



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 08:17 AM UTC

PDB ID : 7S01 / pdb_00007s01
Title : X-ray structure of the phage AR9 non-virion RNA polymerase holoenzyme in complex with a forked oligonucleotide containing the P077 promoter
Authors : Leiman, P.G.; Sokolova, M.L.; Gordeeva, J.; Fraser, A.; Severinov, K.V.
Deposited on : 2021-08-28
Resolution : 3.40 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
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

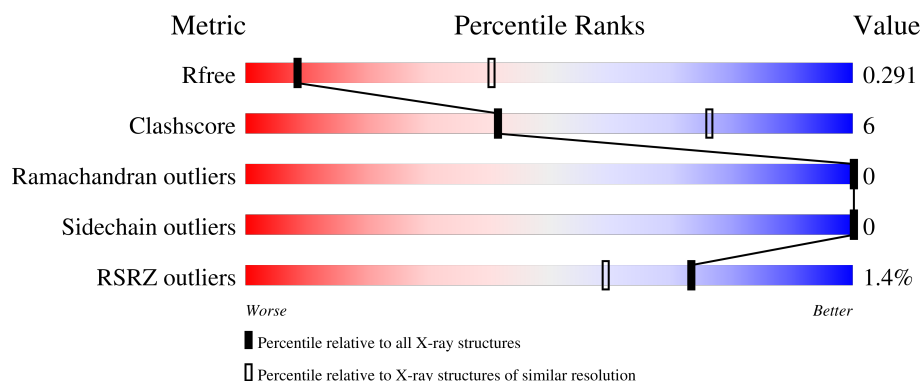
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 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	1001 (3.44-3.36)
Clashscore	190562	1022 (3.44-3.36)
Ramachandran outliers	187476	1012 (3.44-3.36)
Sidechain outliers	187428	1012 (3.44-3.36)
RSRZ outliers	180081	1001 (3.44-3.36)

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	464	% 82% 18%
2	d	426	3% 83% 17%
3	c	496	79% 18% .
4	D	631	% 82% 18%
5	C	665	2% 82% 16% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	N	14	 50% 43% 7%
6	t	14	 86% 14%
7	T	32	 62% 28% 9%
7	n	32	 56% 12% 31%

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 23345 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 DNA-directed RNA polymerase subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	463	Total	C	N	O	S	0	0	0
			3847	2478	620	729	20			

- Molecule 2 is a protein called DNA-directed RNA polymerase beta' subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	d	426	Total	C	N	O	S	0	0	0
			3488	2256	561	663	8			

- Molecule 3 is a protein called DNA-directed RNA polymerase beta subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	c	484	Total	C	N	O	S	0	0	0
			4003	2580	658	754	11			

- Molecule 4 is a protein called DNA-directed RNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	631	Total	C	N	O	S	0	0	0
			5140	3276	837	1009	18			

- Molecule 5 is a protein called DNA-directed RNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	C	649	Total	C	N	O	S	0	0	0
			5271	3337	868	1041	25			

- Molecule 6 is a DNA chain called Non-template strand of the forked DNA oligonucleotide (downstream copy).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	N	13	Total	C	N	O	P	0	0	0
			267	128	49	77	13			
6	t	14	Total	C	N	O	P	0	0	0
			289	138	54	83	14			

- Molecule 7 is a DNA chain called Template strand of the forked DNA oligonucleotide (downstream copy) containing the P077 AR9 promoter motif.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	T	29	Total	C	N	O	P	0	0	0
			589	283	99	178	29			
7	n	22	Total	C	N	O	P	0	0	0
			450	217	80	131	22			

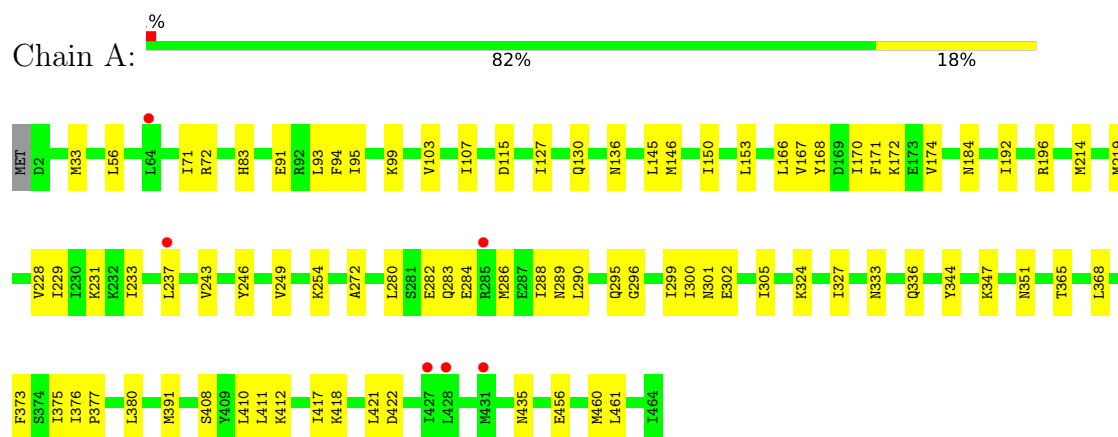
- Molecule 8 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	D	1	Total	Zn	0	0
			1	1		

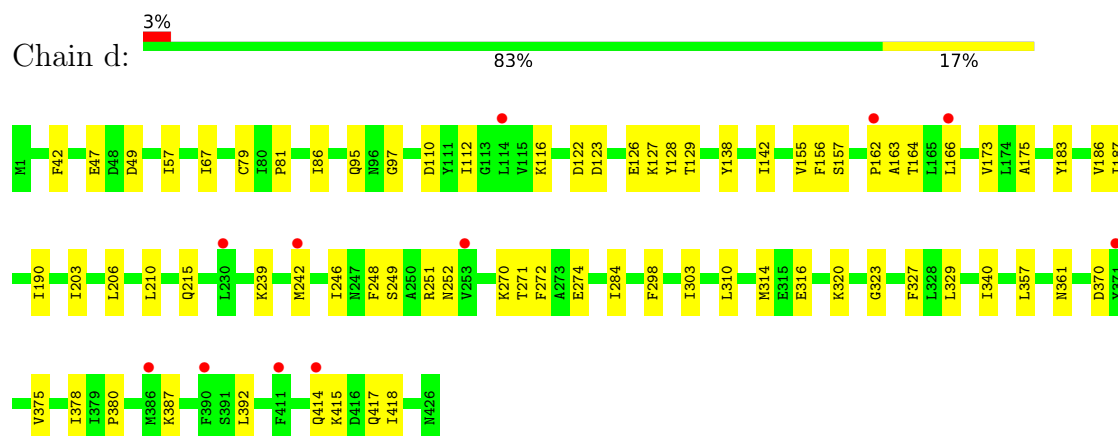
3 Residue-property plots [i](#)

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.

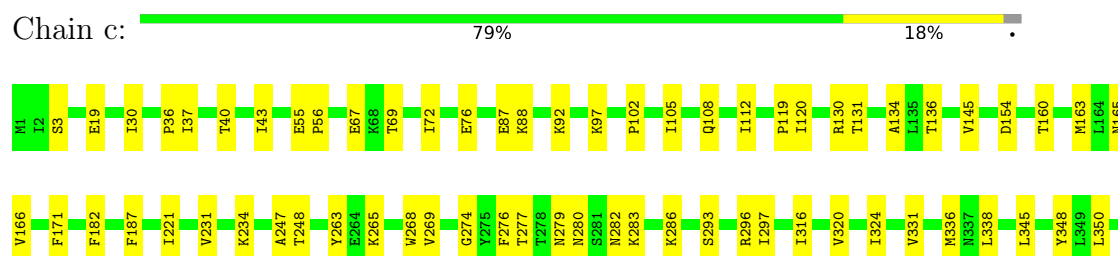
- Molecule 1: DNA-directed RNA polymerase subunit



- Molecule 2: DNA-directed RNA polymerase beta' subunit

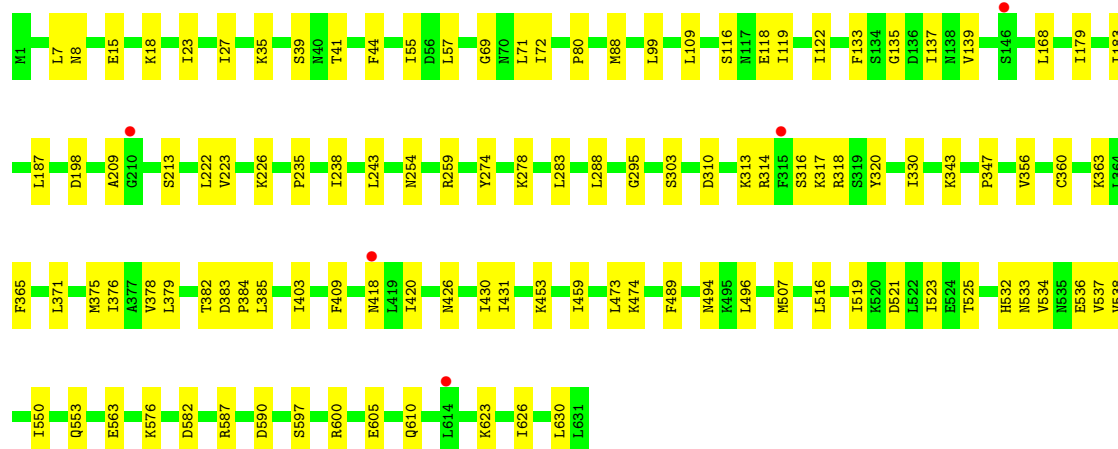
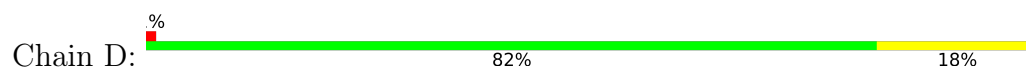


- Molecule 3: DNA-directed RNA polymerase beta subunit

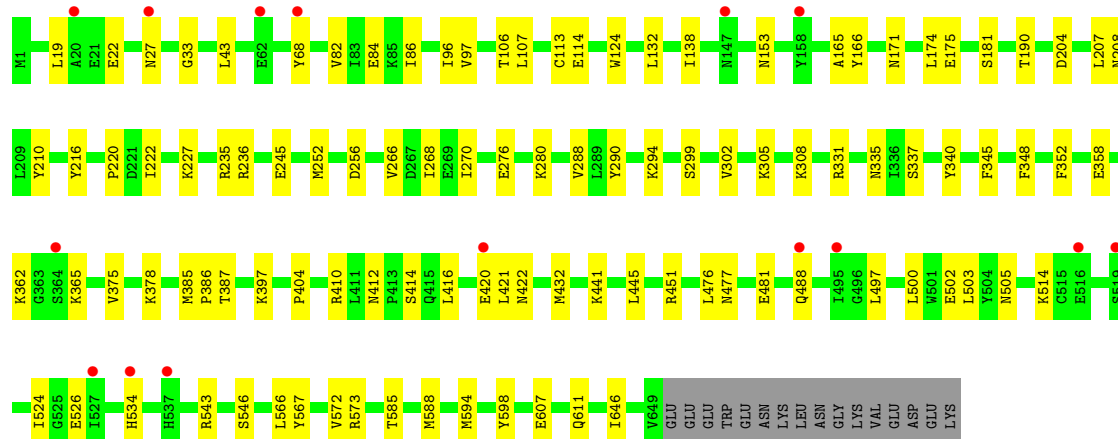
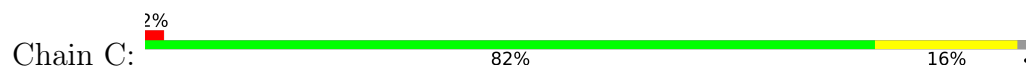




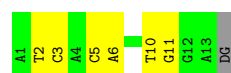
• Molecule 4: DNA-directed RNA polymerase



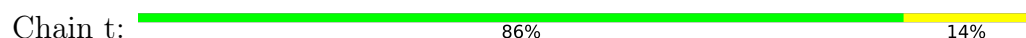
• Molecule 5: DNA-directed RNA polymerase



• Molecule 6: Non-template strand of the forked DNA oligonucleotide (downstream copy)



• Molecule 6: Non-template strand of the forked DNA oligonucleotide (downstream copy)





- Molecule 7: Template strand of the forked DNA oligonucleotide (downstream copy) containing the P077 AR9 promoter motif



- Molecule 7: Template strand of the forked DNA oligonucleotide (downstream copy) containing the P077 AR9 promoter motif



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	176.93Å 110.45Å 222.38Å 90.00° 98.52° 90.00°	Depositor
Resolution (Å)	49.35 – 3.40 49.35 – 3.40	Depositor EDS
% Data completeness (in resolution range)	96.9 (49.35-3.40) 97.0 (49.35-3.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.13 (at 3.40Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.239 , 0.293 0.239 , 0.291	Depositor DCC
R_{free} test set	2847 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	109.7	Xtriage
Anisotropy	0.316	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 112.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	23345	wwPDB-VP
Average B, all atoms (Å ²)	133.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.32% 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](#)

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

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	0.11	0/3909	0.27	0/5259
2	d	0.09	0/3558	0.26	0/4804
3	c	0.09	0/4082	0.26	0/5502
4	D	0.10	0/5222	0.27	0/7038
5	C	0.10	0/5370	0.28	0/7241
6	N	0.18	0/299	0.37	0/459
6	t	0.23	0/324	0.38	0/498
7	T	0.62	2/659 (0.3%)	0.59	0/1013
7	n	0.19	0/504	0.39	0/775
All	All	0.15	2/23927 (0.0%)	0.29	0/32589

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	T	24	DU	O3'-P	5.04	1.61	1.56
7	T	22	DA	O3'-P	5.00	1.61	1.56

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	3847	0	3926	55	0
2	d	3488	0	3485	53	0
3	c	4003	0	4045	62	0
4	D	5140	0	5155	71	0
5	C	5271	0	5153	77	0
6	N	267	0	148	3	0
6	t	289	0	159	1	0
7	T	589	0	324	4	0
7	n	450	0	251	4	0
8	D	1	0	0	0	0
All	All	23345	0	22646	295	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 295 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)
3:c:165:ASN:HD22	3:c:248:THR:HA	1.43	0.83
5:C:235:ARG:HG2	5:C:256:ASP:OD1	1.80	0.82
5:C:276:GLU:HG2	5:C:280:LYS:HE2	1.66	0.77
2:d:166:LEU:HB2	2:d:173:VAL:HG23	1.66	0.77
5:C:114:GLU:HB2	5:C:124:TRP:HE1	1.50	0.76

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	461/464 (99%)	452 (98%)	9 (2%)	0	100	100
2	d	424/426 (100%)	418 (99%)	6 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	c	482/496 (97%)	473 (98%)	9 (2%)	0	100	100
4	D	629/631 (100%)	623 (99%)	6 (1%)	0	100	100
5	C	647/665 (97%)	625 (97%)	22 (3%)	0	100	100
All	All	2643/2682 (98%)	2591 (98%)	52 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	444/445 (100%)	444 (100%)	0	100	100
2	d	390/390 (100%)	390 (100%)	0	100	100
3	c	457/469 (97%)	457 (100%)	0	100	100
4	D	590/590 (100%)	590 (100%)	0	100	100
5	C	593/608 (98%)	593 (100%)	0	100	100
All	All	2474/2502 (99%)	2474 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
4	D	82	ASN
5	C	287	GLN
4	D	394	HIS
5	C	576	ASN
5	C	126	ASN

5.3.3 RNA ⓘ

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 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	463/464 (99%)	0.31	6 (1%) 75 61	92, 151, 211, 270	0
2	d	426/426 (100%)	0.34	11 (2%) 57 42	77, 116, 183, 248	0
3	c	484/496 (97%)	0.33	2 (0%) 88 81	77, 125, 193, 259	0
4	D	631/631 (100%)	0.38	5 (0%) 82 71	78, 122, 170, 208	0
5	C	649/665 (97%)	0.43	15 (2%) 61 46	66, 101, 144, 190	0
6	N	13/14 (92%)	0.26	0 100 100	210, 230, 349, 352	0
6	t	14/14 (100%)	0.17	0 100 100	185, 246, 330, 351	0
7	T	24/32 (75%)	0.65	0 100 100	126, 233, 285, 363	0
7	n	22/32 (68%)	-0.01	0 100 100	142, 205, 261, 275	0
All	All	2726/2774 (98%)	0.36	39 (1%) 73 59	66, 121, 198, 363	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	C	488	GLN	3.5
1	A	427	ILE	3.1
2	d	253	VAL	3.0
5	C	534	HIS	3.0
3	c	446	GLY	2.9

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 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(Å ²)	Q<0.9
8	ZN	D	801	1/1	0.97	0.05	116,116,116,116	0

6.5 Other polymers [i](#)

There are no such residues in this entry.