



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2024 – 08:51 PM EST

PDB ID : 3N7N  
Title : Structure of Csm1/Lrs4 complex  
Authors : Corbett, K.D.; Harrison, S.C.  
Deposited on : 2010-05-27  
Resolution : 3.90 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

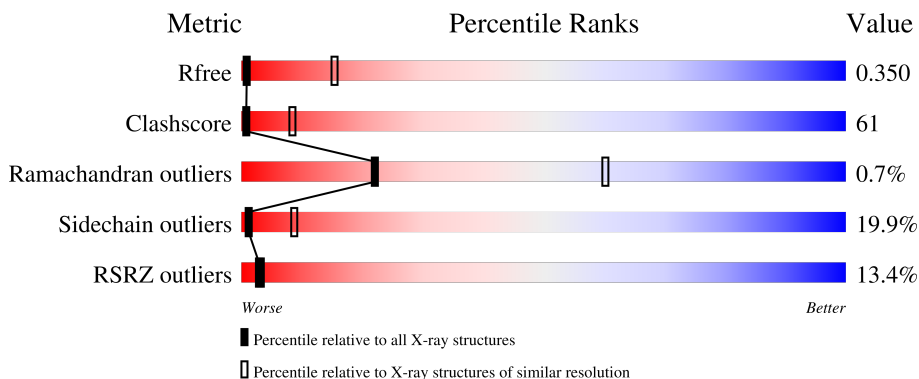
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.90 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	1002 (4.14-3.66)
Clashscore	141614	1004 (4.12-3.68)
Ramachandran outliers	138981	1021 (4.14-3.66)
Sidechain outliers	138945	1014 (4.14-3.66)
RSRZ outliers	127900	1275 (4.20-3.60)

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 of 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	190	
1	B	190	
1	C	190	
1	D	190	
2	E	95	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	95	 21% 11% 68%

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5496 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 Monopolin complex subunit CSM1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	163	1313	840	208	263	2	0	0	0
1	B	164	1326	847	212	265	2	0	0	0
1	C	158	1271	810	203	256	2	0	0	0
1	D	158	1281	821	202	256	2	0	0	0

- Molecule 2 is a protein called Monopolin complex subunit LRS4.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	E	31	155	93	31	31	0	0	0
2	F	30	150	90	30	30	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	?	-	LEU	deletion	UNP Q04087
E	?	-	ASN	deletion	UNP Q04087
E	?	-	ASN	deletion	UNP Q04087
E	?	-	ASN	deletion	UNP Q04087
E	?	-	LYS	deletion	UNP Q04087
E	?	-	GLY	deletion	UNP Q04087
E	?	-	ASP	deletion	UNP Q04087
F	?	-	LEU	deletion	UNP Q04087
F	?	-	ASN	deletion	UNP Q04087
F	?	-	ASN	deletion	UNP Q04087
F	?	-	ASN	deletion	UNP Q04087
F	?	-	LYS	deletion	UNP Q04087

*Continued on next page...*

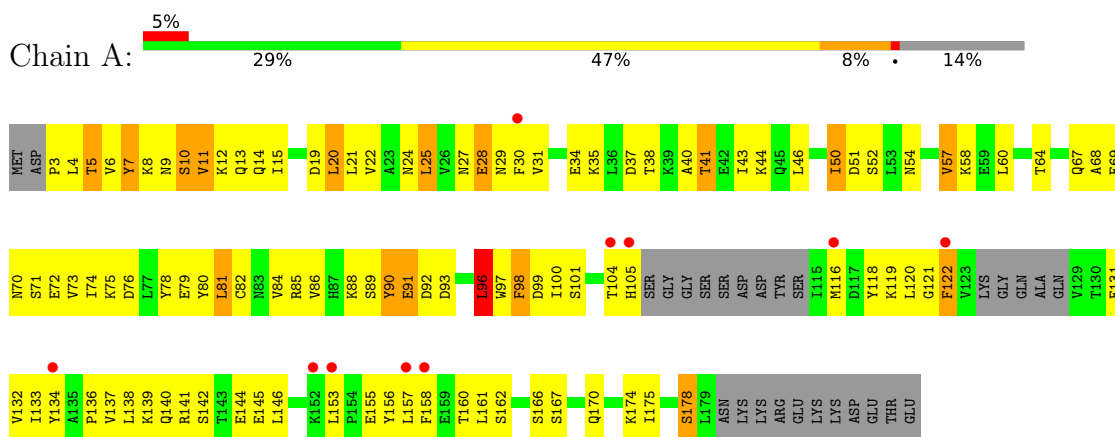
*Continued from previous page...*

<b>Chain</b>	<b>Residue</b>	<b>Modelled</b>	<b>Actual</b>	<b>Comment</b>	<b>Reference</b>
F	?	-	GLY	deletion	UNP Q04087
F	?	-	ASP	deletion	UNP Q04087

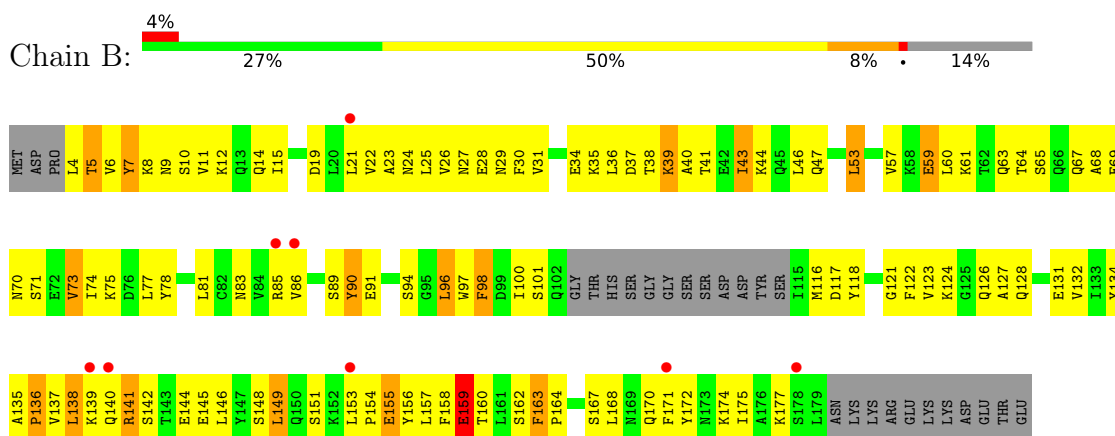
### 3 Residue-property plots [i](#)

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.

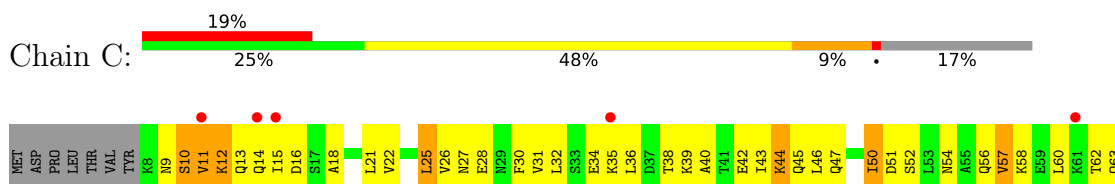
- Molecule 1: Monopolin complex subunit CSM1

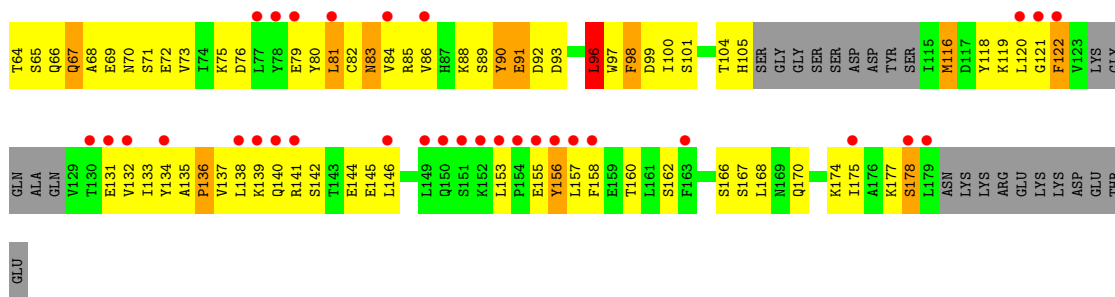


- Molecule 1: Monopolin complex subunit CSM1

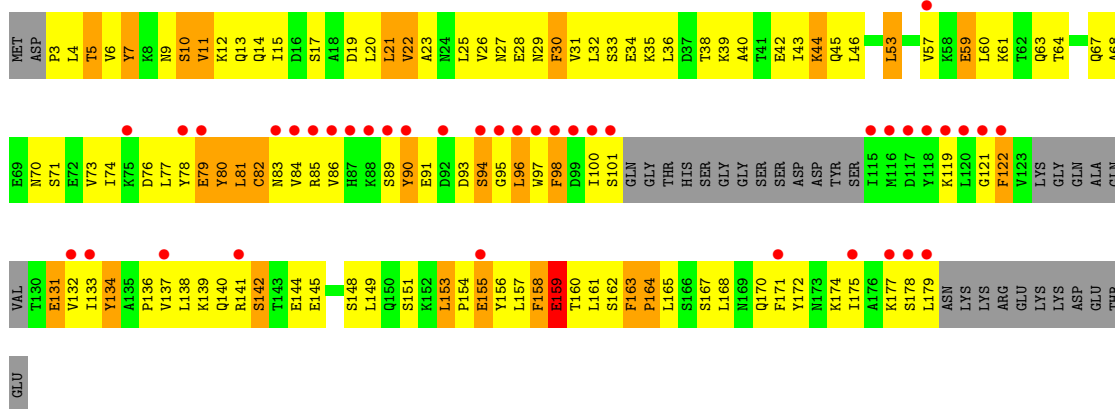
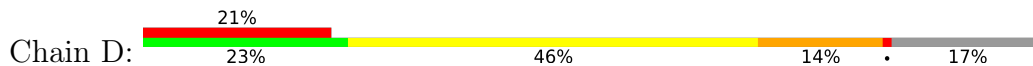


- Molecule 1: Monopolin complex subunit CSM1

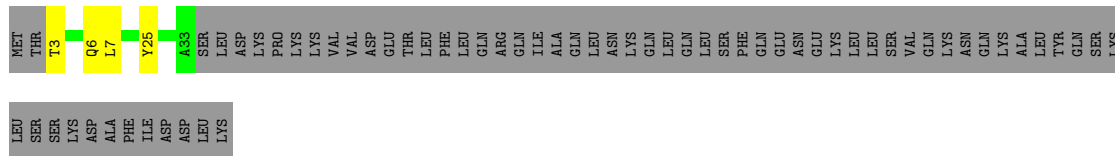




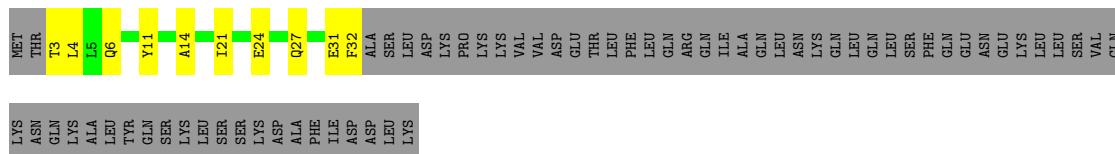
• Molecule 1: Monopolin complex subunit CSM1



• Molecule 2: Monopolin complex subunit LRS4



• Molecule 2: Monopolin complex subunit LRS4



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	152.62Å 152.62Å 118.79Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 3.90 49.96 – 3.91	Depositor EDS
% Data completeness (in resolution range)	44.5 (50.00-3.90) 44.7 (49.96-3.91)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.82 (at 3.88Å)	Xtrriage
Refinement program	CNS 1.3, PHENIX (phenix.refine: 1.6.1_357)	Depositor
R, $R_{free}$	0.333 , 0.355 0.328 , 0.350	Depositor DCC
$R_{free}$ test set	321 reflections (4.83%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	65.7	Xtrriage
Anisotropy	1.103	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 295.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.25$	Xtrriage
Estimated twinning fraction	0.056 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.85	EDS
Total number of atoms	5496	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	358.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.60% 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](#)

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	0.67	0/1332	0.83	2/1800 (0.1%)
1	B	0.57	1/1345 (0.1%)	0.74	1/1817 (0.1%)
1	C	0.68	0/1288	0.85	2/1738 (0.1%)
1	D	0.60	1/1300 (0.1%)	0.74	1/1756 (0.1%)
2	E	0.25	0/154	0.41	0/214
2	F	0.21	0/149	0.34	0/207
All	All	0.61	2/5568 (0.0%)	0.77	6/7532 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	159	GLU	CG-CD	5.49	1.60	1.51
1	D	159	GLU	CG-CD	5.46	1.60	1.51

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	96	LEU	CB-CG-CD1	-5.75	101.22	111.00
1	D	149	LEU	CA-CB-CG	-5.73	102.13	115.30
1	A	81	LEU	CA-CB-CG	-5.72	102.15	115.30
1	B	149	LEU	CA-CB-CG	-5.69	102.22	115.30
1	A	96	LEU	CB-CG-CD1	-5.68	101.34	111.00

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	1313	0	1310	149	0
1	B	1326	0	1328	175	0
1	C	1271	0	1265	214	1
1	D	1281	0	1281	217	1
2	E	155	0	67	14	0
2	F	150	0	62	19	0
All	All	5496	0	5313	659	1

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 61.

The worst 5 of 659 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:26:VAL:HG12	1:C:30:PHE:CE2	1.24	1.69
1:A:7:TYR:CD1	1:B:7:TYR:CE1	2.02	1.47
1:D:14:GLN:HB3	2:F:14:ALA:CB	1.46	1.43
1:C:36:LEU:CD1	1:D:36:LEU:HD12	1.54	1.36
1:D:14:GLN:CB	2:F:14:ALA:HB1	1.53	1.35

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:39:LYS:CE	1:D:30:PHE:CD2[2_545]	1.98	0.22

## 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	157/190 (83%)	143 (91%)	14 (9%)	0	100	100
1	B	160/190 (84%)	148 (92%)	11 (7%)	1 (1%)	25	63

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	152/190 (80%)	139 (91%)	12 (8%)	1 (1%)	22	60
1	D	152/190 (80%)	132 (87%)	17 (11%)	3 (2%)	7	40
2	E	29/95 (30%)	29 (100%)	0	0	100	100
2	F	28/95 (30%)	27 (96%)	1 (4%)	0	100	100
All	All	678/950 (71%)	618 (91%)	55 (8%)	5 (1%)	22	60

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	33	SER
1	D	80	TYR
1	B	136	PRO
1	C	136	PRO
1	D	158	PHE

### 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	151/176 (86%)	125 (83%)	26 (17%)	2	14
1	B	152/176 (86%)	123 (81%)	29 (19%)	1	10
1	C	146/176 (83%)	121 (83%)	25 (17%)	2	14
1	D	148/176 (84%)	109 (74%)	39 (26%)	0	4
All	All	597/704 (85%)	478 (80%)	119 (20%)	1	9

5 of 119 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	12	LYS
1	D	140	GLN
1	C	98	PHE
1	D	134	TYR
1	D	164	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	13	GLN
1	D	169	ASN
1	D	83	ASN
1	B	169	ASN
1	C	102	GLN

### 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 monosaccharides 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, 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	163/190 (85%)	-0.00	10 (6%) 21 15	195, 331, 461, 500	0
1	B	164/190 (86%)	-0.10	8 (4%) 29 24	149, 319, 447, 500	0
1	C	158/190 (83%)	1.10	37 (23%) 0 0	182, 398, 500, 500	0
1	D	158/190 (83%)	1.72	39 (24%) 0 0	142, 400, 500, 500	0
2	E	31/95 (32%)	-0.52	0 100 100	107, 270, 426, 483	0
2	F	30/95 (31%)	-0.93	0 100 100	201, 272, 384, 500	0
All	All	704/950 (74%)	0.55	94 (13%) 3 3	107, 349, 500, 500	0

The worst 5 of 94 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	116	MET	21.0
1	D	87	HIS	19.0
1	D	85	ARG	17.5
1	D	115	ILE	17.3
1	D	101	SER	15.9

### 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 monosaccharides 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.