



wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 06:23 am BST

PDB ID : 5IW7
Title : Crystal structure of yeast Tsr1, a pre-40S ribosome synthesis factor
Authors : McCaughan, U.M.; Jayachandran, U.; Cook, A.G.
Deposited on : 2016-03-22
Resolution : 3.60 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

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
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

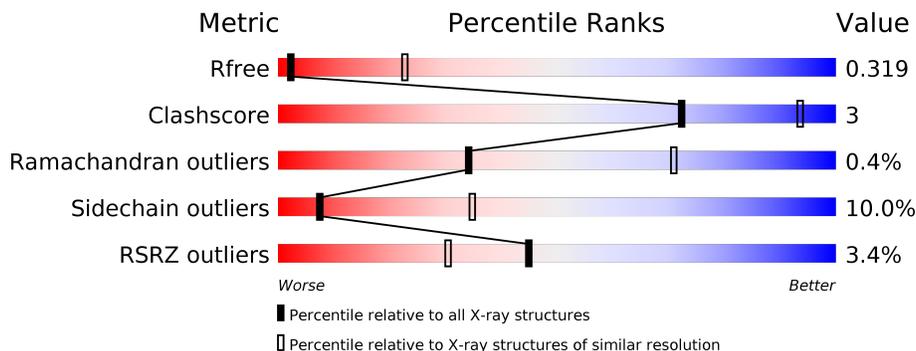
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.60 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}	130704	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RSRZ outliers	127900	1161 (3.70-3.50)

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 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	704	 3% 70% 13% • 15%
1	B	704	 3% 73% 8% • 19%
1	C	704	 3% 69% 10% • 19%
1	D	704	 3% 65% 8% 27%

2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 16065 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 Ribosome biogenesis protein TSR1, Ribosome biogenesis protein TSR1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	595	4531	2905	771	843	12	0	0	0
1	B	569	4015	2543	691	771	10	0	0	0
1	C	567	4093	2615	697	771	10	0	0	0
1	D	517	3426	2145	604	669	8	0	0	0

There are 112 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	24	MET	-	initiating methionine	UNP Q07381
A	25	LYS	-	expression tag	UNP Q07381
A	26	HIS	-	expression tag	UNP Q07381
A	27	HIS	-	expression tag	UNP Q07381
A	28	HIS	-	expression tag	UNP Q07381
A	29	HIS	-	expression tag	UNP Q07381
A	30	HIS	-	expression tag	UNP Q07381
A	31	HIS	-	expression tag	UNP Q07381
A	32	SER	-	expression tag	UNP Q07381
A	33	ALA	-	expression tag	UNP Q07381
A	34	GLY	-	expression tag	UNP Q07381
A	35	LEU	-	expression tag	UNP Q07381
A	36	GLU	-	expression tag	UNP Q07381
A	37	VAL	-	expression tag	UNP Q07381
A	38	LEU	-	expression tag	UNP Q07381
A	39	PHE	-	expression tag	UNP Q07381
A	40	GLN	-	expression tag	UNP Q07381
A	41	GLY	-	expression tag	UNP Q07381
A	42	PRO	-	expression tag	UNP Q07381
A	43	ASP	-	expression tag	UNP Q07381

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	44	SER	-	expression tag	UNP Q07381
A	45	MET	-	expression tag	UNP Q07381
A	471	PRO	-	linker	UNP Q07381
A	472	SER	-	linker	UNP Q07381
A	473	SER	-	linker	UNP Q07381
A	474	GLY	-	linker	UNP Q07381
A	475	SER	-	linker	UNP Q07381
A	476	SER	-	linker	UNP Q07381
B	24	MET	-	initiating methionine	UNP Q07381
B	25	LYS	-	expression tag	UNP Q07381
B	26	HIS	-	expression tag	UNP Q07381
B	27	HIS	-	expression tag	UNP Q07381
B	28	HIS	-	expression tag	UNP Q07381
B	29	HIS	-	expression tag	UNP Q07381
B	30	HIS	-	expression tag	UNP Q07381
B	31	HIS	-	expression tag	UNP Q07381
B	32	SER	-	expression tag	UNP Q07381
B	33	ALA	-	expression tag	UNP Q07381
B	34	GLY	-	expression tag	UNP Q07381
B	35	LEU	-	expression tag	UNP Q07381
B	36	GLU	-	expression tag	UNP Q07381
B	37	VAL	-	expression tag	UNP Q07381
B	38	LEU	-	expression tag	UNP Q07381
B	39	PHE	-	expression tag	UNP Q07381
B	40	GLN	-	expression tag	UNP Q07381
B	41	GLY	-	expression tag	UNP Q07381
B	42	PRO	-	expression tag	UNP Q07381
B	43	ASP	-	expression tag	UNP Q07381
B	44	SER	-	expression tag	UNP Q07381
B	45	MET	-	expression tag	UNP Q07381
B	471	PRO	-	linker	UNP Q07381
B	472	SER	-	linker	UNP Q07381
B	473	SER	-	linker	UNP Q07381
B	474	GLY	-	linker	UNP Q07381
B	475	SER	-	linker	UNP Q07381
B	476	SER	-	linker	UNP Q07381
C	24	MET	-	initiating methionine	UNP Q07381
C	25	LYS	-	expression tag	UNP Q07381
C	26	HIS	-	expression tag	UNP Q07381
C	27	HIS	-	expression tag	UNP Q07381
C	28	HIS	-	expression tag	UNP Q07381
C	29	HIS	-	expression tag	UNP Q07381

Continued on next page...

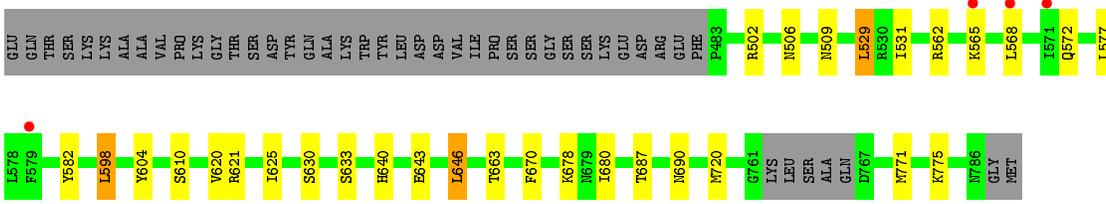
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	30	HIS	-	expression tag	UNP Q07381
C	31	HIS	-	expression tag	UNP Q07381
C	32	SER	-	expression tag	UNP Q07381
C	33	ALA	-	expression tag	UNP Q07381
C	34	GLY	-	expression tag	UNP Q07381
C	35	LEU	-	expression tag	UNP Q07381
C	36	GLU	-	expression tag	UNP Q07381
C	37	VAL	-	expression tag	UNP Q07381
C	38	LEU	-	expression tag	UNP Q07381
C	39	PHE	-	expression tag	UNP Q07381
C	40	GLN	-	expression tag	UNP Q07381
C	41	GLY	-	expression tag	UNP Q07381
C	42	PRO	-	expression tag	UNP Q07381
C	43	ASP	-	expression tag	UNP Q07381
C	44	SER	-	expression tag	UNP Q07381
C	45	MET	-	expression tag	UNP Q07381
C	471	PRO	-	linker	UNP Q07381
C	472	SER	-	linker	UNP Q07381
C	473	SER	-	linker	UNP Q07381
C	474	GLY	-	linker	UNP Q07381
C	475	SER	-	linker	UNP Q07381
C	476	SER	-	linker	UNP Q07381
D	24	MET	-	initiating methionine	UNP Q07381
D	25	LYS	-	expression tag	UNP Q07381
D	26	HIS	-	expression tag	UNP Q07381
D	27	HIS	-	expression tag	UNP Q07381
D	28	HIS	-	expression tag	UNP Q07381
D	29	HIS	-	expression tag	UNP Q07381
D	30	HIS	-	expression tag	UNP Q07381
D	31	HIS	-	expression tag	UNP Q07381
D	32	SER	-	expression tag	UNP Q07381
D	33	ALA	-	expression tag	UNP Q07381
D	34	GLY	-	expression tag	UNP Q07381
D	35	LEU	-	expression tag	UNP Q07381
D	36	GLU	-	expression tag	UNP Q07381
D	37	VAL	-	expression tag	UNP Q07381
D	38	LEU	-	expression tag	UNP Q07381
D	39	PHE	-	expression tag	UNP Q07381
D	40	GLN	-	expression tag	UNP Q07381
D	41	GLY	-	expression tag	UNP Q07381
D	42	PRO	-	expression tag	UNP Q07381
D	43	ASP	-	expression tag	UNP Q07381

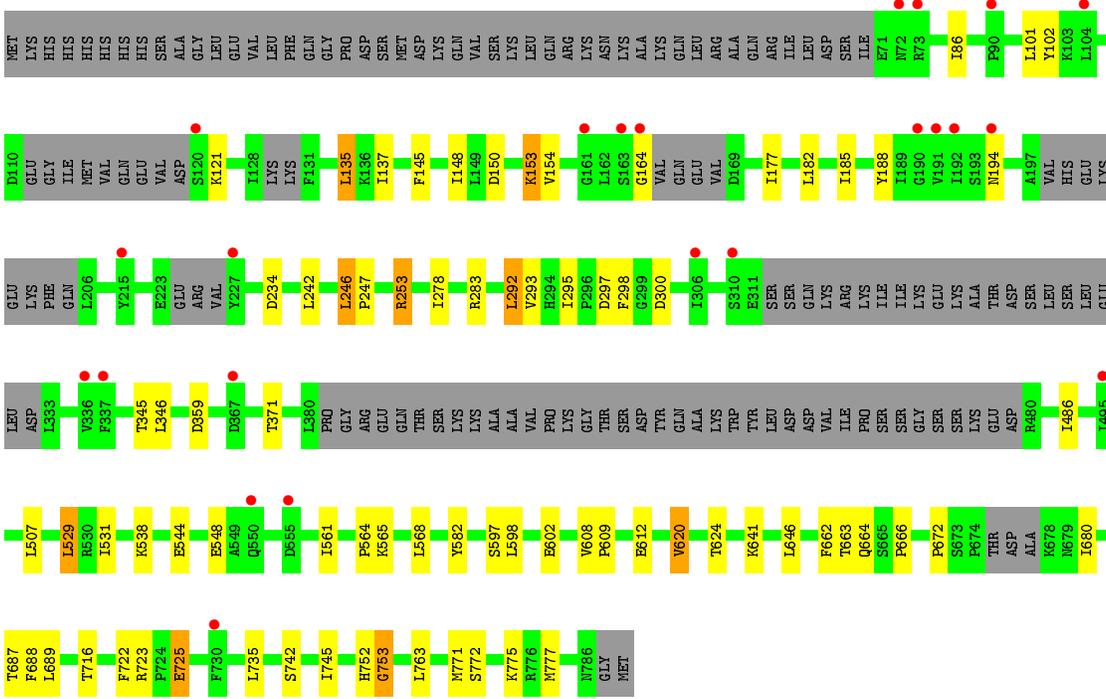
Continued on next page...

Continued from previous page...

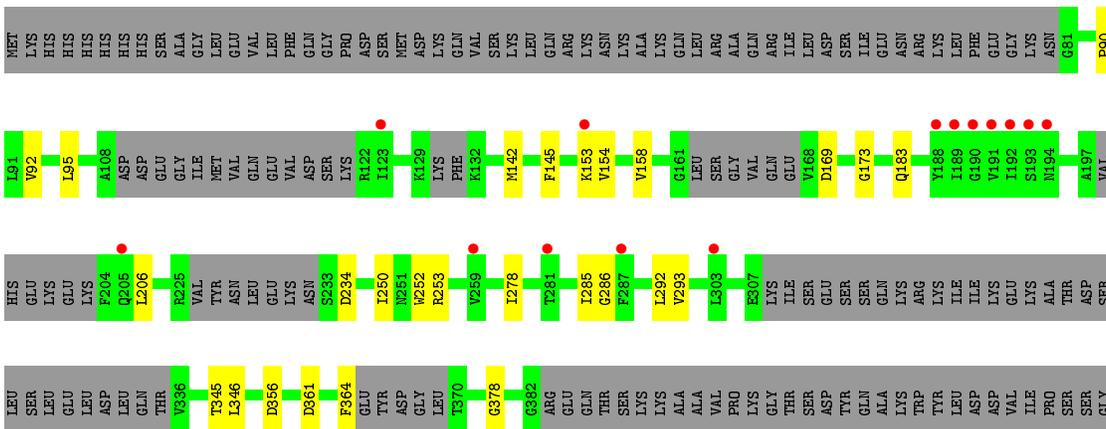
Chain	Residue	Modelled	Actual	Comment	Reference
D	44	SER	-	expression tag	UNP Q07381
D	45	MET	-	expression tag	UNP Q07381
D	471	PRO	-	linker	UNP Q07381
D	472	SER	-	linker	UNP Q07381
D	473	SER	-	linker	UNP Q07381
D	474	GLY	-	linker	UNP Q07381
D	475	SER	-	linker	UNP Q07381
D	476	SER	-	linker	UNP Q07381

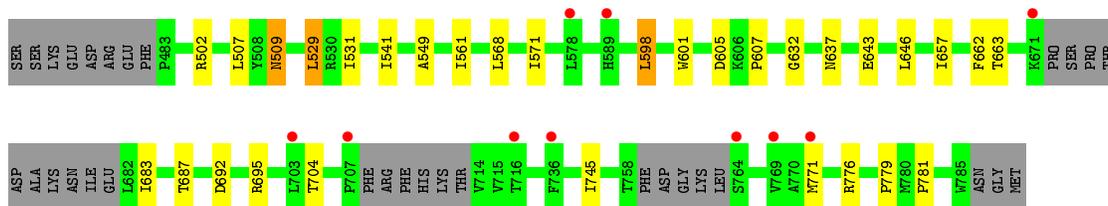


• Molecule 1: Ribosome biogenesis protein TSR1, Ribosome biogenesis protein TSR1



• Molecule 1: Ribosome biogenesis protein TSR1, Ribosome biogenesis protein TSR1





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	65.69Å 174.20Å 320.47Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.77 – 3.60 48.77 – 3.60	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.77-3.60) 99.9 (48.77-3.60)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.16 (at 3.57Å)	Xtrriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.268 , 0.295 0.288 , 0.319	Depositor DCC
R_{free} test set	2228 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	123.0	Xtrriage
Anisotropy	0.438	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 111.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	16065	wwPDB-VP
Average B, all atoms (Å ²)	87.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.57% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.37	0/4639	0.61	0/6324
1	B	0.36	0/4105	0.58	0/5625
1	C	0.37	0/4193	0.61	0/5740
1	D	0.36	0/3494	0.59	0/4811
All	All	0.36	0/16431	0.60	0/22500

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	4531	0	4160	32	0
1	B	4015	0	3368	18	0
1	C	4093	0	3483	32	0
1	D	3426	0	2669	19	0
All	All	16065	0	13680	100	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 3.

The worst 5 of 100 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:514:VAL:HA	1:A:776:ARG:HH12	1.44	0.81
1:D:657:ILE:HG23	1:D:779:PRO:HB2	1.68	0.76
1:C:672:PRO:HA	1:C:680:ILE:HA	1.73	0.69
1:B:302:GLN:NE2	1:B:343:ARG:HD2	2.08	0.68
1:D:286:GLY:HA3	1:D:549:ALA:HB1	1.77	0.66

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	583/704 (83%)	550 (94%)	32 (6%)	1 (0%)	47 79
1	B	555/704 (79%)	522 (94%)	31 (6%)	2 (0%)	34 71
1	C	549/704 (78%)	507 (92%)	38 (7%)	4 (1%)	22 61
1	D	493/704 (70%)	462 (94%)	30 (6%)	1 (0%)	47 79
All	All	2180/2816 (77%)	2041 (94%)	131 (6%)	8 (0%)	34 71

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	375	ASP
1	C	602	GLU
1	C	753	GLY
1	A	81	GLY
1	B	352	GLU

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	449/630 (71%)	397 (88%)	52 (12%)	5	29
1	B	348/630 (55%)	319 (92%)	29 (8%)	11	42
1	C	363/630 (58%)	329 (91%)	34 (9%)	8	38
1	D	264/630 (42%)	237 (90%)	27 (10%)	7	34
All	All	1424/2520 (56%)	1282 (90%)	142 (10%)	7	35

5 of 142 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	355	GLU
1	B	771	MET
1	D	529	LEU
1	B	502	ARG
1	B	625	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	572	GLN
1	B	640	HIS
1	C	587	HIS
1	B	377	HIS
1	D	183	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 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	595/704 (84%)	-0.02	5 (0%) 86 75	8, 43, 112, 216	0
1	B	569/704 (80%)	0.13	24 (4%) 36 23	15, 96, 191, 294	0
1	C	567/704 (80%)	0.12	23 (4%) 37 24	18, 89, 193, 291	0
1	D	517/704 (73%)	0.18	24 (4%) 32 20	29, 114, 263, 298	0
All	All	2248/2816 (79%)	0.10	76 (3%) 45 30	8, 85, 191, 298	0

The worst 5 of 76 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	190	GLY	8.7
1	D	192	ILE	6.2
1	D	191	VAL	5.6
1	D	189	ILE	5.0
1	C	191	VAL	4.6

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

There are no such residues in this entry.