



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:07 AM GMT

PDB ID : 2FLF
Title : Crystal structure of l-fucose-1-phosphate aldolase from *Thermus Thermophilus* HB8
Authors : Jeyakanthan, J.; Yokoyama, S.; Shiro, Y.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on : 2006-01-06
Resolution : 2.70 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
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

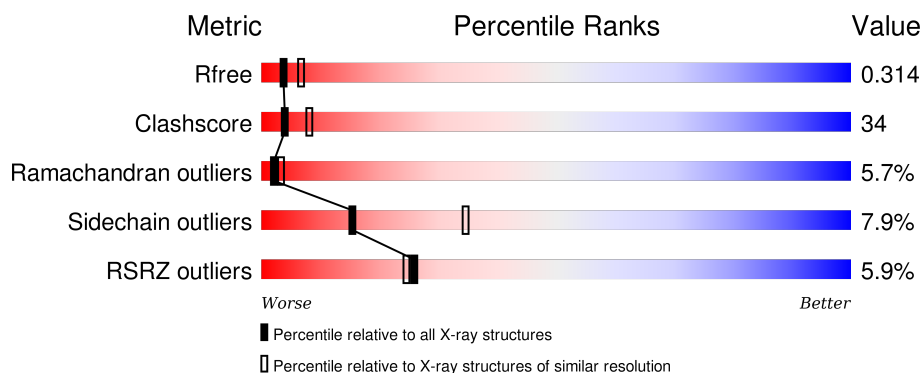
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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}	91344	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	200	<div> <div>4%</div> <div>42%</div> <div>41%</div> <div>11%</div> <div>6%</div> </div>
1	B	200	<div> <div>0%</div> <div>47%</div> <div>43%</div> <div>6%</div> <div>5%</div> </div>
1	C	200	<div> <div>2%</div> <div>49%</div> <div>43%</div> <div>5%</div> </div>
1	D	200	<div> <div>2%</div> <div>44%</div> <div>45%</div> <div>7%</div> <div>5%</div> </div>
1	E	200	<div> <div>22%</div> <div>29%</div> <div>58%</div> <div>8%</div> <div>6%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	200	<div><div></div><div>6%</div><div>36%</div><div>49%</div><div>11%</div><div>5%</div></div>
1	G	200	<div><div></div><div>2%</div><div>45%</div><div>44%</div><div>5%</div><div>6%</div></div>
1	H	200	<div><div></div><div>8%</div><div>38%</div><div>49%</div><div>7%</div><div>7%</div></div>

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 11894 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 fuculose-1-phosphate aldolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	189	Total	C	N	O	S	0	0	0
			1453	923	267	260	3			
1	B	191	Total	C	N	O	S	0	0	0
			1465	931	269	262	3			
1	C	191	Total	C	N	O	S	0	0	0
			1467	932	269	263	3			
1	D	191	Total	C	N	O	S	0	0	0
			1462	928	269	262	3			
1	E	188	Total	C	N	O	S	0	0	0
			1449	921	266	259	3			
1	F	190	Total	C	N	O	S	0	0	0
			1456	925	267	261	3			
1	G	188	Total	C	N	O	S	0	0	0
			1445	918	265	259	3			
1	H	187	Total	C	N	O	S	0	0	0
			1440	915	264	258	3			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	62	Total	O	0	0
			62	62		
2	B	47	Total	O	0	0
			47	47		
2	C	56	Total	O	0	0
			56	56		
2	D	33	Total	O	0	0
			33	33		
2	E	9	Total	O	0	0
			9	9		
2	F	22	Total	O	0	0
			22	22		

Continued on next page...

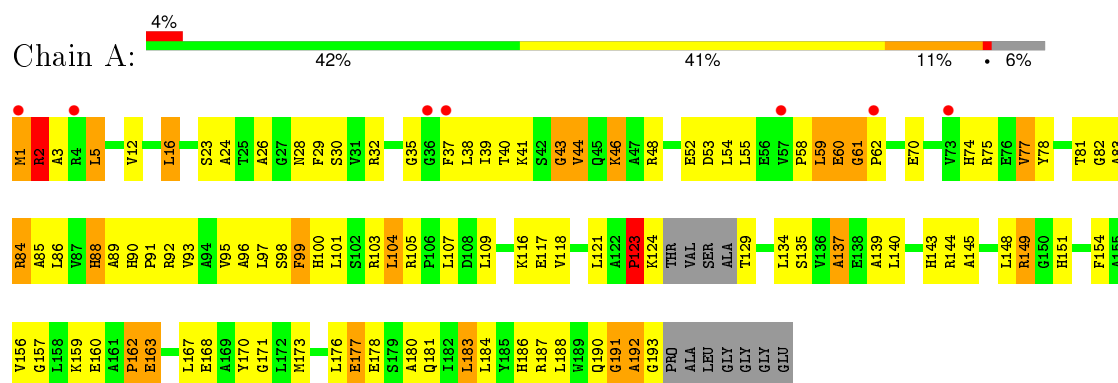
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	20	Total 20	O 20	0	0
2	H	8	Total 8	O 8	0	0

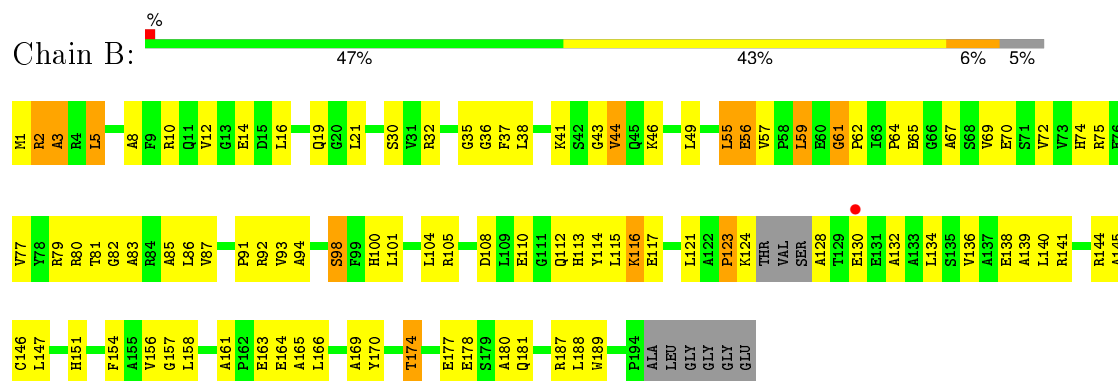
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

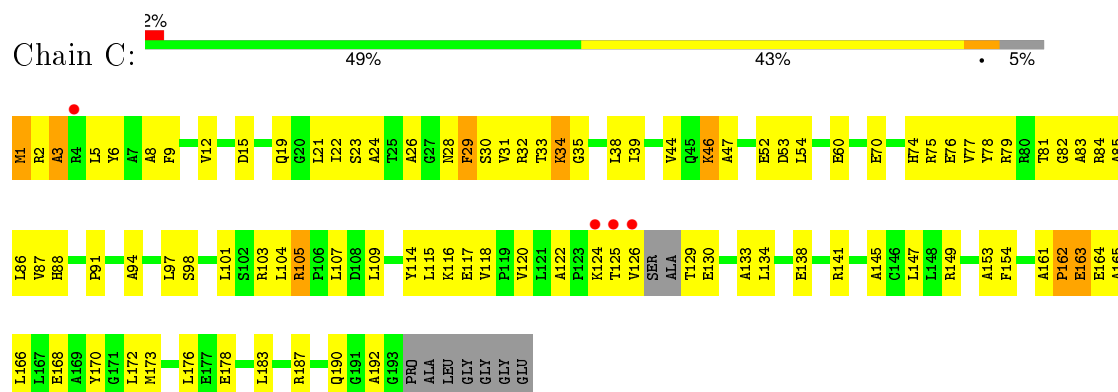
- Molecule 1: fuculose-1-phosphate aldolase



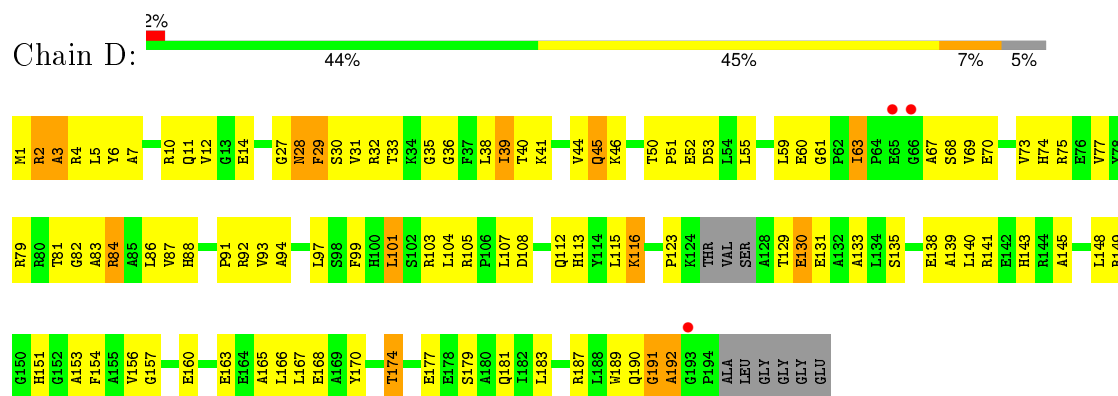
- Molecule 1: fuculose-1-phosphate aldolase

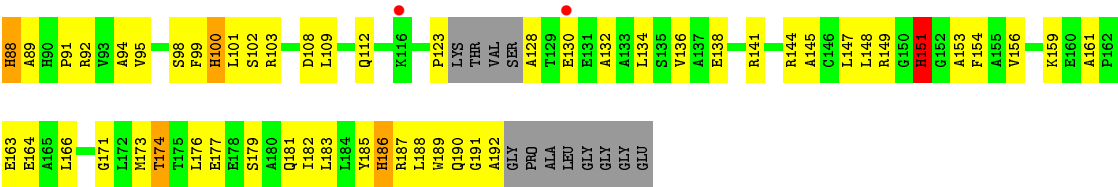


- Molecule 1: fuculose-1-phosphate aldolase

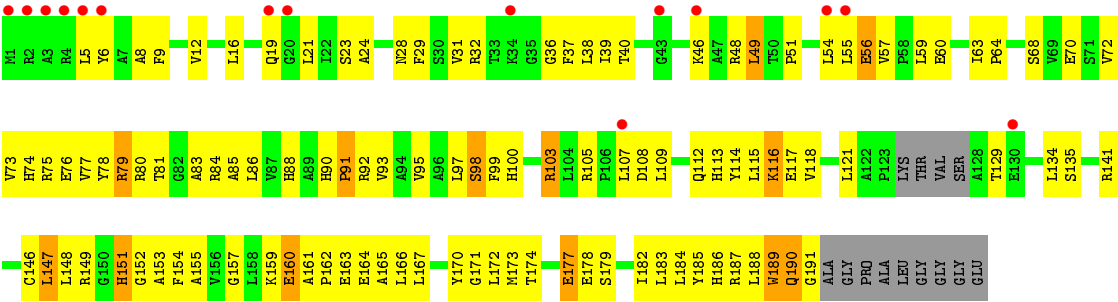


- Molecule 1: fucose-1-phosphate aldolase





● Molecule 1: fuculose-1-phosphate aldolase



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	85.41Å 101.40Å 173.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.97 – 2.70 39.36 – 2.70	Depositor EDS
% Data completeness (in resolution range)	95.5 (19.97-2.70) 95.4 (39.36-2.70)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.46 (at 2.69Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.225 , 0.318 0.223 , 0.314	Depositor DCC
R_{free} test set	2020 reflections (5.04%)	DCC
Wilson B-factor (Å ²)	52.9	Xtriage
Anisotropy	0.403	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 62.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 41911 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	11894	wwPDB-VP
Average B, all atoms (Å ²)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.84% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.46	0/1481	0.72	1/2004 (0.0%)
1	B	0.53	0/1494	0.76	2/2023 (0.1%)
1	C	0.51	0/1495	0.76	1/2024 (0.0%)
1	D	0.52	0/1490	0.75	0/2016
1	E	0.43	0/1477	0.65	2/1999 (0.1%)
1	F	0.44	0/1485	0.65	0/2012
1	G	0.47	0/1473	0.67	0/1995
1	H	0.47	1/1468 (0.1%)	0.74	4/1988 (0.2%)
All	All	0.48	1/11863 (0.0%)	0.72	10/16061 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	H	103	ARG	CG-CD	5.37	1.65	1.51

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	H	141	ARG	NE-CZ-NH2	10.26	125.43	120.30
1	C	124	LYS	N-CA-C	-8.66	87.63	111.00
1	H	103	ARG	NE-CZ-NH2	7.18	123.89	120.30
1	H	103	ARG	NE-CZ-NH1	-6.51	117.05	120.30
1	E	108	ASP	CB-CG-OD1	-5.75	113.13	118.30

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	1453	0	1480	106	0
1	B	1465	0	1492	88	0
1	C	1467	0	1496	82	0
1	D	1462	0	1485	103	0
1	E	1449	0	1477	139	0
1	F	1456	0	1479	126	0
1	G	1445	0	1469	104	0
1	H	1440	0	1464	129	0
2	A	62	0	0	0	0
2	B	47	0	0	2	0
2	C	56	0	0	4	0
2	D	33	0	0	1	0
2	E	9	0	0	1	0
2	F	22	0	0	1	0
2	G	20	0	0	0	0
2	H	8	0	0	0	0
All	All	11894	0	11842	808	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 34.

The worst 5 of 808 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:81:THR:HG22	1:B:83:ALA:H	1.09	1.10
1:B:174:THR:HG21	1:C:24:ALA:H	1.14	1.09
1:G:59:LEU:HD12	1:G:59:LEU:H	1.20	1.05
1:G:55:LEU:HD12	1:G:56:GLU:H	1.25	0.97
1:E:42:SER:HB3	1:E:68:SER:HA	1.45	0.96

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	185/200 (92%)	148 (80%)	24 (13%)	13 (7%)	1	2
1	B	187/200 (94%)	163 (87%)	16 (9%)	8 (4%)	3	7
1	C	187/200 (94%)	169 (90%)	14 (8%)	4 (2%)	9	23
1	D	187/200 (94%)	164 (88%)	13 (7%)	10 (5%)	2	4
1	E	184/200 (92%)	137 (74%)	35 (19%)	12 (6%)	1	2
1	F	186/200 (93%)	130 (70%)	35 (19%)	21 (11%)	0	0
1	G	184/200 (92%)	151 (82%)	24 (13%)	9 (5%)	3	5
1	H	183/200 (92%)	151 (82%)	24 (13%)	8 (4%)	3	6
All	All	1483/1600 (93%)	1213 (82%)	185 (12%)	85 (6%)	2	3

5 of 85 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	60	GLU
1	A	192	ALA
1	B	2	ARG
1	B	116	LYS
1	C	3	ALA

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	146/153 (95%)	130 (89%)	16 (11%)	8	18
1	B	147/153 (96%)	136 (92%)	11 (8%)	17	38
1	C	148/153 (97%)	139 (94%)	9 (6%)	23	49
1	D	146/153 (95%)	133 (91%)	13 (9%)	12	27
1	E	146/153 (95%)	136 (93%)	10 (7%)	20	43
1	F	146/153 (95%)	133 (91%)	13 (9%)	12	27

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	145/153 (95%)	134 (92%)	11 (8%)	16	37
1	H	145/153 (95%)	136 (94%)	9 (6%)	23	49
All	All	1169/1224 (96%)	1077 (92%)	92 (8%)	15	34

5 of 92 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	101	LEU
1	E	28	ASN
1	H	60	GLU
1	D	105	ARG
1	D	167	LEU

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

Mol	Chain	Res	Type
1	D	151	HIS
1	E	90	HIS
1	H	88	HIS
1	E	88	HIS
1	E	143	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 carbohydrates 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 ⓘ

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	189/200 (94%)	0.10	7 (3%) 45 45	24, 54, 90, 110	0
1	B	191/200 (95%)	-0.25	1 (0%) 91 93	19, 40, 75, 104	0
1	C	191/200 (95%)	-0.19	4 (2%) 67 68	18, 40, 76, 108	0
1	D	191/200 (95%)	-0.17	3 (1%) 74 75	19, 45, 81, 99	0
1	E	188/200 (94%)	1.13	44 (23%) 1 1	48, 90, 127, 146	0
1	F	190/200 (95%)	0.42	11 (5%) 26 25	31, 69, 108, 129	0
1	G	188/200 (94%)	0.03	4 (2%) 67 68	26, 56, 93, 117	0
1	H	187/200 (93%)	0.45	15 (8%) 15 13	42, 75, 113, 150	0
All	All	1515/1600 (94%)	0.19	89 (5%) 26 24	18, 58, 110, 150	0

The worst 5 of 89 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	128	ALA	15.0
1	E	1	MET	7.2
1	F	1	MET	6.4
1	E	51	PRO	6.1
1	E	104	LEU	5.6

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 carbohydrates 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.