



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 19, 2023 – 08:04 AM EST

PDB ID : 8GME
Title : Crystal structure of the gp32-Dda-dT17 complex
Authors : He, X.; Yun, M.K.; White, S.W.
Deposited on : 2023-03-24
Resolution : 4.98 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 : 4.02b-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

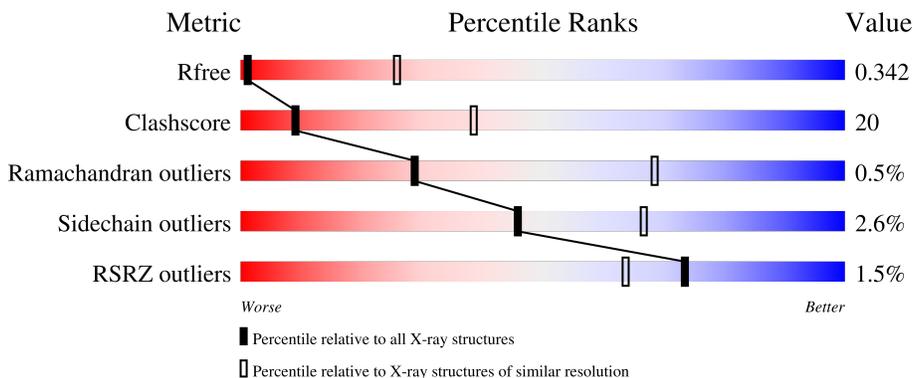
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.98 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	1138 (6.16-3.80)
Clashscore	141614	1213 (6.16-3.80)
Ramachandran outliers	138981	1144 (6.16-3.80)
Sidechain outliers	138945	1121 (6.16-3.80)
RSRZ outliers	127900	1010 (6.22-3.72)

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	301	<div style="display: flex; align-items: center;"> <div style="width: 5%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 50%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 32%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 12%; height: 10px; background-color: grey;"></div> </div> <p style="font-size: small; margin-top: 5px;">% 50% 32% 18%</p>
1	B	301	<div style="display: flex; align-items: center;"> <div style="width: 42%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 35%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 2%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 18%; height: 10px; background-color: grey;"></div> </div> <p style="font-size: small; margin-top: 5px;">42% 35% 22%</p>
2	C	459	<div style="display: flex; align-items: center;"> <div style="width: 5%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 55%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 40%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: grey;"></div> </div> <p style="font-size: small; margin-top: 5px;">% 55% 40%</p>
2	D	459	<div style="display: flex; align-items: center;"> <div style="width: 3%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 47%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 40%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 10%; height: 10px; background-color: grey;"></div> </div> <p style="font-size: small; margin-top: 5px;">3% 47% 40% 12%</p>
3	P	17	<div style="display: flex; align-items: center;"> <div style="width: 29%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 35%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 35%; height: 10px; background-color: grey;"></div> </div> <p style="font-size: small; margin-top: 5px;">29% 35% 35%</p>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	17	 18% 35% 47%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 10967 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 gp32.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	246	1939	1228	317	384	10	0	0	0
1	B	235	1870	1190	306	366	8	0	0	0

- Molecule 2 is a protein called Dda helicase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	439	3506	2267	574	653	12	0	0	0
2	D	406	3250	2099	539	601	11	0	0	0

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-19	MET	-	initiating methionine	UNP A0A6B9WEE3
C	-18	GLY	-	expression tag	UNP A0A6B9WEE3
C	-17	SER	-	expression tag	UNP A0A6B9WEE3
C	-16	SER	-	expression tag	UNP A0A6B9WEE3
C	-15	HIS	-	expression tag	UNP A0A6B9WEE3
C	-14	HIS	-	expression tag	UNP A0A6B9WEE3
C	-13	HIS	-	expression tag	UNP A0A6B9WEE3
C	-12	HIS	-	expression tag	UNP A0A6B9WEE3
C	-11	HIS	-	expression tag	UNP A0A6B9WEE3
C	-10	HIS	-	expression tag	UNP A0A6B9WEE3
C	-9	SER	-	expression tag	UNP A0A6B9WEE3
C	-8	SER	-	expression tag	UNP A0A6B9WEE3
C	-7	GLY	-	expression tag	UNP A0A6B9WEE3
C	-6	LEU	-	expression tag	UNP A0A6B9WEE3
C	-5	VAL	-	expression tag	UNP A0A6B9WEE3
C	-4	PRO	-	expression tag	UNP A0A6B9WEE3

Continued on next page...

Continued from previous page...

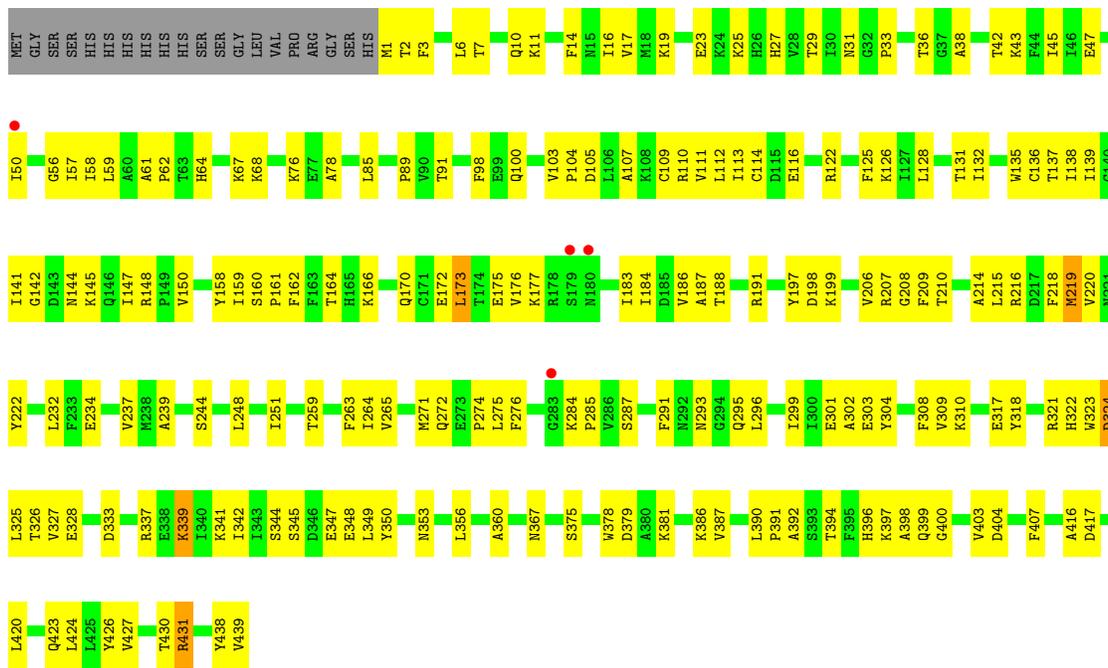
Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	ARG	-	expression tag	UNP A0A6B9WEE3
C	-2	GLY	-	expression tag	UNP A0A6B9WEE3
C	-1	SER	-	expression tag	UNP A0A6B9WEE3
C	0	HIS	-	expression tag	UNP A0A6B9WEE3
C	38	ALA	LYS	conflict	UNP A0A6B9WEE3
C	276	PHE	ILE	conflict	UNP A0A6B9WEE3
C	418	VAL	ALA	conflict	UNP A0A6B9WEE3
D	-19	MET	-	initiating methionine	UNP A0A6B9WEE3
D	-18	GLY	-	expression tag	UNP A0A6B9WEE3
D	-17	SER	-	expression tag	UNP A0A6B9WEE3
D	-16	SER	-	expression tag	UNP A0A6B9WEE3
D	-15	HIS	-	expression tag	UNP A0A6B9WEE3
D	-14	HIS	-	expression tag	UNP A0A6B9WEE3
D	-13	HIS	-	expression tag	UNP A0A6B9WEE3
D	-12	HIS	-	expression tag	UNP A0A6B9WEE3
D	-11	HIS	-	expression tag	UNP A0A6B9WEE3
D	-10	HIS	-	expression tag	UNP A0A6B9WEE3
D	-9	SER	-	expression tag	UNP A0A6B9WEE3
D	-8	SER	-	expression tag	UNP A0A6B9WEE3
D	-7	GLY	-	expression tag	UNP A0A6B9WEE3
D	-6	LEU	-	expression tag	UNP A0A6B9WEE3
D	-5	VAL	-	expression tag	UNP A0A6B9WEE3
D	-4	PRO	-	expression tag	UNP A0A6B9WEE3
D	-3	ARG	-	expression tag	UNP A0A6B9WEE3
D	-2	GLY	-	expression tag	UNP A0A6B9WEE3
D	-1	SER	-	expression tag	UNP A0A6B9WEE3
D	0	HIS	-	expression tag	UNP A0A6B9WEE3
D	38	ALA	LYS	conflict	UNP A0A6B9WEE3
D	276	PHE	ILE	conflict	UNP A0A6B9WEE3
D	418	VAL	ALA	conflict	UNP A0A6B9WEE3

- Molecule 3 is a DNA chain called dT17.

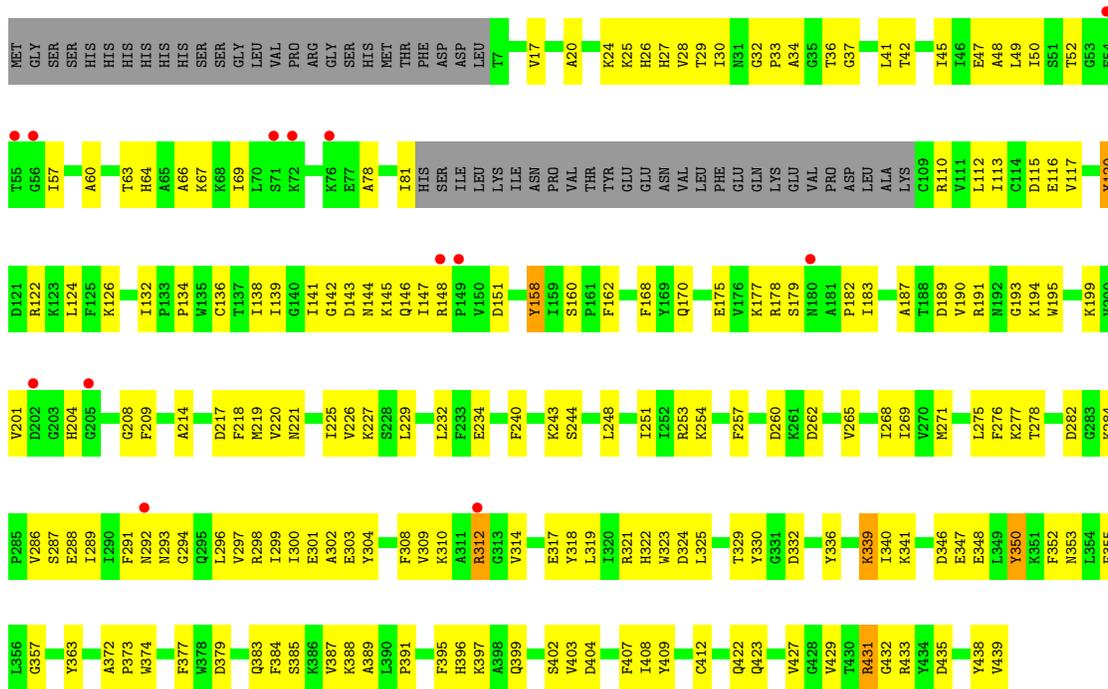
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	P	11	Total	C	N	O	P	0	0	0
			220	110	22	77	11			
3	Q	9	Total	C	N	O	P	0	0	0
			180	90	18	63	9			

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total 1	Zn 1	0	0
4	B	1	Total 1	Zn 1	0	0



• Molecule 2: Dda helicase



• Molecule 3: dT17



● Molecule 3: dT17

Chain Q:  18% 35% 47%

T600	T601	T602	T605	T606	T607	DT							
------	------	------	------	------	------	----	----	----	----	----	----	----	----

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	109.86Å 114.35Å 147.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.56 – 4.98 29.56 – 4.98	Depositor EDS
% Data completeness (in resolution range)	96.9 (29.56-4.98) 96.9 (29.56-4.98)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.43 (at 5.05Å)	Xtrriage
Refinement program	PHENIX 1.14_3260	Depositor
R, R_{free}	0.286 , 0.313 0.319 , 0.342	Depositor DCC
R_{free} test set	7815 reflections (94.93%)	wwPDB-VP
Wilson B-factor (Å ²)	149.3	Xtrriage
Anisotropy	0.428	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.21 , 135.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.23$, $\langle L^2 \rangle = 0.09$	Xtrriage
Estimated twinning fraction	0.229 for k,h,-l	Xtrriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	10967	wwPDB-VP
Average B, all atoms (Å ²)	200.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.19% 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	0.31	0/1980	0.51	0/2671
1	B	0.29	0/1910	0.49	0/2574
2	C	0.27	0/3586	0.50	0/4853
2	D	0.27	0/3323	0.50	0/4491
3	P	0.60	0/241	1.28	0/370
3	Q	0.58	0/197	1.21	0/302
All	All	0.30	0/11237	0.56	0/15261

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	1939	0	1871	67	0
1	B	1870	0	1805	89	0
2	C	3506	0	3516	147	0
2	D	3250	0	3269	141	0
3	P	220	0	133	9	0
3	Q	180	0	109	5	0
4	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	1	0	0	0	0
All	All	10967	0	10703	436	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 20.

The worst 5 of 436 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)
2:C:399:GLN:HG2	2:C:427:VAL:HG11	1.50	0.94
2:D:276:PHE:HB3	2:D:287:SER:HB3	1.50	0.92
2:C:31:ASN:HB3	2:C:172:GLU:HG2	1.61	0.82
2:D:20:ALA:HB1	2:D:26:HIS:HB3	1.61	0.80
2:C:112:LEU:HB3	2:C:138:ILE:HA	1.64	0.79

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	244/301 (81%)	229 (94%)	12 (5%)	3 (1%)	13	50
1	B	229/301 (76%)	212 (93%)	15 (7%)	2 (1%)	17	56
2	C	437/459 (95%)	412 (94%)	24 (6%)	1 (0%)	47	81
2	D	402/459 (88%)	384 (96%)	18 (4%)	0	100	100
All	All	1312/1520 (86%)	1237 (94%)	69 (5%)	6 (0%)	29	68

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	176	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	49	PRO
1	A	221	SER
1	B	49	PRO
1	A	158	GLY

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	211/258 (82%)	204 (97%)	7 (3%)	38	61
1	B	204/258 (79%)	200 (98%)	4 (2%)	55	73
2	C	382/402 (95%)	372 (97%)	10 (3%)	46	67
2	D	353/402 (88%)	344 (98%)	9 (2%)	47	68
All	All	1150/1320 (87%)	1120 (97%)	30 (3%)	46	67

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	347	GLU
2	D	339	LYS
1	B	77	CYS
2	D	431	ARG
2	D	219	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	193	ASN
1	B	199	ASN
2	D	422	GLN
2	D	64	HIS
2	D	399	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](#)

Of 2 ligands modelled in this entry, 2 are 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	246/301 (81%)	-0.47	2 (0%) 86 79	115, 155, 216, 315	0
1	B	235/301 (78%)	-0.40	1 (0%) 92 87	113, 186, 261, 301	0
2	C	439/459 (95%)	-0.32	4 (0%) 84 77	124, 204, 301, 353	0
2	D	406/459 (88%)	-0.18	13 (3%) 47 38	136, 209, 337, 366	0
3	P	11/17 (64%)	0.09	0 100 100	202, 222, 239, 239	0
3	Q	9/17 (52%)	0.27	0 100 100	176, 194, 231, 234	0
All	All	1346/1554 (86%)	-0.31	20 (1%) 73 64	113, 192, 303, 366	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	54	GLU	5.7
1	A	161	PRO	4.5
1	A	160	THR	3.8
2	D	55	THR	3.7
2	D	72	LYS	3.3

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](#)

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
4	ZN	B	500	1/1	0.97	0.09	174,174,174,174	0
4	ZN	A	500	1/1	0.98	0.11	156,156,156,156	0

6.5 Other polymers [i](#)

There are no such residues in this entry.