Full wwPDB X-ray Structure Validation Report ⓘ

Aug 29, 2023 – 08:09 PM JST

PDB ID	:	8GUM
Title	:	Chitin-active AA10 LPMO (GbpA) from Vibrio campbellii
Authors	:	Zhou, Y.; Robinson, R.C.; Suginta, W.
Deposited on	:	2022-09-12
Resolution	:	2.35 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:

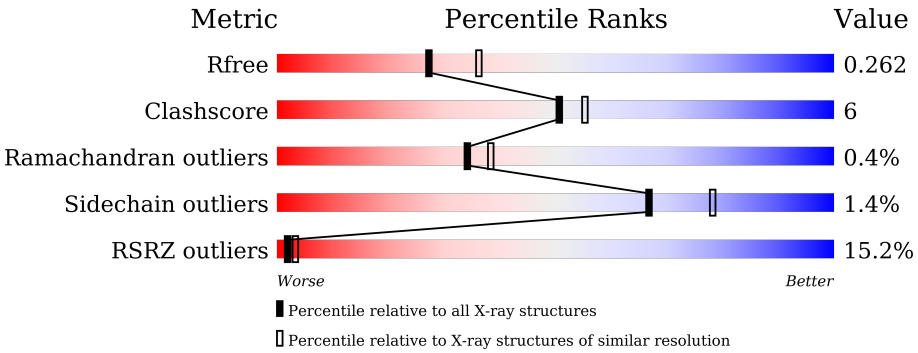
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.35 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	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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	464	11% 86% 14%
1	B	464	16% 65% 12% · 22%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6514 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 GlcNAc-binding protein A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	464	Total	C	N	O	S	0	0	0
			3607	2260	601	734	12			
1	B	360	Total	C	N	O	S	0	0	0
			2795	1741	480	564	10			

- Molecule 2 is water.

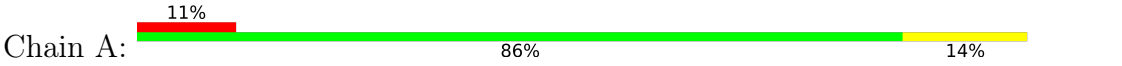
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	57	Total	O	0	0
			57	57		
2	B	55	Total	O	0	0
			55	55		

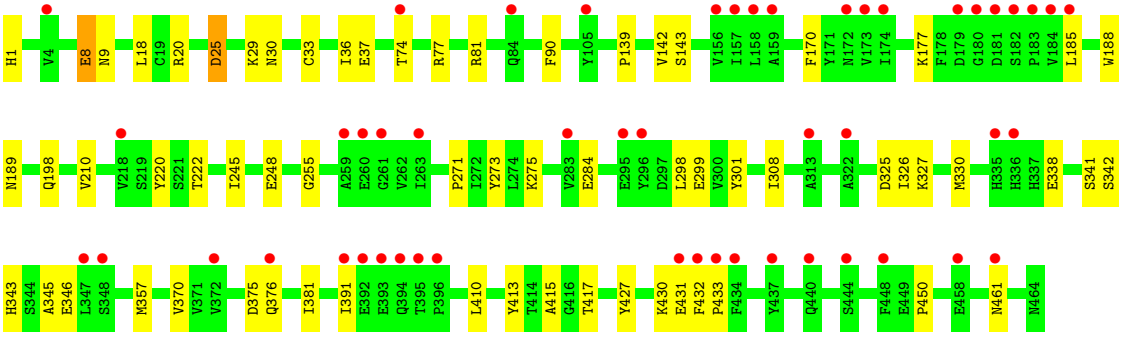
3 Residue-property plots (i)

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: GlcNAc-binding protein A

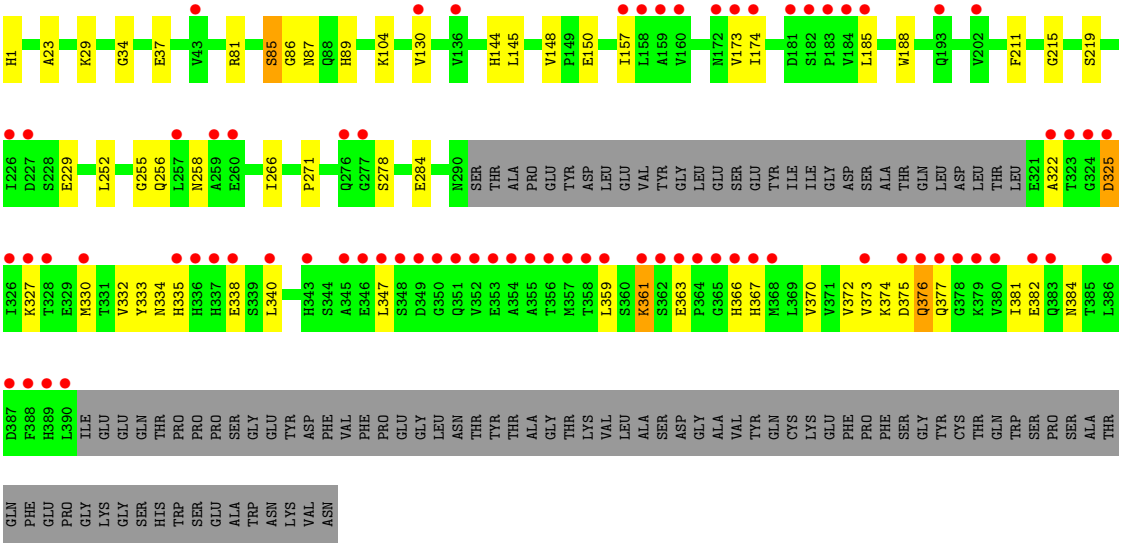
Chain A:





• Molecule 1: GlcNAc-binding protein A

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	113.61Å 113.61Å 215.75Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.10 – 2.35 25.10 – 2.34	Depositor EDS
% Data completeness (in resolution range)	99.6 (25.10-2.35) 99.3 (25.10-2.34)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.52 (at 2.33Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.220 , 0.262 0.220 , 0.262	Depositor DCC
R_{free} test set	2926 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å^2)	58.0	Xtriage
Anisotropy	0.099	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 51.1	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	0.95	EDS
Total number of atoms	6514	wwPDB-VP
Average B, all atoms (Å^2)	90.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.78% 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	0.48	0/3700	0.62	0/5048
1	B	0.47	0/2860	0.63	0/3895
All	All	0.47	0/6560	0.63	0/8943

There are no bond length outliers.

There are no bond angle outliers.

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	3607	0	3364	40	0
1	B	2795	0	2628	34	0
2	A	57	0	0	4	0
2	B	55	0	0	1	0
All	All	6514	0	5992	71	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 6.

All (71) 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:210:VAL:HG21	1:A:220:TYR:HB2	1.48	0.92

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:327:LYS:HG2	1:A:346:GLU:HG2	1.52	0.90
1:B:327:LYS:HG3	1:B:376:GLN:HA	1.65	0.78
1:A:222:THR:HG21	1:A:245:ILE:HA	1.74	0.69
1:A:248:GLU:OE1	2:A:501:HOH:O	2.09	0.69
1:B:322:ALA:HB2	1:B:347:LEU:HD23	1.76	0.68
1:A:29:LYS:NZ	2:A:503:HOH:O	2.27	0.65
1:B:325:ASP:HB3	1:B:376:GLN:HG3	1.79	0.64
1:A:25:ASP:OD1	1:A:25:ASP:N	2.26	0.64
1:A:430:LYS:HD3	1:A:461:ASN:OD1	1.99	0.61
2:A:503:HOH:O	1:B:29:LYS:NZ	2.35	0.60
1:A:284:GLU:HG3	1:B:1:HIS:CG	2.36	0.59
1:B:173:VAL:O	2:B:501:HOH:O	2.16	0.57
1:A:415:ALA:HB1	1:A:431:GLU:HG2	1.86	0.57
1:B:256:GLN:HG3	1:B:266:ILE:HD13	1.87	0.56
1:B:334:ASN:HA	1:B:367:HIS:HA	1.86	0.56
1:A:375:ASP:HB3	1:A:381:ILE:HD11	1.89	0.55
1:B:332:VAL:HB	1:B:340:LEU:HB2	1.89	0.55
1:A:308:ILE:HG12	1:A:391:ILE:HD11	1.90	0.54
1:A:325:ASP:HB3	1:A:376:GLN:HG2	1.90	0.53
1:A:198:GLN:NE2	2:A:510:HOH:O	2.38	0.52
1:B:34:GLY:O	1:B:37:GLU:HG2	2.10	0.51
1:B:185:LEU:HB3	1:B:188:TRP:HD1	1.75	0.50
1:B:255:GLY:O	1:B:271:PRO:HG2	2.11	0.50
1:B:332:VAL:HG11	1:B:359:LEU:HD21	1.93	0.49
1:A:255:GLY:O	1:A:271:PRO:HG2	2.12	0.49
1:B:330:MET:HG2	1:B:370:VAL:O	2.12	0.49
1:A:330:MET:HA	1:A:370:VAL:O	2.12	0.48
1:A:341:SER:OG	1:A:357:MET:HB3	2.13	0.48
1:B:361:LYS:HD2	1:B:363:GLU:OE2	2.13	0.48
1:A:139:PRO:HD2	1:A:142:VAL:CG1	2.44	0.47
1:B:327:LYS:O	1:B:374:LYS:HB2	2.14	0.47
1:A:8:GLU:HB2	1:A:9:ASN:H	1.50	0.47
1:A:20:ARG:HB2	1:A:37:GLU:O	2.15	0.47
1:B:211:PHE:CE1	1:B:215:GLY:HA2	2.50	0.47
1:A:25:ASP:OD1	1:A:81:ARG:NH2	2.48	0.46
1:B:89:HIS:HA	1:B:144:HIS:O	2.16	0.46
1:A:1:HIS:HA	1:A:170:PHE:CZ	2.51	0.45
1:A:298:LEU:HD13	1:A:326:ILE:HD11	1.99	0.45
1:A:413:TYR:CD1	1:A:417:THR:HG21	2.52	0.45
1:A:427:TYR:CD1	1:A:450:PRO:HB3	2.52	0.45
1:B:325:ASP:HB3	1:B:376:GLN:CG	2.46	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:18:LEU:HB2	1:A:30:ASN:ND2	2.32	0.44
1:A:427:TYR:CG	1:A:450:PRO:HB3	2.52	0.44
1:A:330:MET:O	1:A:342:SER:HA	2.17	0.44
1:B:104:LYS:HG2	1:B:130:VAL:HG22	1.99	0.44
1:A:410:LEU:HA	1:A:413:TYR:CE2	2.53	0.44
1:A:1:HIS:CG	1:B:284:GLU:HG3	2.53	0.43
1:A:284:GLU:HG3	1:B:1:HIS:CD2	2.54	0.43
1:B:87:ASN:ND2	1:B:145:LEU:HD21	2.34	0.43
1:B:335:HIS:HB2	1:B:366:HIS:HB3	2.00	0.43
1:A:299:GLU:HG2	1:A:301:TYR:CZ	2.54	0.43
1:A:189:ASN:O	1:A:273:TYR:HA	2.18	0.42
1:B:85:SER:HB3	1:B:150:GLU:HA	2.02	0.42
1:B:157:ILE:O	1:B:174:ILE:N	2.51	0.42
1:B:372:VAL:HG11	1:B:374:LYS:HE3	2.01	0.42
1:A:188:TRP:CZ2	1:A:275:LYS:HD3	2.55	0.42
1:B:86:GLY:O	1:B:148:VAL:N	2.52	0.42
1:B:375:ASP:C	1:B:377:GLN:H	2.23	0.42
1:A:90:PHE:O	1:A:143:SER:HA	2.20	0.41
1:A:343:HIS:NE2	1:A:345:ALA:HB2	2.35	0.41
1:A:432:PHE:HA	1:A:433:PRO:HA	1.74	0.41
1:B:333:TYR:HA	1:B:338:GLU:O	2.19	0.41
1:B:373:VAL:O	1:B:381:ILE:HB	2.19	0.41
1:A:74:THR:OG1	1:A:77:ARG:HB2	2.21	0.41
1:A:33:CYS:O	1:A:36:ILE:HG22	2.20	0.41
1:A:430:LYS:HG3	1:A:431:GLU:N	2.36	0.41
1:B:23:ALA:HB1	1:B:81:ARG:NH1	2.36	0.41
1:B:382:GLU:HG2	1:B:384:ASN:HD21	1.84	0.41
1:A:185:LEU:HB3	1:A:188:TRP:HB2	2.02	0.41
1:B:335:HIS:N	1:B:366:HIS:O	2.54	0.41

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	462/464 (100%)	434 (94%)	28 (6%)	0	100 100
1	B	356/464 (77%)	319 (90%)	34 (10%)	3 (1%)	19 20
All	All	818/928 (88%)	753 (92%)	62 (8%)	3 (0%)	34 38

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	361	LYS
1	B	258	ASN
1	B	376	GLN

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	392/392 (100%)	388 (99%)	4 (1%)	76 85
1	B	304/392 (78%)	298 (98%)	6 (2%)	55 66
All	All	696/784 (89%)	686 (99%)	10 (1%)	67 78

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

Mol	Chain	Res	Type
1	A	8	GLU
1	A	25	ASP
1	A	177	LYS
1	A	338	GLU
1	B	85	SER
1	B	219	SER
1	B	229	GLU
1	B	252	LEU
1	B	278	SER
1	B	325	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	428	GLN
1	B	1	HIS

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	464/464 (100%)	0.54	50 (10%)	5 9	56, 84, 133, 175	0
1	B	360/464 (77%)	1.07	75 (20%)	1 1	48, 81, 164, 190	0
All	All	824/928 (88%)	0.77	125 (15%)	2 3	48, 83, 153, 190	0

All (125) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	347	LEU	10.7
1	B	326	ILE	9.4
1	B	357	MET	7.3
1	B	359	LEU	7.1
1	B	352	VAL	7.1
1	B	336	HIS	6.6
1	A	437	TYR	6.5
1	B	337	HIS	6.5
1	B	356	THR	5.8
1	B	343	HIS	5.7
1	B	350	GLY	5.6
1	A	432	PHE	5.4
1	B	323	THR	5.4
1	B	388	PHE	5.4
1	B	355	ALA	5.3
1	B	184	VAL	5.3
1	B	362	SER	5.2
1	B	325	ASP	5.2
1	B	389	HIS	5.2
1	B	185	LEU	5.1
1	A	183	PRO	5.0
1	B	181	ASP	5.0
1	B	183	PRO	5.0
1	B	349	ASP	5.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	159	ALA	4.7
1	B	335	HIS	4.7
1	B	364	PRO	4.7
1	B	365	GLY	4.7
1	B	257	LEU	4.7
1	A	181	ASP	4.7
1	A	184	VAL	4.6
1	B	354	ALA	4.5
1	A	180	GLY	4.4
1	B	158	LEU	4.4
1	B	367	HIS	4.4
1	B	322	ALA	4.3
1	B	338	GLU	4.1
1	B	340	LEU	4.1
1	B	376	GLN	4.1
1	B	361	LYS	4.0
1	A	434	PHE	4.0
1	B	259	ALA	3.9
1	A	433	PRO	3.9
1	A	335	HIS	3.8
1	A	448	PHE	3.8
1	B	173	VAL	3.8
1	A	393	GLU	3.8
1	B	157	ILE	3.7
1	A	395	THR	3.6
1	A	396	PRO	3.6
1	A	174	ILE	3.6
1	B	363	GLU	3.6
1	A	391	ILE	3.5
1	B	172	ASN	3.5
1	A	394	GLN	3.5
1	B	368	MET	3.5
1	B	324	GLY	3.5
1	A	376	GLN	3.5
1	B	386	LEU	3.4
1	B	260	GLU	3.4
1	B	353	GLU	3.4
1	A	185	LEU	3.3
1	B	366	HIS	3.3
1	A	444	SER	3.3
1	B	346	GLU	3.3
1	B	351	GLN	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	461	ASN	3.2
1	A	263	ILE	3.2
1	B	390	LEU	3.2
1	B	387	ASP	3.2
1	A	173	VAL	3.1
1	B	328	THR	3.1
1	B	377	GLN	3.1
1	B	345	ALA	3.0
1	B	160	VAL	3.0
1	B	358	THR	3.0
1	A	179	ASP	2.9
1	A	458	GLU	2.9
1	A	322	ALA	2.9
1	A	172	ASN	2.9
1	B	348	SER	2.9
1	B	174	ILE	2.8
1	A	348	SER	2.8
1	A	261	GLY	2.8
1	B	227	ASP	2.8
1	B	130	VAL	2.7
1	A	4	VAL	2.7
1	A	182	SER	2.6
1	A	159	ALA	2.6
1	A	440	GLN	2.6
1	A	158	LEU	2.6
1	A	156	VAL	2.5
1	A	283	VAL	2.5
1	A	295	GLU	2.5
1	B	202	VAL	2.5
1	B	330	MET	2.5
1	B	379	LYS	2.5
1	B	226	ILE	2.4
1	B	378	GLY	2.4
1	A	157	ILE	2.4
1	A	313	ALA	2.4
1	A	259	ALA	2.3
1	A	392	GLU	2.3
1	A	347	LEU	2.3
1	B	382	GLU	2.3
1	B	277	GLY	2.2
1	A	84	GLN	2.2
1	A	218	VAL	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	276	GLN	2.2
1	A	260	GLU	2.2
1	B	380	VAL	2.2
1	A	336	HIS	2.1
1	A	431	GLU	2.1
1	B	136	VAL	2.1
1	A	74	THR	2.1
1	B	193	GLN	2.1
1	B	375	ASP	2.1
1	B	383	GLN	2.1
1	B	43	VAL	2.1
1	B	373	VAL	2.1
1	B	182	SER	2.1
1	A	372	VAL	2.1
1	A	105	TYR	2.1
1	A	296	TYR	2.1
1	B	327	LYS	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.