Full wwPDB X-ray Structure Validation Report ⓘ

Jun 19, 2024 – 08:20 AM EDT

PDB ID	:	3WPY
Title	:	Microbacterium saccharophilum K-1 beta-fructofuranosidase mutant T47S/S 200T/F447V/P500S
Authors	:	Yokoi, G.; Mori, M.; Sato, S.; Miyazaki, T.; Nishikawa, A.; Tonozuka, T.
Deposited on	:	2014-01-17
Resolution	:	2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

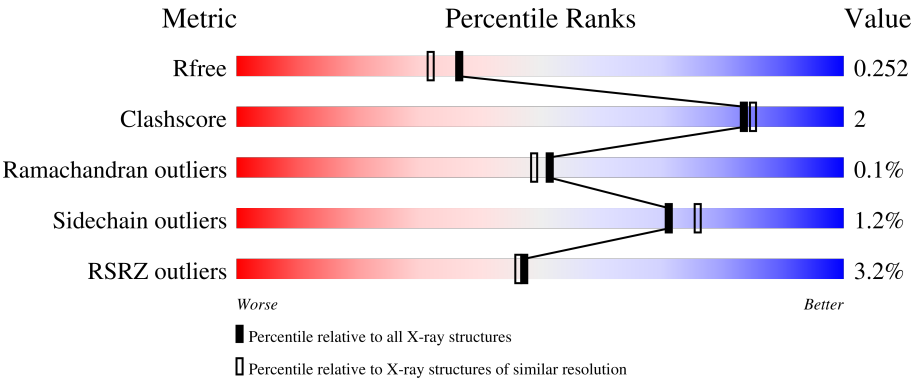
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	542	92% 7%
1	B	542	94% 5%
1	C	542	7% 91% 8%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 13374 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 Beta-fructofuranosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	538	Total	C	N	O	S	0	2	0
			4178	2644	706	818	10			
1	B	538	Total	C	N	O	S	0	1	0
			4178	2645	707	816	10			
1	C	538	Total	C	N	O	S	0	2	0
			4178	2644	706	818	10			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	47	SER	THR	engineered mutation	UNP Q8VW87
A	200	THR	SER	engineered mutation	UNP Q8VW87
A	447	VAL	PHE	engineered mutation	UNP Q8VW87
A	500	SER	PRO	engineered mutation	UNP Q8VW87
B	47	SER	THR	engineered mutation	UNP Q8VW87
B	200	THR	SER	engineered mutation	UNP Q8VW87
B	447	VAL	PHE	engineered mutation	UNP Q8VW87
B	500	SER	PRO	engineered mutation	UNP Q8VW87
C	47	SER	THR	engineered mutation	UNP Q8VW87
C	200	THR	SER	engineered mutation	UNP Q8VW87
C	447	VAL	PHE	engineered mutation	UNP Q8VW87
C	500	SER	PRO	engineered mutation	UNP Q8VW87

- Molecule 2 is water.

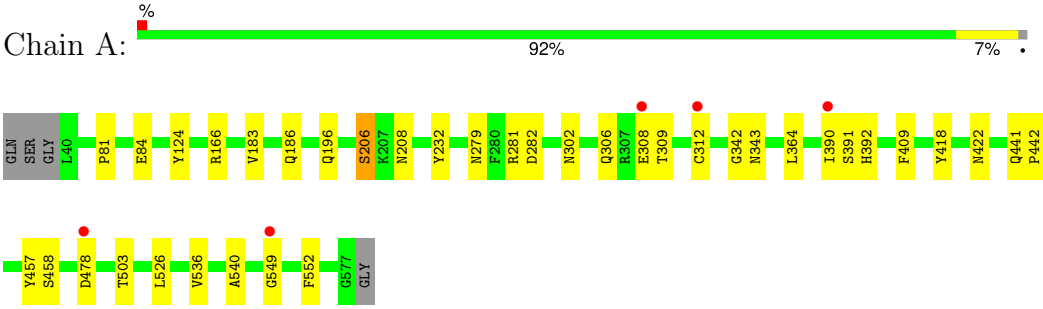
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	328	Total	O	0	0
			328	328		
2	B	309	Total	O	0	0
			309	309		
2	C	203	Total	O	0	0
			203	203		

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.

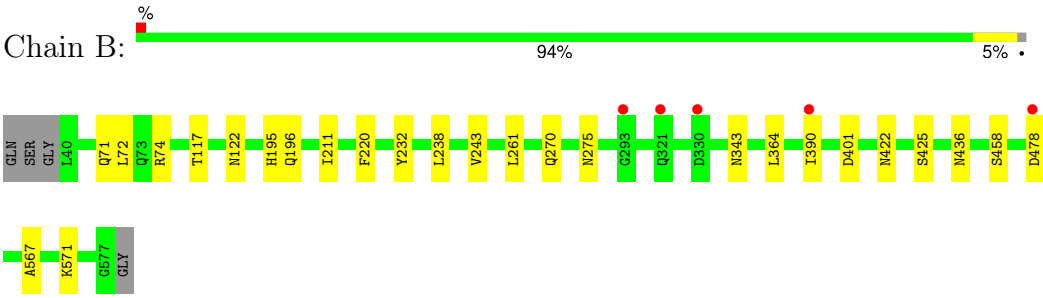
• Molecule 1: Beta-fructofuranosidase

Chain A:



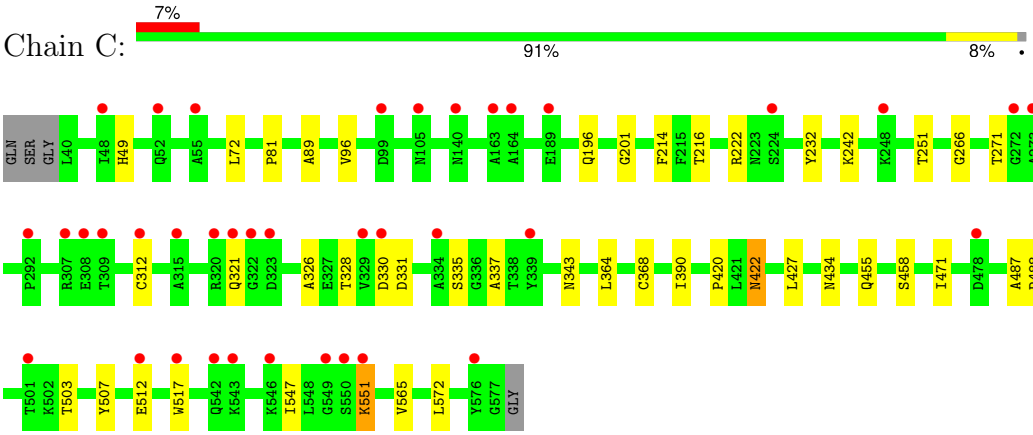
• Molecule 1: Beta-fructofuranosidase

Chain B:



• Molecule 1: Beta-fructofuranosidase

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	57.24Å 81.71Å 90.96Å 66.85° 87.11° 88.48°	Depositor
Resolution (Å)	36.13 – 2.00 36.13 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.1 (36.13-2.00) 97.1 (36.13-2.00)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.74 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.193 , 0.246 0.201 , 0.252	Depositor DCC
R_{free} test set	5012 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	21.7	Xtriage
Anisotropy	0.017	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 52.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.018 for h,-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	13374	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.51% 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	0.55	0/4298	0.72	3/5845 (0.1%)
1	B	0.52	0/4295	0.69	1/5840 (0.0%)
1	C	0.52	0/4298	0.69	0/5845
All	All	0.53	0/12891	0.70	4/17530 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	166	ARG	NE-CZ-NH2	-6.81	116.89	120.30
1	B	401	ASP	CB-CG-OD1	5.74	123.46	118.30
1	A	166	ARG	NE-CZ-NH1	5.65	123.12	120.30
1	A	282	ASP	CB-CG-OD1	5.01	122.81	118.30

There are no chirality outliers.

There are no planarity outliers.

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	4178	0	3941	21	0
1	B	4178	0	3944	13	0
1	C	4178	0	3941	25	0
2	A	328	0	0	2	0
2	B	309	0	0	2	0
2	C	203	0	0	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	13374	0	11826	59	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 2.

All (59) 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:312:CYS:HB2	2:A:661:HOH:O	1.77	0.83
1:A:536:VAL:HG13	1:A:540:ALA:HB3	1.60	0.83
1:A:422:ASN:HD21	1:A:503:THR:H	1.37	0.72
1:A:302:ASN:HD22	1:A:342:GLY:HA2	1.58	0.68
1:A:124:TYR:CD2	1:A:526:LEU:HD13	2.33	0.63
1:A:81:PRO:HG2	1:A:84:GLU:HG3	1.86	0.58
1:C:488:PRO:HD3	1:C:517:TRP:CZ3	2.38	0.57
1:C:422:ASN:HD21	1:C:503:THR:H	1.54	0.55
1:C:471:ILE:CD1	1:C:487:ALA:HB2	2.39	0.53
1:C:321:GLN:HB3	2:C:742:HOH:O	2.07	0.52
1:A:206[A]:SER:OG	1:A:208:ASN:OD1	2.28	0.52
1:B:270:GLN:HE22	1:B:275:ASN:HD22	1.59	0.50
1:A:124:TYR:HD2	1:A:526:LEU:HD13	1.72	0.50
1:A:409:PHE:HB3	1:A:418:TYR:HB3	1.94	0.50
1:C:222:ARG:HD2	2:C:690:HOH:O	2.11	0.49
1:C:89:ALA:O	1:C:507:TYR:OH	2.17	0.49
1:B:567:ALA:O	1:B:571:LYS:HG3	2.13	0.48
1:C:279:ASN:HD22	1:C:281:ARG:HH12	1.62	0.48
1:A:279:ASN:HD22	1:A:281:ARG:HH12	1.62	0.48
1:C:471:ILE:HD11	1:C:487:ALA:HB2	1.94	0.48
1:A:390:ILE:CD1	1:A:458[A]:SER:HA	2.43	0.48
1:A:536:VAL:HG13	1:A:540:ALA:CB	2.38	0.47
1:C:81:PRO:O	1:C:434:ASN:ND2	2.47	0.47
1:C:96:VAL:H	1:C:455:GLN:NE2	2.12	0.47
1:C:337:ALA:O	1:C:368:CYS:HA	2.15	0.47
1:B:390:ILE:CD1	1:B:458:SER:HA	2.45	0.47
1:B:436:ASN:ND2	2:B:861:HOH:O	2.49	0.46
1:C:216:THR:HG21	1:C:281:ARG:HB3	1.98	0.46
1:B:195:HIS:HB2	1:B:220:PHE:HB2	1.96	0.46
1:A:392:HIS:CE1	1:A:457:TYR:CZ	3.03	0.45
1:B:422:ASN:HD21	1:B:503:THR:H	1.65	0.45
1:C:266:GLY:HA2	1:C:271:THR:HG23	1.99	0.45
1:C:390:ILE:CD1	1:C:458[A]:SER:HA	2.47	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:390:ILE:CD1	1:A:458[B]:SER:HA	2.47	0.44
1:C:343:ASN:HB2	1:C:364:LEU:O	2.18	0.44
1:A:183:VAL:O	1:A:186:GLN:OE1	2.36	0.44
1:B:71:GLN:HE22	1:B:74:ARG:HH11	1.66	0.44
1:A:549:GLY:O	1:A:552:PHE:HB3	2.18	0.43
1:A:306:GLN:HB3	1:A:309:THR:OG1	2.17	0.43
1:C:242:LYS:HD2	2:C:738:HOH:O	2.17	0.43
1:B:238:LEU:HB2	1:B:261:LEU:HD11	2.00	0.43
1:C:201:GLY:HA3	1:C:214:PHE:O	2.18	0.43
1:A:343:ASN:HB2	1:A:364:LEU:O	2.19	0.42
1:C:471:ILE:HD11	1:C:487:ALA:CB	2.49	0.42
1:A:312:CYS:CB	2:A:661:HOH:O	2.53	0.42
1:B:117:THR:HA	1:B:122:ASN:O	2.19	0.42
1:A:441:GLN:HB3	1:A:442:PRO:HD2	2.01	0.42
1:B:211:ILE:HD11	1:B:243:VAL:HG23	2.01	0.42
1:C:390:ILE:CD1	1:C:458[B]:SER:HA	2.49	0.42
1:B:270:GLN:NE2	1:B:275:ASN:HD22	2.18	0.41
2:B:679:HOH:O	1:C:512:GLU:HG2	2.20	0.41
1:B:72:LEU:HD23	1:B:425:SER:HB2	2.03	0.41
1:B:343:ASN:HB2	1:B:364:LEU:O	2.21	0.41
1:C:72:LEU:HD21	1:C:427:LEU:HB2	2.03	0.41
1:C:547:ILE:HD13	1:C:572:LEU:HD11	2.02	0.41
1:C:331:ASP:O	1:C:335:SER:N	2.53	0.40
1:C:551:LYS:HG2	1:C:565:VAL:HG11	2.03	0.40
1:A:446:ASP:HB3	1:A:449:GLN:HB2	2.03	0.40
1:C:49:HIS:CD2	1:C:420:PRO:HD2	2.57	0.40

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	538/542 (99%)	517 (96%)	21 (4%)	0	100	100
1	B	537/542 (99%)	521 (97%)	16 (3%)	0	100	100
1	C	538/542 (99%)	515 (96%)	21 (4%)	2 (0%)	34	30
All	All	1613/1626 (99%)	1553 (96%)	58 (4%)	2 (0%)	51	49

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	422	ASN
1	C	326	ALA

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	438/438 (100%)	431 (98%)	7 (2%)	62	67
1	B	437/438 (100%)	434 (99%)	3 (1%)	84	88
1	C	438/438 (100%)	431 (98%)	7 (2%)	62	67
All	All	1313/1314 (100%)	1296 (99%)	17 (1%)	71	74

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	196	GLN
1	A	206[A]	SER
1	A	206[B]	SER
1	A	232	TYR
1	A	308	GLU
1	A	391	SER
1	A	478	ASP
1	B	196	GLN
1	B	232	TYR
1	B	478	ASP
1	C	196	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	232	TYR
1	C	251	THR
1	C	312	CYS
1	C	328	THR
1	C	330	ASP
1	C	551	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (24) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	52	GLN
1	A	186	GLN
1	A	228	ASN
1	A	279	ASN
1	A	302	ASN
1	A	306	GLN
1	A	422	ASN
1	A	436	ASN
1	B	71	GLN
1	B	140	ASN
1	B	223	ASN
1	B	228	ASN
1	B	270	GLN
1	B	279	ASN
1	B	340	GLN
1	B	422	ASN
1	B	436	ASN
1	C	191	GLN
1	C	228	ASN
1	C	279	ASN
1	C	422	ASN
1	C	436	ASN
1	C	455	GLN
1	C	542	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

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, 95^{th} 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	538/542 (99%)	-0.16	5 (0%)	84	83	12, 21, 38, 51	2 (0%)
1	B	538/542 (99%)	-0.15	6 (1%)	80	79	12, 23, 40, 52	1 (0%)
1	C	538/542 (99%)	0.44	40 (7%)	14	13	15, 31, 51, 63	2 (0%)
All	All	1614/1626 (99%)	0.04	51 (3%)	47	46	12, 25, 44, 63	5 (0%)

All (51) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	330	ASP	4.1
1	B	293	GLY	4.1
1	C	312	CYS	3.8
1	C	329	VAL	3.8
1	C	334	ALA	3.7
1	C	105	ASN	3.7
1	C	550	SER	3.5
1	C	551	LYS	3.3
1	C	164	ALA	3.2
1	A	312	CYS	3.2
1	C	542	GLN	3.2
1	A	390	ILE	3.2
1	C	498	GLY	3.1
1	C	495	GLY	3.0
1	B	321	GLN	2.9
1	C	320	ARG	2.9
1	C	224	SER	2.8
1	C	512	GLU	2.8
1	C	309	THR	2.8
1	C	546	LYS	2.7
1	C	321	GLN	2.7
1	C	339	TYR	2.7
1	C	323	ASP	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	99	ASP	2.6
1	B	497	GLY	2.6
1	C	517	TRP	2.6
1	C	248	LYS	2.5
1	C	52	GLN	2.4
1	C	189	GLU	2.4
1	A	478	ASP	2.4
1	C	576	TYR	2.4
1	B	390	ILE	2.4
1	C	308	GLU	2.4
1	C	48	ILE	2.4
1	C	543	LYS	2.3
1	C	478	ASP	2.3
1	C	549	GLY	2.3
1	C	315	ALA	2.3
1	C	292	PRO	2.3
1	C	273	ALA	2.2
1	B	478	ASP	2.2
1	B	330	ASP	2.2
1	C	55	ALA	2.1
1	C	501	THR	2.1
1	A	308	GLU	2.0
1	C	140	ASN	2.0
1	C	272	GLY	2.0
1	C	307	ARG	2.0
1	C	322	GLY	2.0
1	C	163	ALA	2.0
1	A	549	GLY	2.0

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.