



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 11:19 PM UTC

PDB ID : 4OY2 / pdb_00004oy2
Title : Crystal structure of TAF1-TAF7, a TFIID subcomplex
Authors : Bhattacharya, S.; Lou, X.; Rajashankar, K.; Jacobson, R.H.; Webb, P.
Deposited on : 2014-02-10
Resolution : 2.90 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

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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

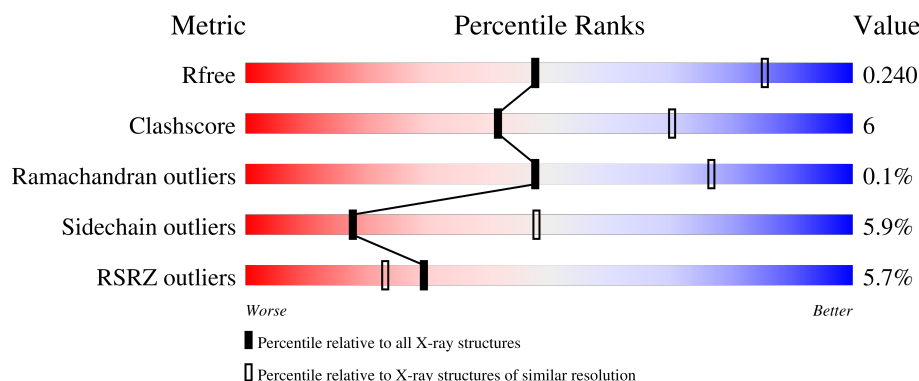
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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}	180053	2481 (2.90-2.90)
Clashscore	190562	2690 (2.90-2.90)
Ramachandran outliers	187476	2623 (2.90-2.90)
Sidechain outliers	187428	2625 (2.90-2.90)
RSRZ outliers	180081	2481 (2.90-2.90)

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 ≥ 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	506	 3% 75% 16% 7%
1	C	506	 10% 77% 13% 9%
1	E	506	 2% 71% 19% 8%
2	B	235	 3% 77% 20% 2%
2	D	235	 9% 77% 19% 3%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	235	4% 71% 24%

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 16474 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 Transcription initiation factor TFIID subunit 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	470	Total	C	N	O	S	Se	0	3	0
			3726	2358	660	696	1	11			
1	C	460	Total	C	N	O	S	Se	0	2	0
			3526	2231	619	664	1	11			
1	E	468	Total	C	N	O	S	Se	0	8	0
			3787	2401	670	704	1	11			

- Molecule 2 is a protein called Transcription initiation factor TFIID subunit 7.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	229	Total	C	N	O	S	Se	0	0	0
			1825	1157	301	361	1	5			
2	D	225	Total	C	N	O	S	Se	0	1	0
			1774	1123	295	350	1	5			
2	F	225	Total	C	N	O	S	Se	0	3	0
			1814	1149	298	361	1	5			

- Molecule 3 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Zn	0	0
			1	1		

- Molecule 4 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (CCD ID: TRS) (formula: C₄H₁₂NO₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	D	1	Total	C	N	O	0	0
			8	4	1	3		

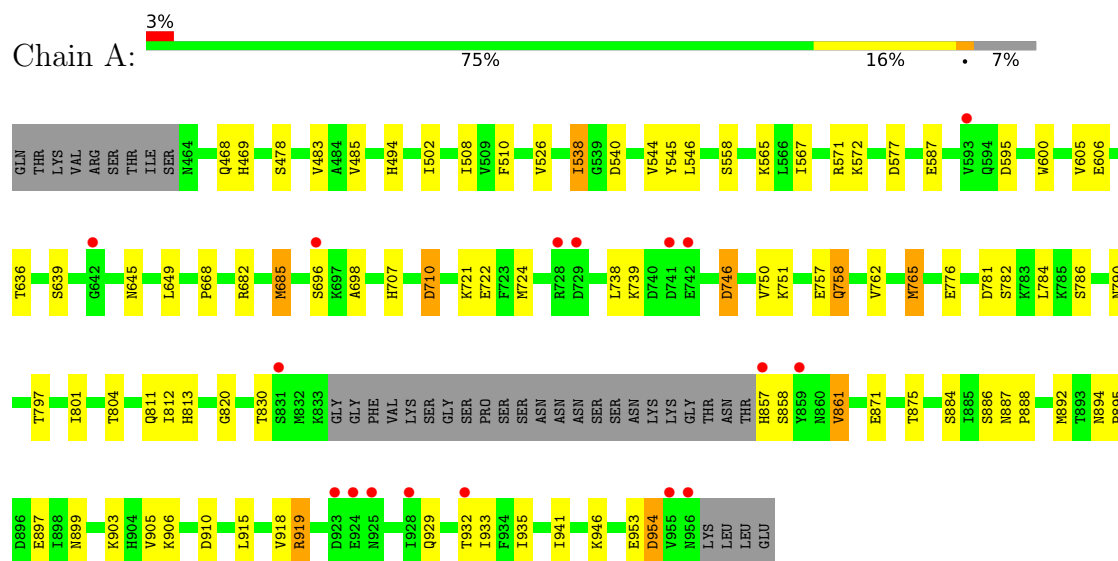
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	3	Total	O	0	0
			3	3		
5	C	2	Total	O	0	0
			2	2		
5	E	5	Total	O	0	0
			5	5		
5	D	2	Total	O	0	0
			2	2		
5	F	1	Total	O	0	0
			1	1		

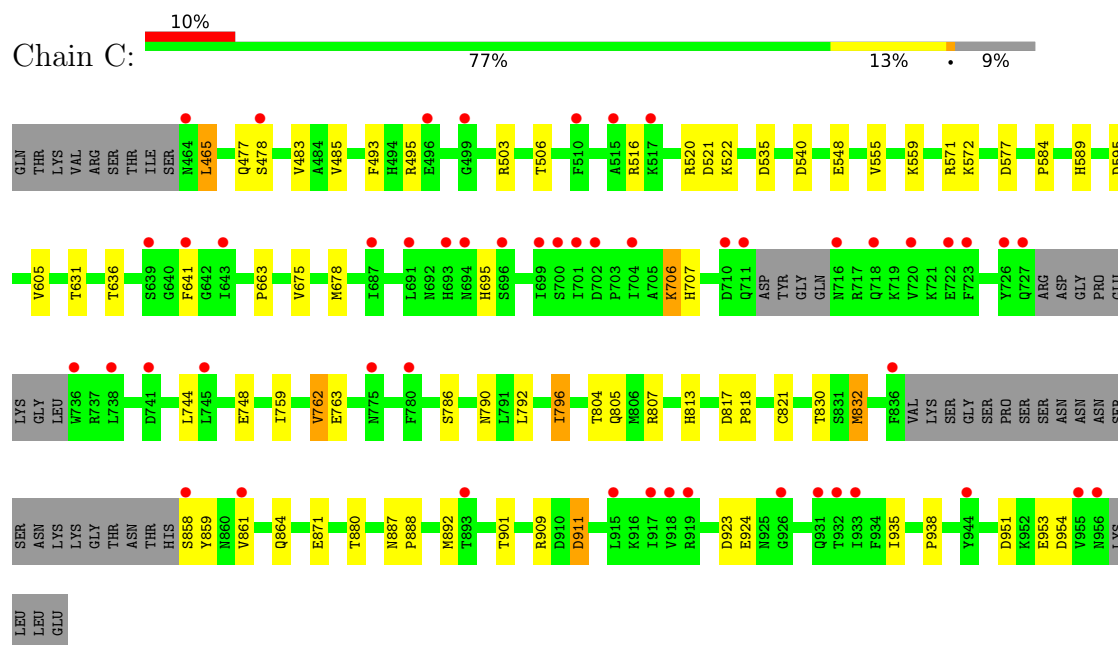
3 Residue-property plots [i](#)

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: Transcription initiation factor TFIID subunit 1



• Molecule 1: Transcription initiation factor TFIID subunit 1



Chain E:

Amino Acid	Count
GLN	1462
THR	469
LYS	5470
VAL	1476
ARG	4477
SER	4478
THR	4482
THR	4492
THR	4499
THR	1508
THR	1509
THR	1510
R516	1516
K517	1517
R520	1520
D521	1521
Q534	1534
D535	1535
L538	1538
T541	1541
L546	1546
M547	1547
E548	1548
V555	1555
K559	1559
F560	1560
G561	1561
M562	1562
A563	1563
N564	1564
K565	1565
Y569	1569
Y570	1570
R571	1571
N574	1574
E575	1575
R576	1576

Chain B:

3% 77% 20% ..

LYS L87 R88 R89 R90 E97 A98 Y99 D100 D105 I124 Q125 I143 N147 H150 V158 L167 P169 T169 V170 V173 A174 K175 S176 V177 D178 R179 K180 K184 T185 F186 D187 M191 A209 T212 E221 W228 E229 L235 A255 L256 L269

Y270 D282 I286 D287 Y288 V292 M295 Q299 D300 E305 V306 S314 E315 LEU GLN ALA ARG GLN

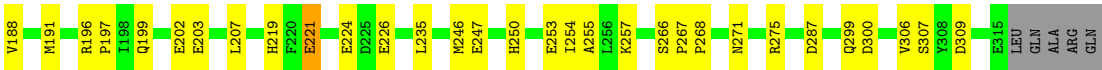
Chain D:

9% 77% 19%

LYS LEU ARG ARG LYS LYS P91 A98 Y99 D100 S101 E102 A103 S104 D105 I106 L111 I112 I117 I120 Q125 L126 E127 N131 K144 K146 N147 H150 A151 T154 V158 F159 Y160 L167 V170 I171 K175 S176 V177 D178 R179 T185 F186 D187 V188

Q199 E200 E201 E202 E203 V204 Y205 E213 D214 L215 K218 E221 E224 K231 K236 G240 A241 P242 L243 S244 D245 L254 K280 L286 E290 D294 M295 L296 Q299 E305 V306 S307 Y308 S314 E315 LEU GLN ALA ARG GLN

Chain F:



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	122.22Å 123.66Å 229.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	83.68 – 2.90 83.68 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.8 (83.68-2.90) 99.9 (83.68-2.90)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.05 (at 2.91Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.209 , 0.243 0.206 , 0.240	Depositor DCC
R_{free} test set	3936 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	70.6	Xtriage
Anisotropy	0.078	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 59.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.014 for k,h,-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	16474	wwPDB-VP
Average B, all atoms (Å ²)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.59% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, TRS

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	0.31	0/3809	0.77	5/5140 (0.1%)
1	C	0.33	0/3598	0.78	4/4861 (0.1%)
1	E	0.32	0/3884	0.75	4/5227 (0.1%)
2	B	0.32	0/1853	0.75	0/2502
2	D	0.32	0/1804	0.79	0/2439
2	F	0.31	0/1850	0.74	2/2496 (0.1%)
All	All	0.32	0/16798	0.76	15/22665 (0.1%)

There are no bond length outliers.

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	792	LEU	CA-C-N	6.34	125.84	119.24
1	C	792	LEU	C-N-CA	6.34	125.84	119.24
1	C	887	ASN	CA-C-N	5.75	125.87	119.32
1	C	887	ASN	C-N-CA	5.75	125.87	119.32
1	A	954	ASP	N-CA-C	-5.61	105.03	112.94

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-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 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	3726	0	3572	52	0
1	C	3526	0	3290	36	0
1	E	3787	0	3713	65	0
2	B	1825	0	1781	34	0
2	D	1774	0	1704	25	0
2	F	1814	0	1776	36	0
3	A	1	0	0	0	0
4	D	8	0	12	1	0
5	A	3	0	0	0	0
5	C	2	0	0	0	0
5	D	2	0	0	0	0
5	E	5	0	0	0	0
5	F	1	0	0	0	0
All	All	16474	0	15848	208	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 6.

The worst 5 of 208 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:832:MSE:HG2	1:C:861:VAL:HG12	1.56	0.88
1:E:682:ARG:NH1	2:F:98:ALA:O	2.22	0.73
1:A:682:ARG:NH1	2:B:98:ALA:O	2.21	0.72
1:E:571:ARG:NH2	1:E:595:ASP:O	2.23	0.72
2:F:202:GLU:HG3	2:F:246:MSE:HE1	1.74	0.70

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	469/506 (93%)	449 (96%)	18 (4%)	2 (0%)	30	59
1	C	454/506 (90%)	433 (95%)	21 (5%)	0	100	100
1	E	472/506 (93%)	451 (96%)	20 (4%)	1 (0%)	43	72
2	B	227/235 (97%)	220 (97%)	7 (3%)	0	100	100
2	D	224/235 (95%)	216 (96%)	8 (4%)	0	100	100
2	F	226/235 (96%)	212 (94%)	14 (6%)	0	100	100
All	All	2072/2223 (93%)	1981 (96%)	88 (4%)	3 (0%)	48	77

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	685	MSE
1	E	561	GLY
1	A	812	ILE

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	396/449 (88%)	369 (93%)	27 (7%)	14	42
1	C	360/449 (80%)	339 (94%)	21 (6%)	18	49
1	E	414/449 (92%)	385 (93%)	29 (7%)	14	40
2	B	200/209 (96%)	191 (96%)	9 (4%)	24	58
2	D	190/209 (91%)	181 (95%)	9 (5%)	23	56
2	F	201/209 (96%)	192 (96%)	9 (4%)	24	58
All	All	1761/1974 (89%)	1657 (94%)	104 (6%)	18	48

5 of 104 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	601	ASN
1	E	796	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	F	199	GLN
1	E	606	GLU
1	E	740	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	715	GLN
2	D	219	HIS
2	F	174	ASN
2	F	131	ASN
1	C	775	ASN

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
4	TRS	D	401	-	7,7,7	0.31	0	9,9,9	0.24	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	TRS	D	401	-	-	2/9/9/9	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	401	TRS	N-C-C3-O3
4	D	401	TRS	C2-C-C3-O3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	401	TRS	1	0

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	459/506 (90%)	0.25	17 (3%) 45 37	45, 76, 120, 153	3 (0%)
1	C	449/506 (88%)	0.67	50 (11%) 10 9	36, 84, 163, 191	2 (0%)
1	E	457/506 (90%)	0.27	11 (2%) 59 50	28, 69, 102, 133	8 (1%)
2	B	224/235 (95%)	0.19	6 (2%) 56 47	52, 73, 103, 149	0
2	D	220/235 (93%)	0.80	22 (10%) 12 10	34, 86, 131, 141	1 (0%)
2	F	220/235 (93%)	0.32	9 (4%) 41 33	36, 76, 110, 131	3 (1%)
All	All	2029/2223 (91%)	0.41	115 (5%) 29 23	28, 76, 128, 191	17 (0%)

The worst 5 of 115 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	736	TRP	5.4
1	C	836	PHE	5.1
2	D	315	GLU	4.9
1	C	956	ASN	4.6
1	C	932	THR	4.5

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 oligosaccharides in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
4	TRS	D	401	8/8	0.69	0.15	97,103,107,108	0
3	ZN	A	1001	1/1	0.90	0.09	134,134,134,134	0

6.5 Other polymers [i](#)

There are no such residues in this entry.