



wwPDB X-ray Structure Validation Summary Report

Mar 5, 2026 – 05:33 PM UTC

PDB ID : 4I2V / pdb_00004i2v
Title : X-ray structure of the unliganded uridine phosphorylase from *Yersinia pseudotuberculosis* at 2.12Å resolution
Authors : Lashkov, A.A.; Balaev, V.V.; Prokofev, I.I.; Betzel, C.; Gabdoulkhakov, A.G.; Mikhailov, A.M.
Deposited on : 2012-11-23
Resolution : 2.12 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

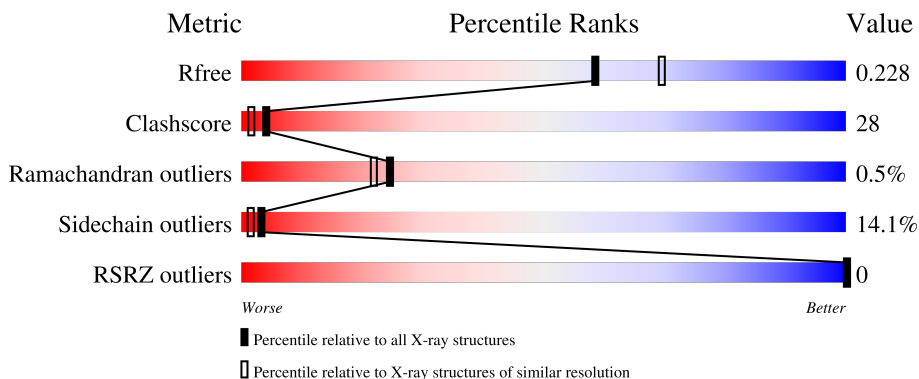
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.12 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.



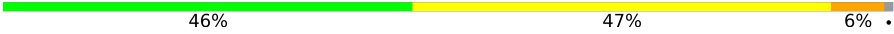
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	8290 (2.14-2.10)
Clashscore	190562	8817 (2.14-2.10)
Ramachandran outliers	187476	8738 (2.14-2.10)
Sidechain outliers	187428	8739 (2.14-2.10)
RSRZ outliers	180081	8294 (2.14-2.10)

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	253	43% 49% 7% .
1	B	253	45% 42% 12% .
1	C	253	40% 48% 11% .
1	D	253	40% 50% 8% .
1	E	253	44% 45% 10% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	253	 46% 47% 6%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11322 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 Uridine phosphorylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	250	1875	1173	330	360	12	0	0	0
1	B	250	1875	1173	330	360	12	0	0	0
1	C	250	1875	1173	330	360	12	0	0	0
1	D	250	1875	1173	330	360	12	0	0	0
1	E	250	1875	1173	330	360	12	0	0	0
1	F	250	1875	1173	330	360	12	0	0	0

- Molecule 2 is water.

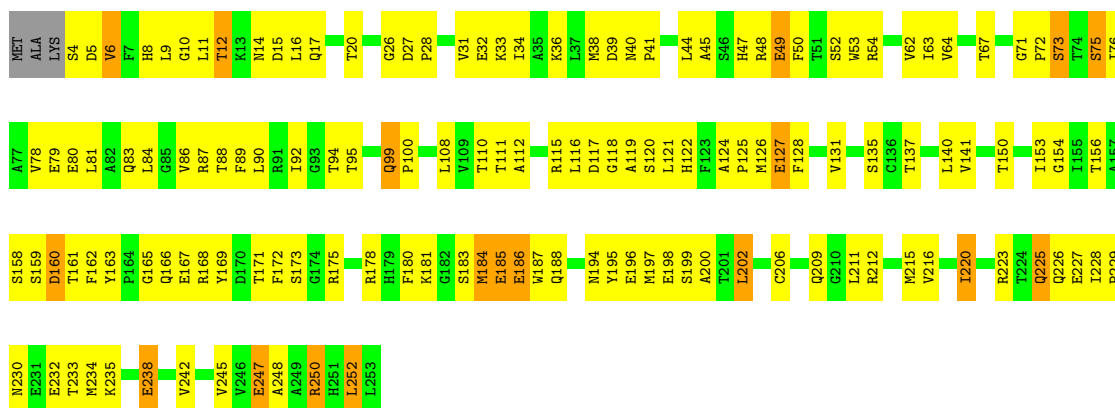
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	12	Total 12	O 12	0	0
2	B	17	Total 17	O 17	0	0
2	C	13	Total 13	O 13	0	0
2	D	17	Total 17	O 17	0	0
2	E	4	Total 4	O 4	0	0
2	F	9	Total 9	O 9	0	0

3 Residue-property plots

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.

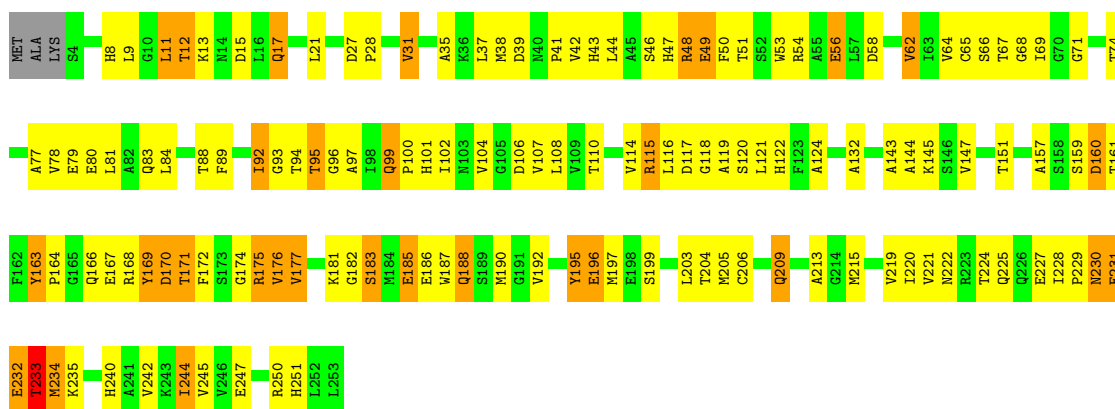
- Molecule 1: Uridine phosphorylase

Chain A: 



- Molecule 1: Uridine phosphorylase

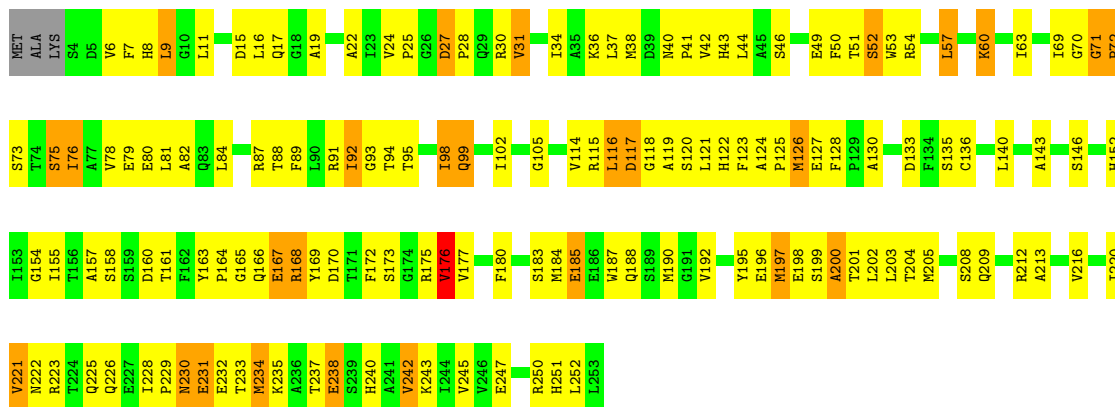
Chain B: 



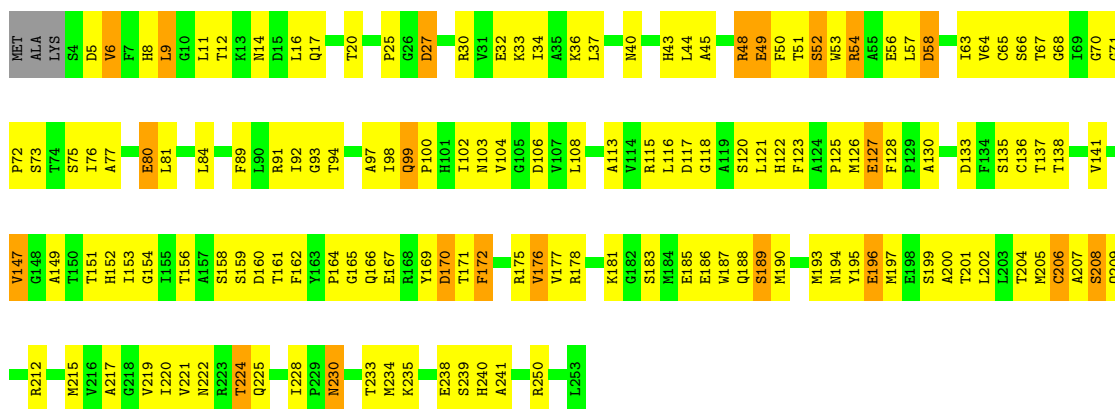
- Molecule 1: Uridine phosphorylase

Chain C: 

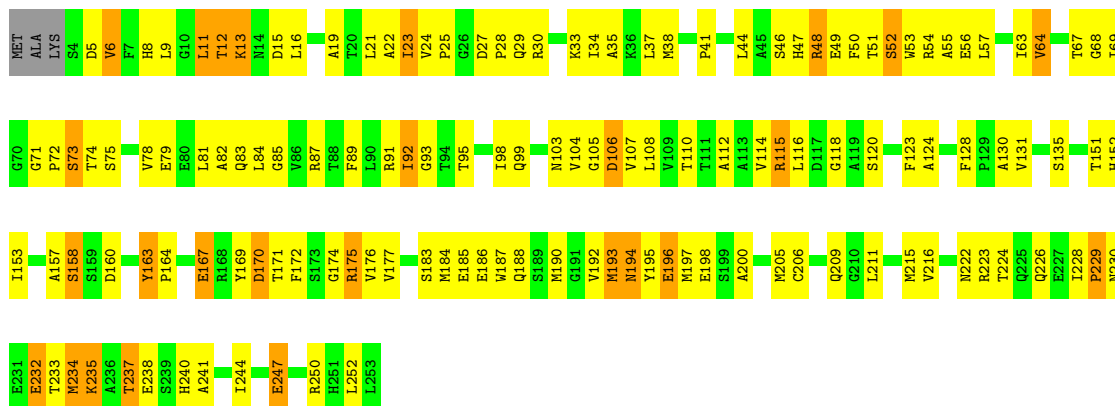


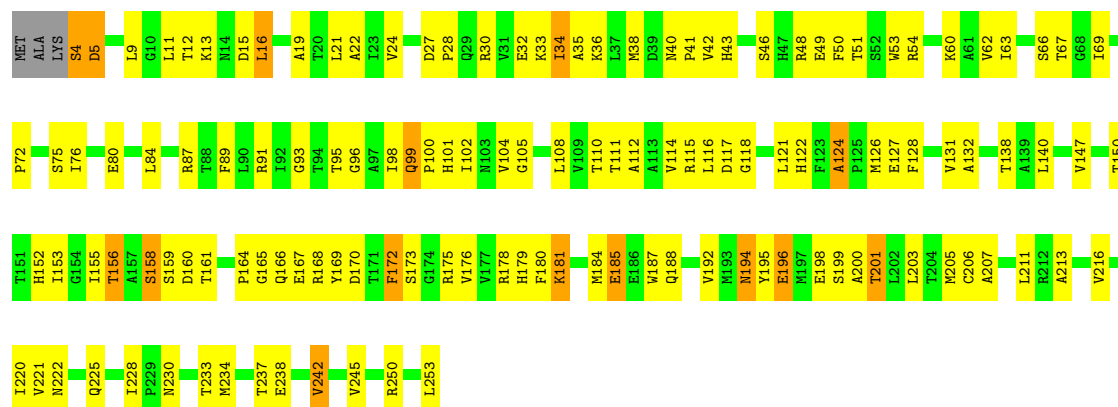


- Molecule 1: Uridine phosphorylase



- Molecule 1: Uridine phosphorylase





4 Data and refinement statistics i

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, α , β , γ	156.68Å 156.68Å 48.46Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.61 – 2.12 29.61 – 2.12	Depositor EDS
% Data completeness (in resolution range)	97.6 (29.61-2.12) 97.7 (29.61-2.12)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.71 (at 2.12Å)	Xtriage
Refinement program	PHENIX 1.8.1_1168	Depositor
R, R_{free}	0.200 , 0.242 0.202 , 0.228	Depositor DCC
R_{free} test set	3680 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	31.6	Xtriage
Anisotropy	0.246	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 16.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.41$, $\langle L^2 \rangle = 0.23$	Xtriage
Estimated twinning fraction	0.427 for -h,-k,l 0.075 for h,-h-k,-l 0.074 for -k,-h,-l	Xtriage
Reported twinning fraction	0.460 for -h,-k,l	Depositor
Outliers	0 of 73600 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	11322	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.49% 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	0.60	0/1908	1.03	6/2590 (0.2%)
1	B	0.61	0/1908	1.11	12/2590 (0.5%)
1	C	0.59	0/1908	1.12	14/2590 (0.5%)
1	D	0.60	0/1908	1.04	7/2590 (0.3%)
1	E	0.55	0/1908	1.06	13/2590 (0.5%)
1	F	0.59	0/1908	1.05	9/2590 (0.3%)
All	All	0.59	0/11448	1.07	61/15540 (0.4%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1
1	E	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 61 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	73	SER	N-CA-C	-8.64	101.81	111.14
1	C	71	GLY	CA-C-N	8.62	127.93	118.97
1	C	71	GLY	C-N-CA	8.62	127.93	118.97
1	F	124	ALA	CA-C-N	7.41	127.08	119.82
1	F	124	ALA	C-N-CA	7.41	127.08	119.82

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	232	GLU	Peptide
1	E	232	GLU	Peptide

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	1875	0	1869	123	0
1	B	1875	0	1869	108	0
1	C	1875	0	1869	116	0
1	D	1875	0	1869	124	0
1	E	1875	0	1869	101	0
1	F	1875	0	1869	97	0
2	A	12	0	0	1	0
2	B	17	0	0	0	0
2	C	13	0	0	1	0
2	D	17	0	0	4	0
2	E	4	0	0	0	0
2	F	9	0	0	1	0
All	All	11322	0	11214	620	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 28.

The worst 5 of 620 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:122:HIS:HB3	1:B:164:PRO:HB2	1.43	1.00
1:F:16:LEU:HG	1:F:63:ILE:HG13	1.48	0.93
1:D:49:GLU:HG2	1:D:68:GLY:HA2	1.55	0.88
1:E:167:GLU:HG3	1:E:183:SER:HB2	1.54	0.87
1:E:164:PRO:HB2	1:F:122:HIS:HB3	1.57	0.85

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	248/253 (98%)	228 (92%)	20 (8%)	0	100	100
1	B	248/253 (98%)	236 (95%)	11 (4%)	1 (0%)	30	28
1	C	248/253 (98%)	227 (92%)	17 (7%)	4 (2%)	7	3
1	D	248/253 (98%)	229 (92%)	18 (7%)	1 (0%)	30	28
1	E	248/253 (98%)	228 (92%)	18 (7%)	2 (1%)	16	12
1	F	248/253 (98%)	233 (94%)	15 (6%)	0	100	100
All	All	1488/1518 (98%)	1381 (93%)	99 (7%)	8 (0%)	24	22

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	126	MET
1	C	221	VAL
1	C	116	LEU
1	B	176	VAL
1	C	176	VAL

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	201/203 (99%)	175 (87%)	26 (13%)	4	2
1	B	201/203 (99%)	170 (85%)	31 (15%)	2	1
1	C	201/203 (99%)	171 (85%)	30 (15%)	3	1

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	201/203 (99%)	170 (85%)	31 (15%)	2	1
1	E	201/203 (99%)	169 (84%)	32 (16%)	2	1
1	F	201/203 (99%)	181 (90%)	20 (10%)	7	4
All	All	1206/1218 (99%)	1036 (86%)	170 (14%)	3	1

5 of 170 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	235	LYS
1	E	226	GLN
1	E	11	LEU
1	E	92	ILE
1	F	4	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	152	HIS
1	D	251	HIS
1	B	225	GLN
1	B	99	GLN
1	D	8	HIS

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](#)

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	250/253 (98%)	-1.00	0 100 100	17, 31, 50, 61	0
1	B	250/253 (98%)	-1.05	0 100 100	16, 30, 44, 55	0
1	C	250/253 (98%)	-1.05	0 100 100	18, 30, 46, 53	0
1	D	250/253 (98%)	-1.06	0 100 100	16, 31, 46, 58	0
1	E	250/253 (98%)	-1.00	0 100 100	20, 32, 53, 66	0
1	F	250/253 (98%)	-1.03	0 100 100	19, 34, 50, 59	0
All	All	1500/1518 (98%)	-1.03	0 100 100	16, 32, 50, 66	0

There are no RSRZ outliers to report.

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.