



wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 01:41 PM GMT

PDB ID : 1Q9C
Title : Crystal Structure of the Histone domain of Son of Sevenless
Authors : Sondermann, H.; Soisson, S.M.; Bar-Sagi, D.; Kuriyan, J.
Deposited on : 2003-08-24
Resolution : 3.21 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

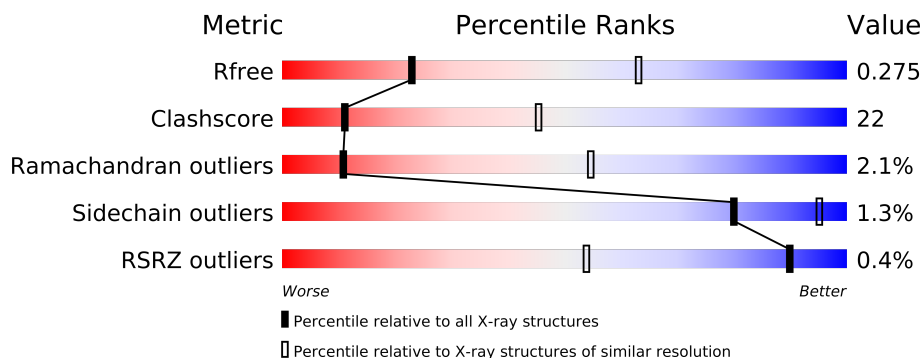
The following versions of software and data (see [references](#)) were used in the production of this report:

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 3.21 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	1205 (3.30-3.14)
Clashscore	79885	1072 (3.28-3.16)
Ramachandran outliers	78287	1052 (3.28-3.16)
Sidechain outliers	78261	1051 (3.28-3.16)
RSRZ outliers	66119	1206 (3.30-3.14)

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. 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	191	
1	B	191	
1	C	191	
1	D	191	
1	E	191	
1	F	191	
1	G	191	
1	H	191	
1	I	191	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 12555 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 Son of sevenless protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	164	Total	C	N	O	S	Se	0	0	0
			1320	851	220	243	2	4			
1	B	172	Total	C	N	O	S	Se	0	0	0
			1395	895	241	253	2	4			
1	C	162	Total	C	N	O	S	Se	0	0	0
			1299	837	218	238	2	4			
1	D	178	Total	C	N	O	S	Se	0	0	0
			1443	925	247	265	2	4			
1	E	176	Total	C	N	O	S	Se	0	0	0
			1423	913	241	263	2	4			
1	F	174	Total	C	N	O	S	Se	0	0	0
			1404	903	235	260	2	4			
1	G	174	Total	C	N	O	S	Se	0	0	0
			1411	905	243	257	2	4			
1	H	175	Total	C	N	O	S	Se	0	0	0
			1417	910	240	261	2	4			
1	I	178	Total	C	N	O	S	Se	0	0	0
			1443	925	247	265	2	4			

There are 45 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	58	MSE	MET	MODIFIED RESIDUE	UNP Q07889
A	145	ALA	VAL	SEE REMARK 999	UNP Q07889
A	166	MSE	MET	MODIFIED RESIDUE	UNP Q07889
A	173	MSE	MET	MODIFIED RESIDUE	UNP Q07889
A	175	MSE	MET	MODIFIED RESIDUE	UNP Q07889
B	58	MSE	MET	MODIFIED RESIDUE	UNP Q07889
B	145	ALA	VAL	SEE REMARK 999	UNP Q07889
B	166	MSE	MET	MODIFIED RESIDUE	UNP Q07889
B	173	MSE	MET	MODIFIED RESIDUE	UNP Q07889
B	175	MSE	MET	MODIFIED RESIDUE	UNP Q07889
C	58	MSE	MET	MODIFIED RESIDUE	UNP Q07889

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	145	ALA	VAL	SEE REMARK 999	UNP Q07889
C	166	MSE	MET	MODIFIED RESIDUE	UNP Q07889
C	173	MSE	MET	MODIFIED RESIDUE	UNP Q07889
C	175	MSE	MET	MODIFIED RESIDUE	UNP Q07889
D	58	MSE	MET	MODIFIED RESIDUE	UNP Q07889
D	145	ALA	VAL	SEE REMARK 999	UNP Q07889
D	166	MSE	MET	MODIFIED RESIDUE	UNP Q07889
D	173	MSE	MET	MODIFIED RESIDUE	UNP Q07889
D	175	MSE	MET	MODIFIED RESIDUE	UNP Q07889
E	58	MSE	MET	MODIFIED RESIDUE	UNP Q07889
E	145	ALA	VAL	SEE REMARK 999	UNP Q07889
E	166	MSE	MET	MODIFIED RESIDUE	UNP Q07889
E	173	MSE	MET	MODIFIED RESIDUE	UNP Q07889
E	175	MSE	MET	MODIFIED RESIDUE	UNP Q07889
F	58	MSE	MET	MODIFIED RESIDUE	UNP Q07889
F	145	ALA	VAL	SEE REMARK 999	UNP Q07889
F	166	MSE	MET	MODIFIED RESIDUE	UNP Q07889
F	173	MSE	MET	MODIFIED RESIDUE	UNP Q07889
F	175	MSE	MET	MODIFIED RESIDUE	UNP Q07889
G	58	MSE	MET	MODIFIED RESIDUE	UNP Q07889
G	145	ALA	VAL	SEE REMARK 999	UNP Q07889
G	166	MSE	MET	MODIFIED RESIDUE	UNP Q07889
G	173	MSE	MET	MODIFIED RESIDUE	UNP Q07889
G	175	MSE	MET	MODIFIED RESIDUE	UNP Q07889
H	58	MSE	MET	MODIFIED RESIDUE	UNP Q07889
H	145	ALA	VAL	SEE REMARK 999	UNP Q07889
H	166	MSE	MET	MODIFIED RESIDUE	UNP Q07889
H	173	MSE	MET	MODIFIED RESIDUE	UNP Q07889
H	175	MSE	MET	MODIFIED RESIDUE	UNP Q07889
I	58	MSE	MET	MODIFIED RESIDUE	UNP Q07889
I	145	ALA	VAL	SEE REMARK 999	UNP Q07889
I	166	MSE	MET	MODIFIED RESIDUE	UNP Q07889
I	173	MSE	MET	MODIFIED RESIDUE	UNP Q07889
I	175	MSE	MET	MODIFIED RESIDUE	UNP Q07889

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: Son of sevenless protein

Chain A:



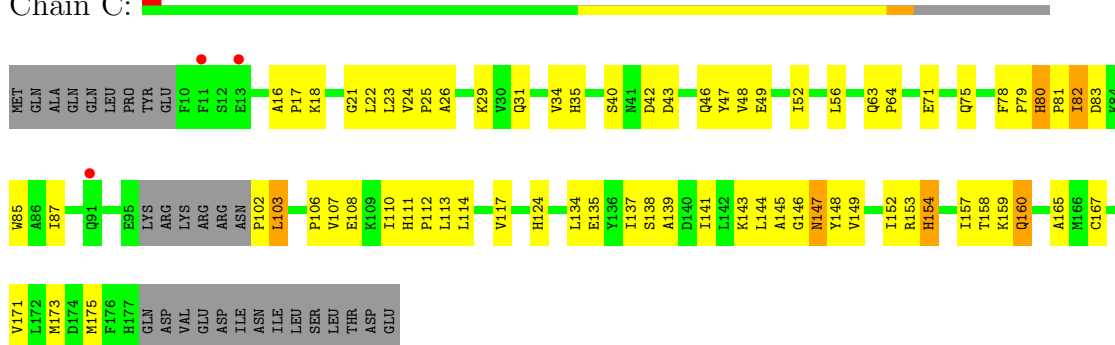
- Molecule 1: Son of sevenless protein

Chain B:



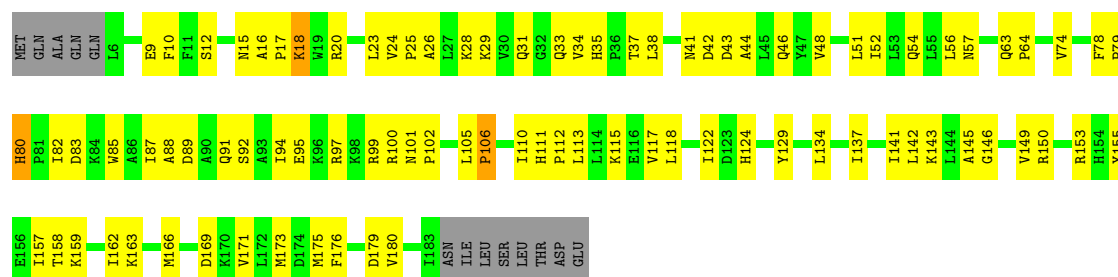
- Molecule 1: Son of sevenless protein

Chain C:



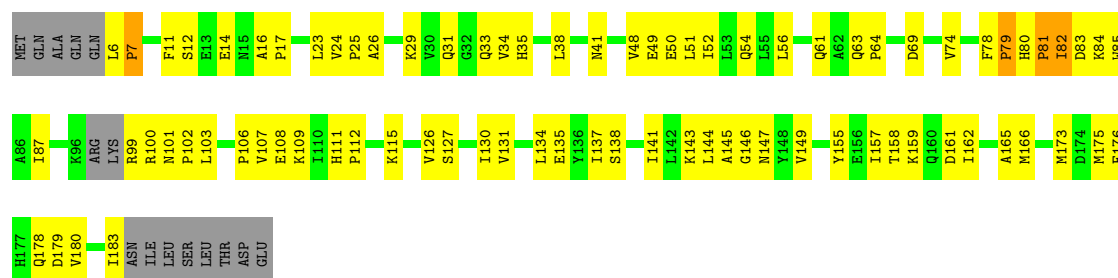
- Molecule 1: Son of sevenless protein

Chain D:



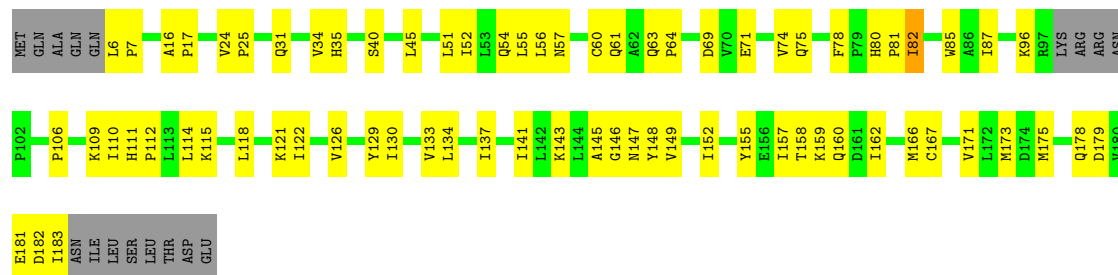
- Molecule 1: Son of sevenless protein

Chain E:



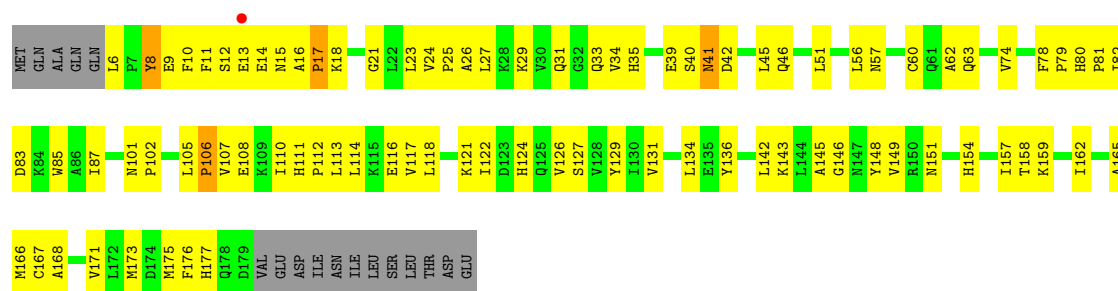
- Molecule 1: Son of sevenless protein

Chain F:



- Molecule 1: Son of sevenless protein

Chain G:



- Molecule 1: Son of sevenless protein

Chain H:

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	93.40Å 109.20Å 212.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.90 – 3.21 97.12 – 3.21	Depositor EDS
% Data completeness (in resolution range)	96.7 (19.90-3.21) 96.7 (97.12-3.21)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.15 (at 3.19Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.255 , 0.292 0.247 , 0.275	Depositor DCC
R_{free} test set	2436 reflections (6.98%)	DCC
Wilson B-factor (Å ²)	82.3	Xtriage
Anisotropy	0.229	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 40.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 35037 reflections	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	12555	wwPDB-VP
Average B, all atoms (Å ²)	82.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.49% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities 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	0.23	0/1344	0.50	2/1816 (0.1%)
1	B	0.23	0/1421	0.50	0/1917
1	C	0.24	0/1322	0.47	1/1785 (0.1%)
1	D	0.23	0/1469	0.49	2/1984 (0.1%)
1	E	0.24	0/1448	0.55	2/1956 (0.1%)
1	F	0.24	0/1429	0.53	2/1930 (0.1%)
1	G	0.23	0/1437	0.48	1/1940 (0.1%)
1	H	0.23	0/1441	0.50	1/1944 (0.1%)
1	I	0.25	0/1469	0.51	2/1984 (0.1%)
All	All	0.24	0/12780	0.50	13/17256 (0.1%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	80	HIS	C-N-CD	-10.65	97.16	120.60
1	F	80	HIS	C-N-CD	-8.77	101.30	120.60
1	I	80	HIS	C-N-CD	-8.04	102.92	120.60
1	A	80	HIS	C-N-CD	-7.48	104.13	120.60
1	E	80	HIS	C-N-CA	7.42	153.16	122.00

There are no chirality outliers.

There are no planarity outliers.

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	1320	0	1334	48	0
1	B	1395	0	1422	65	0
1	C	1299	0	1320	65	0
1	D	1443	0	1466	77	0
1	E	1423	0	1439	71	0
1	F	1404	0	1421	49	0
1	G	1411	0	1436	77	0
1	H	1417	0	1439	63	0
1	I	1443	0	1466	63	0
All	All	12555	0	12743	568	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 22.

The worst 5 of 568 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:81:PRO:HB2	1:B:85:TRP:HE1	1.12	1.10
1:B:81:PRO:CB	1:B:85:TRP:HE1	1.80	0.94
1:B:79:PRO:HD2	1:B:175:MSE:HE2	1.52	0.91
1:A:60:CYS:HB3	1:A:118:LEU:HD21	1.52	0.91
1:G:15:ASN:HA	1:G:18:LYS:HE2	1.54	0.89

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	160/191 (84%)	141 (88%)	18 (11%)	1 (1%)	33	83
1	B	170/191 (89%)	145 (85%)	20 (12%)	5 (3%)	7	43
1	C	158/191 (83%)	138 (87%)	18 (11%)	2 (1%)	18	68
1	D	176/191 (92%)	159 (90%)	15 (8%)	2 (1%)	21	72
1	E	172/191 (90%)	144 (84%)	22 (13%)	6 (4%)	6	37
1	F	170/191 (89%)	151 (89%)	15 (9%)	4 (2%)	9	50

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	G	172/191 (90%)	144 (84%)	21 (12%)	7 (4%)	4	32
1	H	171/191 (90%)	154 (90%)	14 (8%)	3 (2%)	13	60
1	I	176/191 (92%)	157 (89%)	17 (10%)	2 (1%)	21	72
All	All	1525/1719 (89%)	1333 (87%)	160 (10%)	32 (2%)	11	54

5 of 32 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	82	ILE
1	E	178	GLN
1	H	96	LYS
1	B	82	ILE
1	G	8	TYR

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	146/168 (87%)	144 (99%)	2 (1%)	78	95
1	B	154/168 (92%)	153 (99%)	1 (1%)	92	98
1	C	144/168 (86%)	140 (97%)	4 (3%)	56	89
1	D	160/168 (95%)	159 (99%)	1 (1%)	92	98
1	E	158/168 (94%)	156 (99%)	2 (1%)	80	95
1	F	156/168 (93%)	154 (99%)	2 (1%)	80	95
1	G	156/168 (93%)	156 (100%)	0	100	100
1	H	157/168 (94%)	155 (99%)	2 (1%)	80	95
1	I	160/168 (95%)	156 (98%)	4 (2%)	60	90
All	All	1391/1512 (92%)	1373 (99%)	18 (1%)	80	95

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	69	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	183	ILE
1	I	123	ASP
1	C	160	GLN
1	D	18	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	57	ASN
1	F	154	HIS
1	I	111	HIS
1	F	61	GLN
1	F	160	GLN

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, 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	164/191 (85%)	-0.00	0 100 100	21, 69, 171, 195	0
1	B	172/191 (90%)	-0.04	1 (0%) 86 43	22, 79, 178, 196	0
1	C	162/191 (84%)	0.02	3 (1%) 64 19	26, 87, 169, 195	0
1	D	178/191 (93%)	-0.19	0 100 100	21, 56, 162, 195	0
1	E	176/191 (92%)	-0.04	0 100 100	17, 85, 180, 196	0
1	F	174/191 (91%)	-0.14	0 100 100	15, 59, 162, 195	0
1	G	174/191 (91%)	-0.09	1 (0%) 86 43	21, 73, 162, 194	0
1	H	175/191 (91%)	-0.09	1 (0%) 86 43	15, 64, 162, 194	0
1	I	178/191 (93%)	-0.13	0 100 100	21, 67, 169, 195	0
All	All	1553/1719 (90%)	-0.08	6 (0%) 90 54	15, 70, 171, 196	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	118	LEU	3.3
1	C	91	GLN	2.9
1	G	13	GLU	2.4
1	H	175	MSE	2.4
1	C	11	PHE	2.2

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.