Full wwPDB X-ray Structure Validation Report ⓘ

Nov 14, 2023 – 07:14 PM JST

PDB ID	:	6AAV
Title	:	Crystal structure of alpha-glucosyl transfer enzyme, XgtA at 1.72 angstrom resolution
Authors	:	Kurumizaka, H.; Arimura, Y.; Kirimura, K.; Watanabe, R.
Deposited on	:	2018-07-19
Resolution	:	1.72 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:

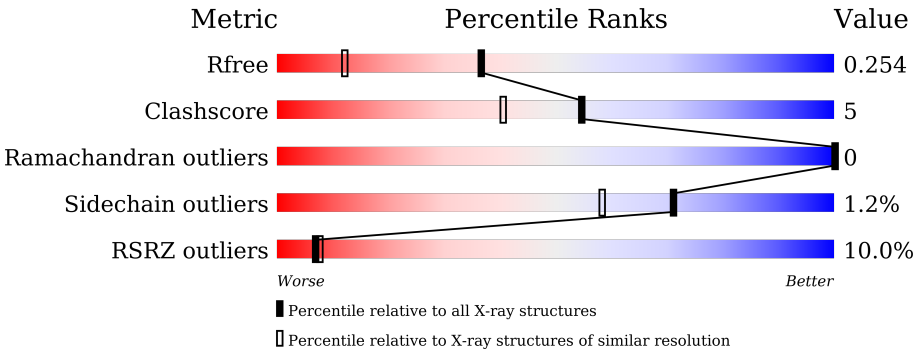
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 1.72 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	5722 (1.74-1.70)
Clashscore	141614	6152 (1.74-1.70)
Ramachandran outliers	138981	6051 (1.74-1.70)
Sidechain outliers	138945	6051 (1.74-1.70)
RSRZ outliers	127900	5629 (1.74-1.70)

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 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	544	7% 90% 8%
1	B	544	13% 81% 16%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 9068 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 Alpha-glucosyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	536	Total	C	N	O	S	0	0	0
			4233	2703	735	781	14			
1	B	535	Total	C	N	O	S	0	0	0
			4228	2701	734	779	14			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLY	-	expression tag	UNP Q76LB0
A	-4	SER	-	expression tag	UNP Q76LB0
A	-3	HIS	-	expression tag	UNP Q76LB0
A	-2	MET	-	expression tag	UNP Q76LB0
A	-1	ALA	-	expression tag	UNP Q76LB0
A	0	SER	-	expression tag	UNP Q76LB0
B	-5	GLY	-	expression tag	UNP Q76LB0
B	-4	SER	-	expression tag	UNP Q76LB0
B	-3	HIS	-	expression tag	UNP Q76LB0
B	-2	MET	-	expression tag	UNP Q76LB0
B	-1	ALA	-	expression tag	UNP Q76LB0
B	0	SER	-	expression tag	UNP Q76LB0

- Molecule 2 is water.

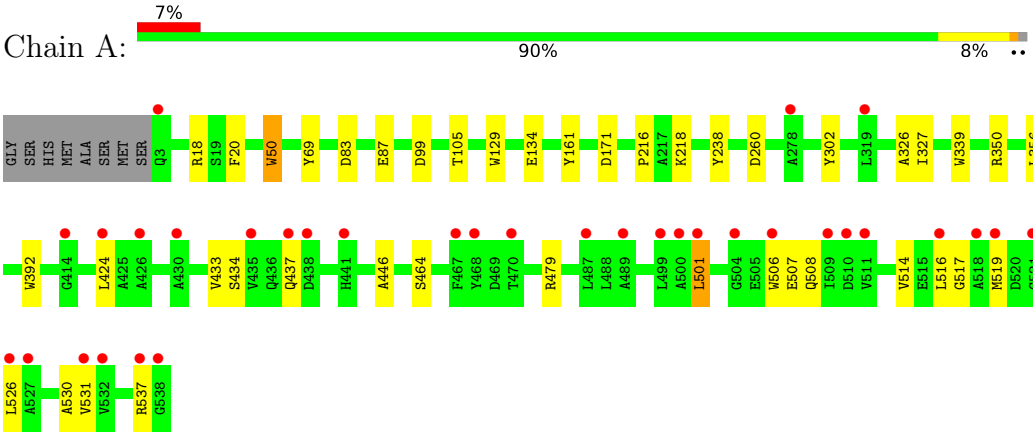
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	355	Total	O	0	0
			355	355		
2	B	252	Total	O	0	0
			252	252		

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.

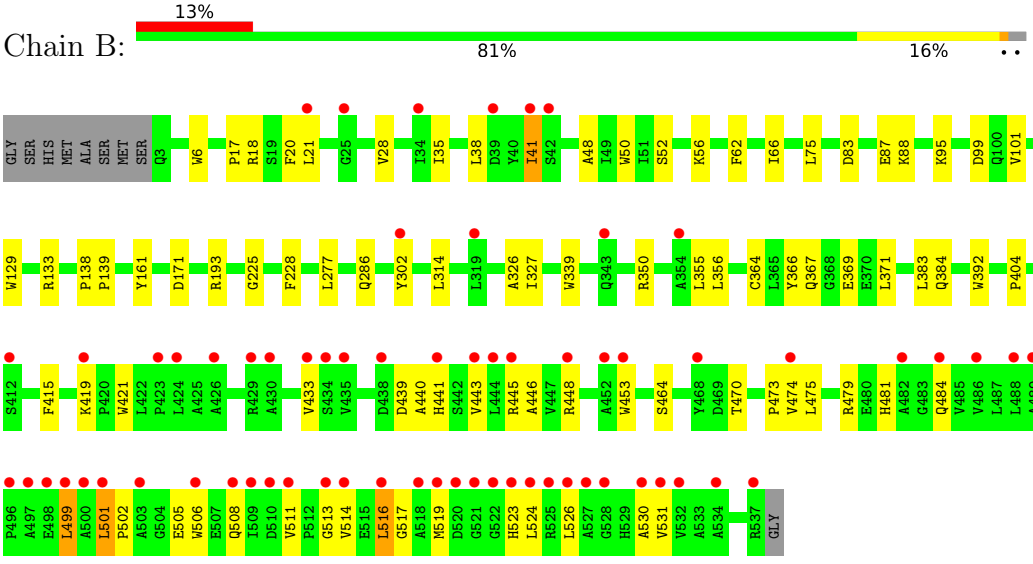
• Molecule 1: Alpha-glucosyltransferase

Chain A:



• Molecule 1: Alpha-glucosyltransferase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α, β, γ	73.07Å 83.48Å 180.78Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.86 – 1.72 48.86 – 1.72	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.86-1.72) 100.0 (48.86-1.72)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.66 (at 1.72Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.218 , 0.254 0.217 , 0.254	Depositor DCC
R_{free} test set	6005 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	27.5	Xtriage
Anisotropy	0.521	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 33.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	9068	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.34% 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	0.47	0/4361	0.65	2/5946 (0.0%)
1	B	0.45	1/4356 (0.0%)	0.67	5/5941 (0.1%)
All	All	0.46	1/8717 (0.0%)	0.66	7/11887 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	419	LYS	CD-CE	5.73	1.65	1.51

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	516	LEU	CB-CG-CD2	-10.12	93.79	111.00
1	B	505	GLU	CA-CB-CG	6.14	126.91	113.40
1	B	419	LYS	CB-CG-CD	-5.70	96.77	111.60
1	A	501	LEU	CA-CB-CG	5.66	128.32	115.30
1	B	41	ILE	CA-CB-CG1	5.47	121.39	111.00
1	A	260	ASP	CB-CG-OD1	5.46	123.21	118.30
1	B	499	LEU	CB-CG-CD2	-5.40	101.82	111.00

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	4233	0	4023	30	0
1	B	4228	0	4020	55	0
2	A	355	0	0	3	0
2	B	252	0	0	1	0
All	All	9068	0	8043	85	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 5.

All (85) 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:501:LEU:HD21	1:B:506:TRP:HD1	1.54	0.72
1:A:134:GLU:H	1:A:134:GLU:CD	1.95	0.67
1:A:83:ASP:O	1:A:87:GLU:HG3	1.95	0.66
1:B:501:LEU:HG	1:B:502:PRO:HD2	1.76	0.66
1:B:516:LEU:HD23	1:B:517:GLY:O	1.96	0.65
1:B:133:ARG:HD3	1:B:138:PRO:O	1.98	0.63
1:A:522:GLY:O	1:A:523:HIS:ND1	2.28	0.62
1:B:83:ASP:OD2	1:B:193:ARG:NH1	2.32	0.62
1:B:516:LEU:HD23	1:B:516:LEU:C	2.19	0.62
1:A:519:MET:HB3	1:A:524:LEU:HD23	1.80	0.61
1:A:216:PRO:HG2	1:A:238:TYR:HB2	1.84	0.59
1:B:327:ILE:HG12	1:B:356:LEU:HD22	1.84	0.59
1:A:433:VAL:O	1:A:437:GLN:HB3	2.04	0.58
1:A:516:LEU:HD23	1:A:517:GLY:O	2.03	0.57
1:B:501:LEU:HD21	1:B:506:TRP:CD1	2.38	0.57
1:B:28:VAL:HG11	1:B:75:LEU:HD11	1.87	0.56
1:B:516:LEU:CD2	1:B:517:GLY:O	2.55	0.55
1:A:508:GLN:HB2	1:A:519:MET:SD	2.46	0.55
1:A:18:ARG:HD2	2:A:865:HOH:O	2.06	0.54
1:B:516:LEU:HD23	1:B:517:GLY:N	2.22	0.54
1:B:440:ALA:HA	1:B:445:ARG:HG3	1.90	0.54
1:A:327:ILE:HG12	1:A:356:LEU:HD22	1.90	0.53
1:B:410:ALA:HB1	1:B:411:PRO:HD2	1.92	0.52
1:B:83:ASP:O	1:B:87:GLU:HG2	2.10	0.52
1:B:514:VAL:HG21	1:B:531:VAL:HG11	1.91	0.51
1:A:302:TYR:HB2	1:A:339:TRP:CD1	2.46	0.51
1:A:20:PHE:HB3	2:A:743:HOH:O	2.12	0.50
1:A:326:ALA:HB2	1:A:364:CYS:HB2	1.93	0.50
1:A:383:LEU:HD11	1:A:392:TRP:CE3	2.47	0.49
1:B:526:LEU:HD23	1:B:530:ALA:C	2.33	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:526:LEU:HD23	1:B:530:ALA:O	2.14	0.48
1:B:62:PHE:CZ	1:B:384:GLN:HB3	2.49	0.48
1:A:507:GLU:HG2	1:A:537:ARG:NH2	2.29	0.48
1:B:21:LEU:HD13	1:B:407:TRP:CD1	2.48	0.47
1:B:366:TYR:O	1:B:369:GLU:HG3	2.14	0.47
1:B:6:TRP:CE3	1:B:95:LYS:HG3	2.49	0.47
1:B:474:VAL:HG11	1:B:499:LEU:CD2	2.45	0.47
1:B:511:VAL:HG12	1:B:514:VAL:HG22	1.97	0.47
1:B:326:ALA:HB2	1:B:364:CYS:HB2	1.97	0.47
1:B:453:TRP:CH2	1:B:511:VAL:HG22	2.51	0.46
1:B:277:LEU:HD22	1:B:314:LEU:HD13	1.97	0.46
1:B:38:LEU:HA	1:B:41:ILE:HG13	1.97	0.45
1:A:519:MET:CB	1:A:524:LEU:HD23	2.46	0.45
1:A:464:SER:O	1:A:479:ARG:HA	2.17	0.45
1:B:355:LEU:HD11	1:B:475:LEU:HD23	1.99	0.45
1:B:35:ILE:HG23	1:B:88:LYS:HD3	1.98	0.45
1:B:383:LEU:HD11	1:B:392:TRP:CE3	2.52	0.45
1:B:350:ARG:HG3	1:B:443:VAL:HG22	1.99	0.44
1:B:499:LEU:CD2	1:B:524:LEU:HD11	2.48	0.44
1:B:508:GLN:HB2	1:B:519:MET:HE2	1.99	0.44
1:A:526:LEU:HB3	1:A:530:ALA:HB3	2.00	0.44
1:B:302:TYR:HB2	1:B:339:TRP:CD1	2.53	0.43
1:B:470:THR:OG1	1:B:474:VAL:O	2.34	0.43
1:A:508:GLN:OE1	1:A:519:MET:HG2	2.17	0.43
1:B:481:HIS:O	1:B:484:GLN:HB3	2.19	0.43
1:A:501:LEU:HD21	1:A:506:TRP:CD1	2.53	0.43
1:B:66:ILE:HD12	1:B:101:VAL:HB	2.01	0.43
1:A:350:ARG:HD2	1:A:446:ALA:CB	2.48	0.43
1:B:464:SER:O	1:B:479:ARG:HA	2.19	0.43
1:B:129:TRP:CZ3	1:B:161:TYR:HB3	2.54	0.43
1:A:50:TRP:CD1	1:A:50:TRP:C	2.93	0.42
1:B:20:PHE:HB3	2:B:714:HOH:O	2.20	0.42
1:B:519:MET:HA	1:B:523:HIS:O	2.18	0.42
1:B:367:GLN:O	1:B:367:GLN:HG3	2.20	0.42
1:B:225:GLY:HA3	1:B:228:PHE:O	2.20	0.42
1:B:371:LEU:HD21	1:B:433:VAL:HG22	2.02	0.42
1:B:404:PRO:HB3	1:B:415:PHE:CD1	2.55	0.42
1:B:56:LYS:HB3	1:B:56:LYS:HE3	1.86	0.41
1:A:218:LYS:HE2	1:A:238:TYR:CE1	2.55	0.41
1:A:507:GLU:HG2	1:A:537:ARG:HH21	1.84	0.41
1:A:434:SER:O	1:A:437:GLN:HG2	2.20	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:514:VAL:HG21	1:A:531:VAL:HB	2.03	0.41
1:A:134:GLU:CD	1:A:134:GLU:N	2.71	0.41
1:B:17:PRO:HD3	1:B:52:SER:HB2	2.03	0.41
1:B:109:HIS:CE1	1:B:111:TRP:CG	3.09	0.41
1:B:446:ALA:HB1	1:B:513:GLY:O	2.20	0.41
1:A:129:TRP:CZ3	1:A:161:TYR:HB3	2.55	0.41
1:B:48:ALA:HB2	1:B:95:LYS:HB2	2.03	0.41
1:B:138:PRO:HA	1:B:139:PRO:HD3	1.93	0.41
1:B:473:PRO:HG2	1:B:491:ASN:OD1	2.20	0.40
1:A:424:LEU:HD12	2:A:791:HOH:O	2.20	0.40
1:B:404:PRO:HB3	1:B:415:PHE:CG	2.56	0.40
1:B:18:ARG:NH1	1:B:421:TRP:CD2	2.90	0.40
1:B:439:ASP:OD2	1:B:441:HIS:HB2	2.21	0.40
1:A:69:TYR:OH	1:A:105:THR:HG22	2.22	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	534/544 (98%)	525 (98%)	9 (2%)	0	100	100
1	B	533/544 (98%)	523 (98%)	10 (2%)	0	100	100
All	All	1067/1088 (98%)	1048 (98%)	19 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	431/437 (99%)	427 (99%)	4 (1%)	78	69
1	B	431/437 (99%)	425 (99%)	6 (1%)	67	52
All	All	862/874 (99%)	852 (99%)	10 (1%)	71	58

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	50	TRP
1	A	99	ASP
1	A	171	ASP
1	A	523	HIS
1	B	50	TRP
1	B	99	ASP
1	B	171	ASP
1	B	286	GLN
1	B	448	ARG
1	B	501	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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)

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, 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	536/544 (98%)	0.48	37 (6%) 16 19	19, 33, 66, 83	0
1	B	535/544 (98%)	0.96	70 (13%) 3 3	19, 42, 74, 91	0
All	All	1071/1088 (98%)	0.72	107 (9%) 7 8	19, 36, 71, 91	0

All (107) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	521	GLY	6.2
1	B	526	LEU	6.1
1	B	511	VAL	6.0
1	B	482	ALA	5.9
1	B	499	LEU	5.9
1	B	441	HIS	5.4
1	B	433	VAL	5.3
1	B	524	LEU	5.3
1	B	521	GLY	5.2
1	A	3	GLN	5.2
1	B	438	ASP	5.1
1	A	524	LEU	5.0
1	B	510	ASP	5.0
1	A	522	GLY	4.9
1	B	522	GLY	4.8
1	B	429	ARG	4.7
1	B	513	GLY	4.7
1	B	519	MET	4.7
1	B	41	ILE	4.5
1	A	501	LEU	4.4
1	A	518	ALA	4.3
1	B	532	VAL	4.3
1	B	492	LEU	4.2
1	B	434	SER	4.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	319	LEU	4.0
1	B	523	HIS	4.0
1	B	426	ALA	4.0
1	A	509	ILE	4.0
1	B	530	ALA	4.0
1	B	525	ARG	3.9
1	B	531	VAL	3.9
1	B	503	ALA	3.8
1	B	474	VAL	3.8
1	B	302	TYR	3.7
1	A	526	LEU	3.7
1	B	506	TRP	3.7
1	B	319	LEU	3.7
1	B	516	LEU	3.6
1	B	528	GLY	3.6
1	B	39	ASP	3.5
1	B	486	VAL	3.4
1	B	343	GLN	3.4
1	A	499	LEU	3.4
1	A	523	HIS	3.3
1	B	443	VAL	3.3
1	B	25	GLY	3.3
1	B	520	ASP	3.3
1	B	497	ALA	3.3
1	A	487	LEU	3.2
1	B	527	ALA	3.2
1	B	501	LEU	3.2
1	A	531	VAL	3.2
1	B	354	ALA	3.1
1	B	444	LEU	3.1
1	B	468	TYR	3.1
1	A	426	ALA	3.1
1	B	518	ALA	3.1
1	B	534	ALA	3.1
1	B	537	ARG	3.1
1	A	511	VAL	3.1
1	B	484	GLN	3.0
1	B	424	LEU	3.0
1	B	407	TRP	3.0
1	A	414	GLY	2.9
1	B	435	VAL	2.9
1	A	537	ARG	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	411	PRO	2.8
1	B	500	ALA	2.8
1	B	34	ILE	2.8
1	B	488	LEU	2.8
1	A	438	ASP	2.8
1	B	514	VAL	2.7
1	B	452	ALA	2.7
1	B	430	ALA	2.6
1	A	441	HIS	2.6
1	B	445	ARG	2.6
1	B	453	TRP	2.5
1	A	430	ALA	2.5
1	B	498	GLU	2.5
1	A	510	ASP	2.5
1	B	496	PRO	2.5
1	B	490	PHE	2.5
1	B	509	ILE	2.5
1	B	448	ARG	2.4
1	A	468	TYR	2.4
1	A	424	LEU	2.4
1	A	500	ALA	2.4
1	A	470	THR	2.4
1	A	504	GLY	2.4
1	A	516	LEU	2.4
1	A	532	VAL	2.4
1	B	419	LYS	2.4
1	A	506	TRP	2.3
1	B	412	SER	2.3
1	B	508	GLN	2.2
1	B	489	ALA	2.2
1	A	538	GLY	2.2
1	A	519	MET	2.2
1	A	489	ALA	2.2
1	A	527	ALA	2.2
1	A	467	PHE	2.1
1	B	42	SER	2.1
1	B	423	PRO	2.1
1	A	278	ALA	2.1
1	A	435	VAL	2.0
1	B	21	LEU	2.0
1	A	437	GLN	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.