WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jun 17, 2024 – 07:39 PM EDT

PDB ID	:	3FKD
Title	:	The crystal structure of L-threonine-O-3-phosphate decarboxylase from Porphyromonas gingivalis
Authors	:	Zhang, Z.; Eswaramoorthy, S.; Burley, S.K.; Swaminathan, S.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on	:	2008-12-16
Resolution	:	2.50 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

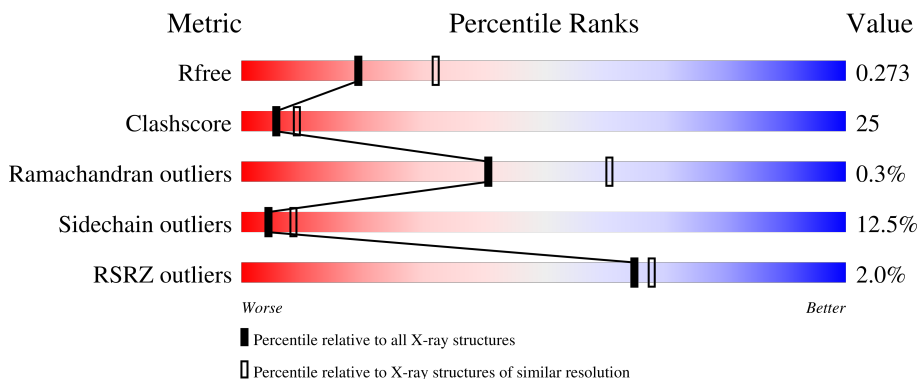|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.50 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	Percentile Ranks	Value
Rfree		0.273
Clashscore		25
Ramachandran outliers		0.3%
Sidechain outliers		12.5%
RSRZ outliers		2.0%

Worse   *Better*

▌Percentile relative to all X-ray structures

▯Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 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 <=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	350	61% 29% • 7%
1	B	350	% 55% 34% • 7%
1	C	350	% 52% 34% 7% 7%
1	D	350	6% 46% 34% 8% • 11%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 10610 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 L-threonine-O-3-phosphate decarboxylase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	326	Total	C	N	O	S	Se	0	0	0
			2624	1677	454	481	5	7			
1	B	325	Total	C	N	O	S	Se	0	0	0
			2615	1672	453	478	5	7			
1	C	325	Total	C	N	O	S	Se	0	0	0
			2607	1666	451	478	5	7			
1	D	310	Total	C	N	O	S	Se	0	0	0
			2466	1571	432	452	5	6			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	343	GLU	-	expression tag	UNP Q7MXY2
A	344	GLY	-	expression tag	UNP Q7MXY2
A	345	HIS	-	expression tag	UNP Q7MXY2
A	346	HIS	-	expression tag	UNP Q7MXY2
A	347	HIS	-	expression tag	UNP Q7MXY2
A	348	HIS	-	expression tag	UNP Q7MXY2
A	349	HIS	-	expression tag	UNP Q7MXY2
A	350	HIS	-	expression tag	UNP Q7MXY2
B	343	GLU	-	expression tag	UNP Q7MXY2
B	344	GLY	-	expression tag	UNP Q7MXY2
B	345	HIS	-	expression tag	UNP Q7MXY2
B	346	HIS	-	expression tag	UNP Q7MXY2
B	347	HIS	-	expression tag	UNP Q7MXY2
B	348	HIS	-	expression tag	UNP Q7MXY2
B	349	HIS	-	expression tag	UNP Q7MXY2
B	350	HIS	-	expression tag	UNP Q7MXY2
C	343	GLU	-	expression tag	UNP Q7MXY2
C	344	GLY	-	expression tag	UNP Q7MXY2
C	345	HIS	-	expression tag	UNP Q7MXY2
C	346	HIS	-	expression tag	UNP Q7MXY2
C	347	HIS	-	expression tag	UNP Q7MXY2

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	348	HIS	-	expression tag	UNP Q7MXY2
C	349	HIS	-	expression tag	UNP Q7MXY2
C	350	HIS	-	expression tag	UNP Q7MXY2
D	343	GLU	-	expression tag	UNP Q7MXY2
D	344	GLY	-	expression tag	UNP Q7MXY2
D	345	HIS	-	expression tag	UNP Q7MXY2
D	346	HIS	-	expression tag	UNP Q7MXY2
D	347	HIS	-	expression tag	UNP Q7MXY2
D	348	HIS	-	expression tag	UNP Q7MXY2
D	349	HIS	-	expression tag	UNP Q7MXY2
D	350	HIS	-	expression tag	UNP Q7MXY2

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	64	Total 64	O 64	0	0
2	B	75	Total 75	O 75	0	0
2	C	96	Total 96	O 96	0	0
2	D	63	Total 63	O 63	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.

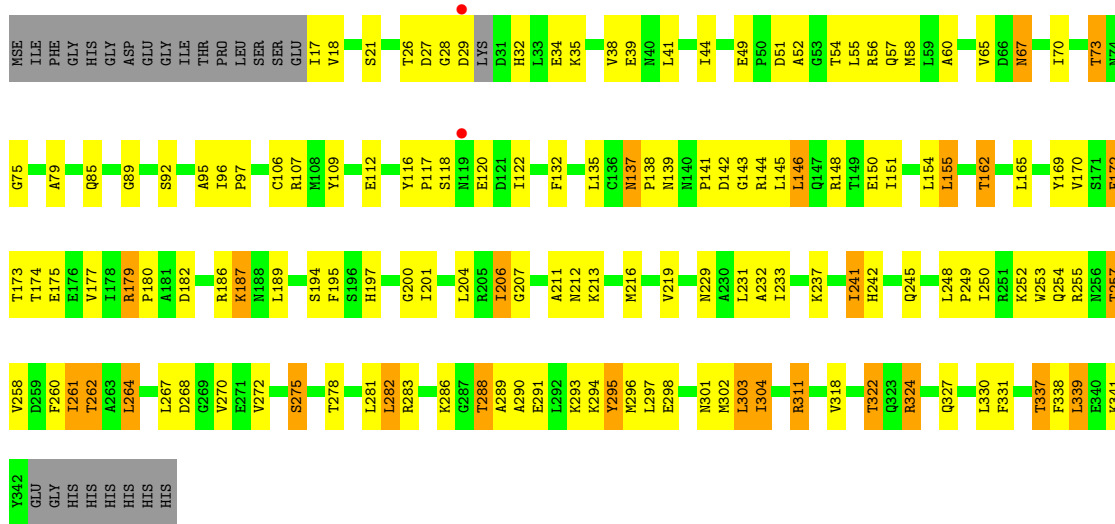- Molecule 1: L-threonine-O-3-phosphate decarboxylase
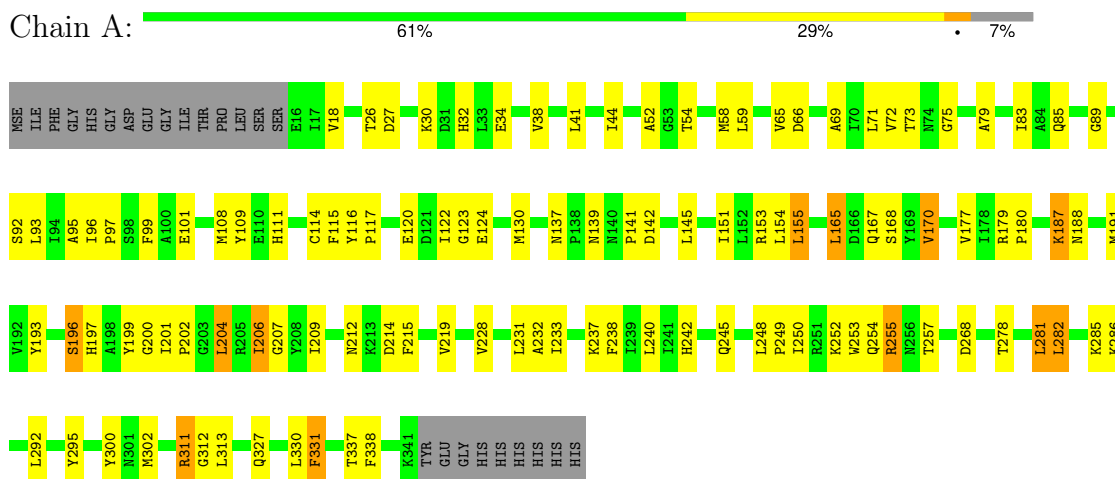
Chain A:

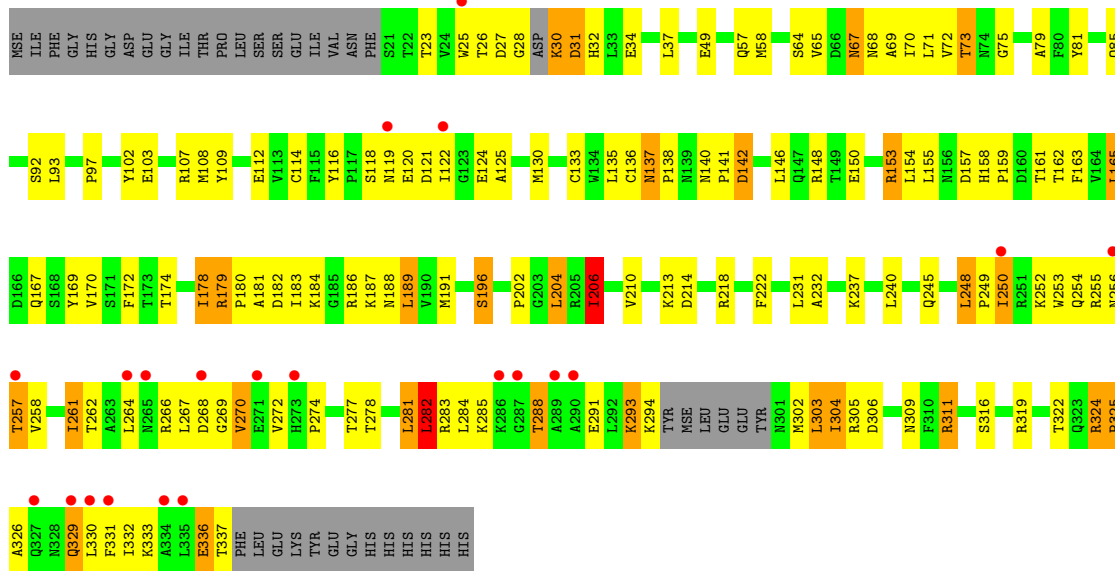- Molecule 1: L-threonine-O-3-phosphate decarboxylase

Chain B:

- Molecule 1: L-threonine-O-3-phosphate decarboxylase

Chain C:

• Molecule 1: L-threonine-O-3-phosphate decarboxylase

Chain D:
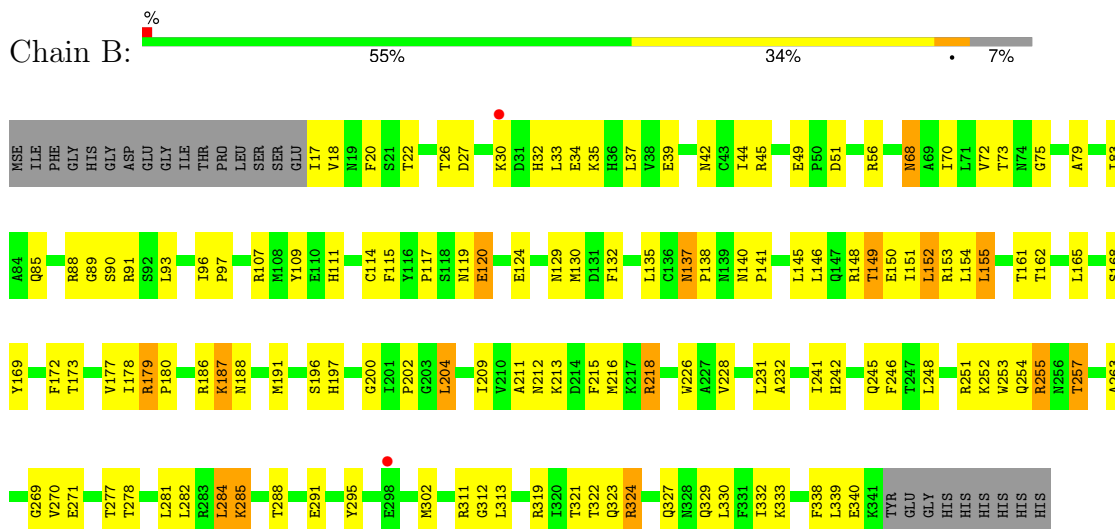
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	87.44Å 101.51Å 93.71Å 90.00° 90.22° 90.00°	Depositor
Resolution (Å)	46.86 – 2.50 93.71 – 2.50	Depositor EDS
% Data completeness (in resolution range)	98.9 (46.86-2.50) 98.9 (93.71-2.50)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	4.70 (at 2.51Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, R_{free}	0.224 , 0.275 0.223 , 0.273	Depositor DCC
R_{free} test set	2877 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	30.4	Xtriage
Anisotropy	0.918	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 23.6	EDS
L-test for twinning[2]	$< |L| > = 0.43, < L^2 > = 0.25$	Xtriage
Estimated twinning fraction	0.309 for h,-k,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	10610	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.29% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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 ⓘ

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	0.66	0/2679	0.77	0/3624
1	B	0.67	0/2670	0.76	0/3612
1	C	0.67	1/2661 (0.0%)	0.78	2/3601 (0.1%)
1	D	0.65	0/2516	0.77	3/3405 (0.1%)
All	All	0.67	1/10526 (0.0%)	0.77	5/14242 (0.0%)

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	A	0	1
1	B	0	1
1	C	0	1
All	All	0	3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	106	CYS	CB-SG	-5.07	1.73	1.81

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	282	LEU	CA-CB-CG	6.69	130.68	115.30
1	D	206	ILE	CB-CA-C	-5.87	99.86	111.60
1	C	282	LEU	CA-CB-CG	5.53	128.02	115.30
1	D	268	ASP	CB-CG-OD2	5.21	122.99	118.30
1	C	268	ASP	CB-CG-OD2	5.20	122.97	118.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	89	GLY	Peptide
1	B	89	GLY	Peptide
1	C	89	GLY	Peptide

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	2624	0	2581	115	0
1	B	2615	0	2575	118	0
1	C	2607	0	2552	133	0
1	D	2466	0	2416	167	0
2	A	64	0	0	4	0
2	B	75	0	0	4	0
2	C	96	0	0	4	0
2	D	63	0	0	5	0
All	All	10610	0	10124	518	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 25.

The worst 5 of 518 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:B:218:ARG:HH11	1:B:218:ARG:HG2	0.99	1.16
1:C:288:THR:HG22	1:C:291:GLU:H	1.14	1.12
1:A:311:ARG:HG3	1:A:311:ARG:HH11	1.06	1.11
1:C:257:THR:HB	1:C:322:THR:HG21	1.28	1.08
1:A:54:THR:HG22	1:A:58:MSE:HE3	1.35	1.08

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	324/350 (93%)	307 (95%)	17 (5%)	0	100	100
1	B	323/350 (92%)	309 (96%)	12 (4%)	2 (1%)	25	43
1	C	321/350 (92%)	307 (96%)	14 (4%)	0	100	100
1	D	304/350 (87%)	282 (93%)	20 (7%)	2 (1%)	22	39
All	All	1272/1400 (91%)	1205 (95%)	63 (5%)	4 (0%)	41	61

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	88	ARG
1	B	340	GLU
1	D	274	PRO
1	D	325	PRO

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	281/296 (95%)	263 (94%)	18 (6%)	17	33
1	B	280/296 (95%)	246 (88%)	34 (12%)	5	9
1	C	278/296 (94%)	237 (85%)	41 (15%)	3	5
1	D	262/296 (88%)	217 (83%)	45 (17%)	2	3
All	All	1101/1184 (93%)	963 (88%)	138 (12%)	4	8

5 of 138 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	189	LEU
1	D	218	ARG
1	D	293	LYS
1	B	285	LYS
1	B	284	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	82	GLN
1	D	327	GLN
1	D	85	GLN
1	D	242	HIS
1	B	245	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	319/350 (91%)	-0.18	0	100	100	16, 29, 45, 52	0
1	B	318/350 (90%)	-0.07	2 (0%)	89	90	17, 29, 47, 54	0
1	C	318/350 (90%)	-0.05	2 (0%)	89	90	18, 29, 50, 58	0
1	D	304/350 (86%)	0.34	21 (6%)	16	17	18, 34, 54, 71	0
All	All	1259/1400 (89%)	0.01	25 (1%)	65	68	16, 30, 50, 71	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	122	ILE	4.9
1	D	329	GLN	4.5
1	D	264	LEU	4.4
1	D	256	ASN	4.2
1	D	290	ALA	4.1

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.