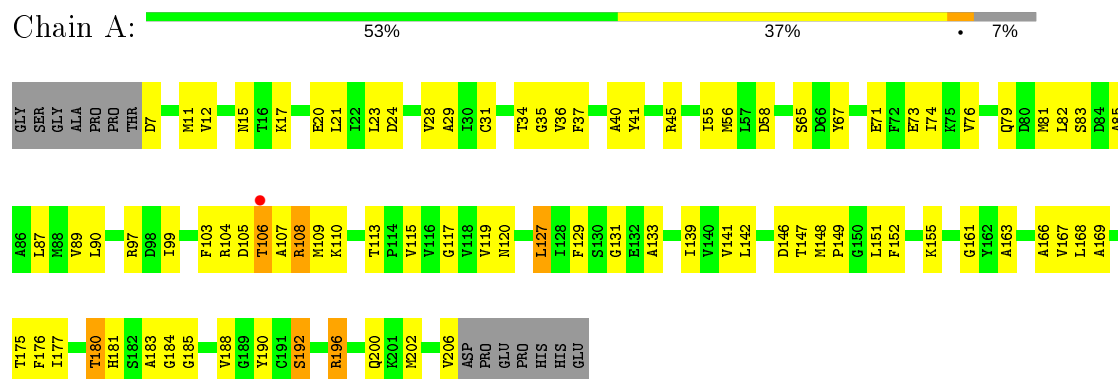
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	18	Total	O	0	0
			18	18		
3	B	7	Total	O	0	0
			7	7		
3	C	12	Total	O	0	0
			12	12		
3	D	9	Total	O	0	0
			9	9		
3	E	2	Total	O	0	0
			2	2		
3	F	1	Total	O	0	0
			1	1		
3	G	2	Total	O	0	0
			2	2		
3	H	1	Total	O	0	0
			1	1		

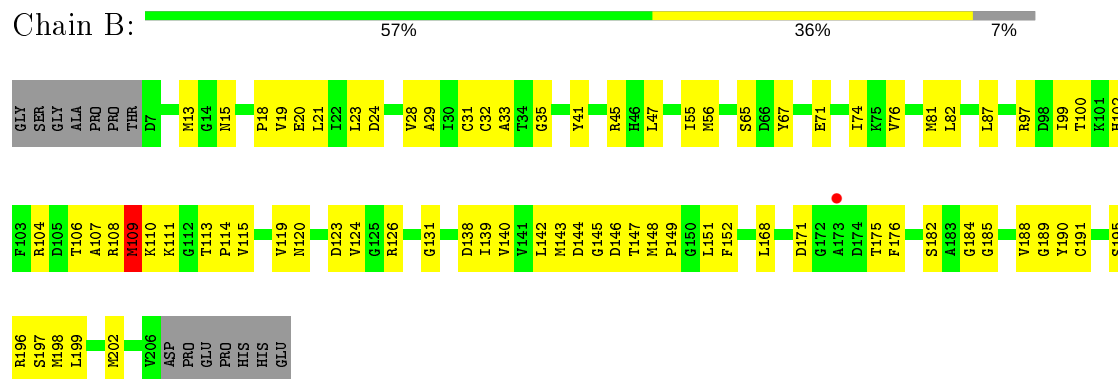
### 3 Residue-property plots

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.

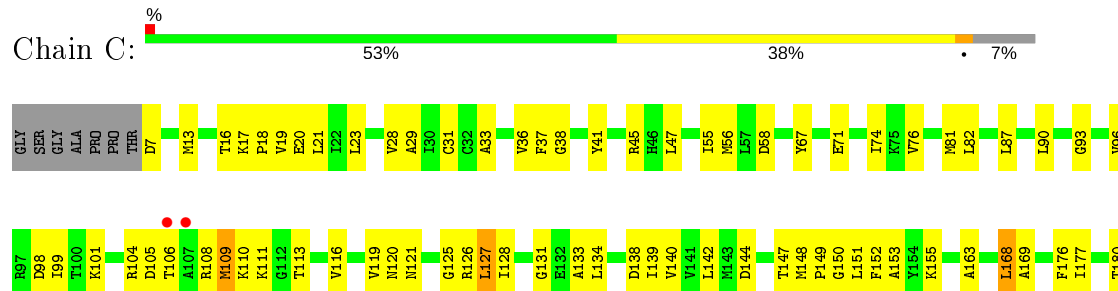
#### • Molecule 1: PICORNAIN 3C



#### • Molecule 1: PICORNAIN 3C



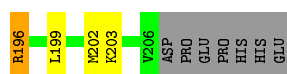
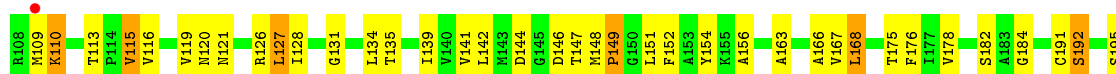
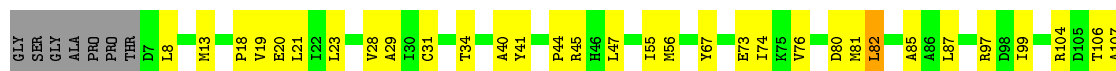
#### • Molecule 1: PICORNAIN 3C





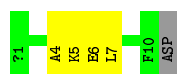
• Molecule 1: PICORNAIN 3C

Chain D: 59% 31% 7%



• Molecule 2: FOOT AND MOUTH DISEASE VIRUS (SEROTYPE A) VARIANT VP1 CAPSID PROTEIN

Chain E: 55% 36% 9%



• Molecule 2: FOOT AND MOUTH DISEASE VIRUS (SEROTYPE A) VARIANT VP1 CAPSID PROTEIN

Chain F: 55% 36% 9%



• Molecule 2: FOOT AND MOUTH DISEASE VIRUS (SEROTYPE A) VARIANT VP1 CAPSID PROTEIN

Chain G: 55% 36% 9%



• Molecule 2: FOOT AND MOUTH DISEASE VIRUS (SEROTYPE A) VARIANT VP1 CAPSID PROTEIN

Chain H: 45% 45% 9%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	64.03Å 75.11Å 86.30Å 90.00° 99.26° 90.00°	Depositor
Resolution (Å)	56.33 – 2.70 56.33 – 2.70	Depositor EDS
% Data completeness (in resolution range)	75.8 (56.33-2.70) 75.9 (56.33-2.70)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.14 (at 2.69Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.236 , 0.290 0.233 , 0.284	Depositor DCC
$R_{free}$ test set	831 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.9	Xtriage
Anisotropy	0.609	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 60.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	6096	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	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 47.57 % 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 9.7442e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: ACE

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	0.41	0/1458	0.65	0/1979
1	B	0.41	0/1463	0.65	0/1984
1	C	0.39	0/1458	0.64	0/1979
1	D	0.39	0/1465	0.64	0/1986
2	E	0.41	0/72	0.84	0/97
2	F	0.39	0/72	0.78	0/97
2	G	0.44	0/72	0.79	0/97
2	H	0.45	0/72	0.83	0/97
All	All	0.40	0/6132	0.65	0/8316

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	1435	0	1368	89	0
1	B	1440	0	1381	90	0
1	C	1435	0	1368	87	0
1	D	1442	0	1388	78	0
2	E	73	0	76	11	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	73	0	76	6	0
2	G	73	0	76	5	0
2	H	73	0	76	7	0
3	A	18	0	0	0	0
3	B	7	0	0	0	0
3	C	12	0	0	1	0
3	D	9	0	0	0	0
3	E	2	0	0	0	0
3	F	1	0	0	0	0
3	G	2	0	0	0	0
3	H	1	0	0	2	0
All	All	6096	0	5809	337	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 28.

The worst 5 of 337 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:45:ARG:HH21	1:A:82:LEU:HB3	1.17	1.06
1:C:45:ARG:HH21	1:C:82:LEU:HB3	1.24	1.01
1:B:45:ARG:HH21	1:B:82:LEU:HB3	1.25	0.97
1:D:74:ILE:HD11	1:D:149:PRO:HG2	1.50	0.94
1:B:13:MET:HG3	1:B:126:ARG:HH11	1.34	0.93

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	198/214 (92%)	179 (90%)	18 (9%)	1 (0%)	29 54

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	198/214 (92%)	178 (90%)	17 (9%)	3 (2%)	10	26
1	C	198/214 (92%)	178 (90%)	17 (9%)	3 (2%)	10	26
1	D	198/214 (92%)	181 (91%)	15 (8%)	2 (1%)	15	37
2	E	8/11 (73%)	7 (88%)	1 (12%)	0	100	100
2	F	8/11 (73%)	7 (88%)	1 (12%)	0	100	100
2	G	8/11 (73%)	7 (88%)	1 (12%)	0	100	100
2	H	8/11 (73%)	7 (88%)	1 (12%)	0	100	100
All	All	824/900 (92%)	744 (90%)	71 (9%)	9 (1%)	14	34

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	109	MET
1	C	186	ASN
1	D	110	LYS
1	B	111	LYS
1	B	107	ALA

### 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	140/169 (83%)	133 (95%)	7 (5%)	24	51
1	B	141/169 (83%)	137 (97%)	4 (3%)	43	73
1	C	140/169 (83%)	134 (96%)	6 (4%)	29	57
1	D	142/169 (84%)	136 (96%)	6 (4%)	30	58
2	E	7/8 (88%)	7 (100%)	0	100	100
2	F	7/8 (88%)	7 (100%)	0	100	100
2	G	7/8 (88%)	7 (100%)	0	100	100
2	H	7/8 (88%)	7 (100%)	0	100	100
All	All	591/708 (84%)	568 (96%)	23 (4%)	32	61

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	196	ARG
1	C	127	LEU
1	D	192	SER
1	C	109	MET
1	C	168	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	91	HIS
1	D	120	ASN
1	C	200	GLN
1	B	91	HIS
1	D	91	HIS

### 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](#)

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 ⓘ

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	200/214 (93%)	-0.12	1 (0%) 91 92	11, 31, 51, 60	0
1	B	200/214 (93%)	-0.12	1 (0%) 91 92	16, 32, 53, 60	0
1	C	200/214 (93%)	-0.18	2 (1%) 82 83	11, 30, 47, 55	0
1	D	200/214 (93%)	-0.17	1 (0%) 91 92	15, 29, 51, 63	0
2	E	9/11 (81%)	-0.16	0 100 100	28, 32, 38, 38	0
2	F	9/11 (81%)	-0.25	0 100 100	29, 32, 40, 42	0
2	G	9/11 (81%)	-0.21	0 100 100	31, 34, 37, 40	0
2	H	9/11 (81%)	-0.18	0 100 100	26, 31, 37, 42	0
All	All	836/900 (92%)	-0.15	5 (0%) 89 91	11, 31, 51, 63	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	106	THR	3.6
1	C	106	THR	3.1
1	C	107	ALA	2.6
1	D	109	MET	2.6
1	B	173	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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.