



Full wwPDB X-ray Structure Validation Report ⓘ

May 13, 2020 – 03:08 am BST

PDB ID : 6L8W
Title : Crystal structure of ugt transferase mutant2
Authors : Li, J.; Shan, N.; Yang, J.G.; Liu, W.D.; Sun, Y.X.
Deposited on : 2019-11-07
Resolution : 2.05 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 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	:	2.11
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)	:	2.11

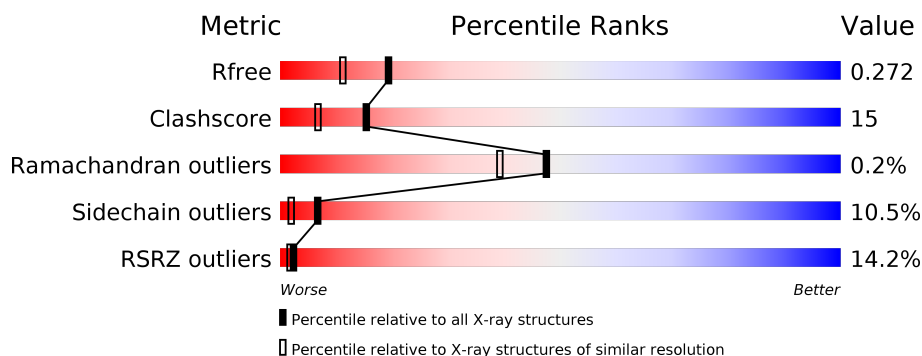
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.05 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	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)
RSRZ outliers	127900	1672 (2.04-2.04)

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 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 $\leq 5\%$. 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	454	<div> <div>14%</div> <div> <div></div> <div>76%</div> <div>16%</div> <div>5%</div> <div></div> </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3546 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 Glycosyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	438	Total	C	N	O	S	0	0	0
			3472	2225	573	656	18			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	15	ALA	SER	engineered mutation	UNP K7NBW3
A	28	HIS	ARG	engineered mutation	UNP K7NBW3
A	47	ARG	HIS	engineered mutation	UNP K7NBW3
A	48	MET	LEU	engineered mutation	UNP K7NBW3
A	76	LEU	MET	engineered mutation	UNP K7NBW3
A	79	TYR	THR	engineered mutation	UNP K7NBW3
A	109	ILE	LEU	engineered mutation	UNP K7NBW3

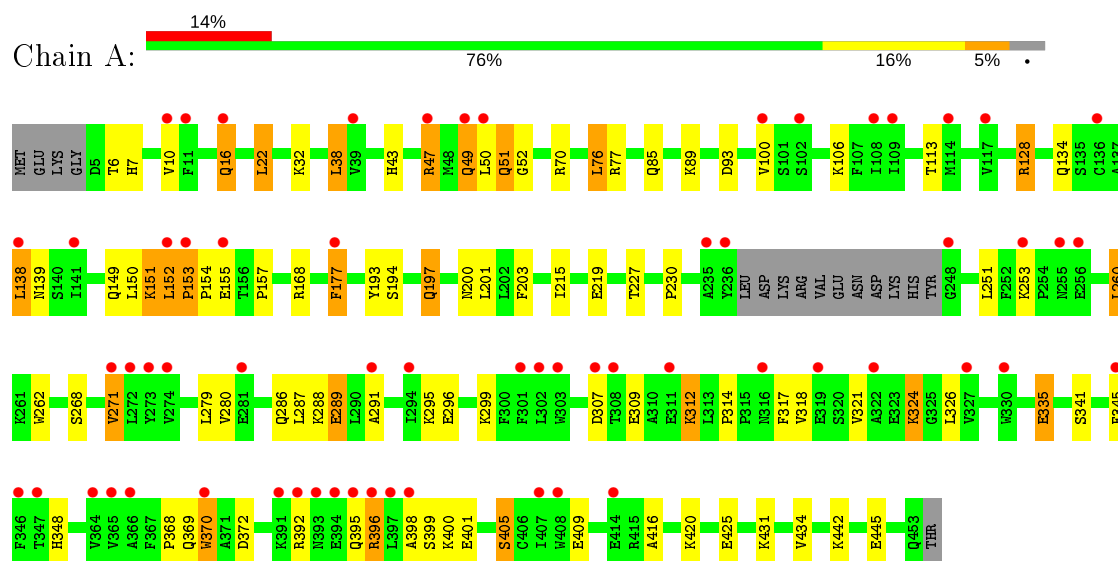
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	74	Total	O	0	0
			74	74		

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: Glycosyltransferase



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	48.77Å 69.56Å 152.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	24.76 – 2.05 24.75 – 2.05	Depositor EDS
% Data completeness (in resolution range)	99.0 (24.76-2.05) 99.1 (24.75-2.05)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.49 (at 2.04Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.207 , 0.269 0.222 , 0.272	Depositor DCC
R_{free} test set	1666 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	45.2	Xtriage
Anisotropy	0.052	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 51.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	3546	wwPDB-VP
Average B, all atoms (Å ²)	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.97% 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, 0.333 respectively for untwinned datasets, and 0.375, 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	0.73	0/3555	0.76	0/4825

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 [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	3472	0	3446	103	0
2	A	74	0	0	3	0
All	All	3546	0	3446	103	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 15.

All (103) 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:16:GLN:HG3	1:A:279:LEU:CD1	1.72	1.19
1:A:16:GLN:HG3	1:A:279:LEU:HD11	1.16	1.09
1:A:6:THR:HG23	1:A:106:LYS:HD2	1.37	1.06
1:A:177:PHE:CD2	1:A:370:TRP:HB3	1.95	1.00

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:16:GLN:CG	1:A:279:LEU:HD11	1.94	0.98
1:A:138:LEU:HD21	1:A:372:ASP:HB3	1.47	0.97
1:A:286:GLN:HE22	1:A:398:ALA:H	0.94	0.93
1:A:260:LEU:HD11	1:A:335:GLU:HG2	1.53	0.88
1:A:197:GLN:H	1:A:197:GLN:HE21	1.21	0.88
1:A:286:GLN:HE22	1:A:398:ALA:N	1.76	0.84
1:A:396:ARG:HH11	1:A:396:ARG:HG2	1.43	0.83
1:A:152:LEU:HB2	1:A:153:PRO:HD3	1.65	0.78
1:A:286:GLN:NE2	1:A:398:ALA:H	1.79	0.78
1:A:47:ARG:NH1	1:A:47:ARG:HB3	2.01	0.75
1:A:177:PHE:CD2	1:A:370:TRP:CB	2.70	0.75
1:A:314:PRO:HG2	1:A:317:PHE:HB2	1.69	0.74
1:A:177:PHE:HD2	1:A:370:TRP:HB3	1.53	0.73
1:A:215:ILE:HD13	1:A:227:THR:HG21	1.70	0.73
1:A:348:HIS:HD2	2:A:523:HOH:O	1.73	0.71
1:A:151:LYS:HD3	1:A:155:GLU:H	1.55	0.71
1:A:177:PHE:HB2	1:A:370:TRP:HB2	1.73	0.71
1:A:51:GLN:H	1:A:51:GLN:NE2	1.90	0.69
1:A:152:LEU:CB	1:A:153:PRO:HD3	2.26	0.65
1:A:151:LYS:H	1:A:151:LYS:HD2	1.61	0.65
1:A:138:LEU:CD2	1:A:372:ASP:HB3	2.22	0.65
1:A:392:ARG:HG3	1:A:398:ALA:HB2	1.79	0.64
1:A:49:GLN:H	1:A:49:GLN:CD	2.01	0.64
1:A:70:ARG:HH21	1:A:70:ARG:HG2	1.61	0.64
1:A:16:GLN:CB	1:A:279:LEU:HD11	2.27	0.64
1:A:296:GLU:OE2	1:A:400:LYS:HD2	1.97	0.64
1:A:47:ARG:HH11	1:A:47:ARG:CG	2.11	0.63
1:A:289:GLU:HG3	1:A:400:LYS:HB3	1.79	0.63
1:A:262:TRP:CG	1:A:326:LEU:HD22	2.34	0.62
1:A:43:HIS:CE1	1:A:47:ARG:HG3	2.35	0.62
1:A:47:ARG:CZ	1:A:47:ARG:HB3	2.30	0.61
1:A:177:PHE:CB	1:A:370:TRP:HB2	2.30	0.61
1:A:47:ARG:HH11	1:A:47:ARG:HG3	1.65	0.60
1:A:396:ARG:HG2	1:A:396:ARG:NH1	2.15	0.60
1:A:16:GLN:HG2	1:A:16:GLN:O	2.02	0.58
1:A:51:GLN:HE21	1:A:52:GLY:H	1.51	0.58
1:A:128:ARG:H	1:A:200:ASN:ND2	2.01	0.58
1:A:100:VAL:O	1:A:100:VAL:HG23	2.04	0.57
1:A:280:VAL:HG11	1:A:369:GLN:HG3	1.87	0.57
1:A:260:LEU:HD23	1:A:260:LEU:H	1.69	0.56
1:A:152:LEU:HB2	1:A:153:PRO:CD	2.34	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:134:GLN:OE1	1:A:138:LEU:CD2	2.54	0.55
1:A:49:GLN:OE1	1:A:49:GLN:N	2.39	0.55
1:A:134:GLN:OE1	1:A:138:LEU:HD23	2.06	0.55
1:A:70:ARG:HH21	1:A:70:ARG:CG	2.19	0.54
1:A:150:LEU:HA	1:A:151:LYS:HE3	1.90	0.54
1:A:396:ARG:HH11	1:A:396:ARG:CG	2.16	0.54
1:A:151:LYS:N	1:A:151:LYS:HD2	2.23	0.53
1:A:51:GLN:H	1:A:51:GLN:HE21	1.56	0.53
1:A:177:PHE:HD2	1:A:370:TRP:CB	2.17	0.52
1:A:260:LEU:CD2	1:A:260:LEU:N	2.72	0.52
1:A:151:LYS:N	1:A:151:LYS:CD	2.72	0.52
1:A:289:GLU:CG	1:A:400:LYS:N	2.73	0.52
1:A:152:LEU:CB	1:A:153:PRO:CD	2.88	0.51
1:A:289:GLU:HG2	1:A:399:SER:HA	1.90	0.51
1:A:139:ASN:HA	1:A:193:TYR:OH	2.11	0.51
1:A:289:GLU:HG3	1:A:400:LYS:CB	2.40	0.51
1:A:260:LEU:HD23	1:A:260:LEU:N	2.27	0.50
1:A:309:GLU:HA	1:A:312:LYS:HE2	1.94	0.49
1:A:47:ARG:NH1	1:A:47:ARG:CB	2.73	0.49
1:A:215:ILE:CD1	1:A:227:THR:HG21	2.42	0.49
1:A:32:LYS:HD3	1:A:445:GLU:CD	2.32	0.49
1:A:157:PRO:HD2	2:A:557:HOH:O	2.11	0.49
1:A:295:LYS:HG3	1:A:321:VAL:HG12	1.95	0.48
1:A:268:SER:HA	1:A:341:SER:HA	1.95	0.48
1:A:47:ARG:NH1	1:A:47:ARG:CG	2.73	0.48
1:A:289:GLU:HG3	1:A:400:LYS:N	2.29	0.47
1:A:22:LEU:HA	1:A:22:LEU:HD13	1.73	0.47
1:A:291:ALA:HB1	1:A:317:PHE:CE1	2.49	0.47
1:A:201:LEU:HD11	1:A:203:PHE:CE2	2.50	0.46
1:A:89:LYS:HE3	1:A:93:ASP:OD1	2.15	0.46
1:A:324:LYS:HD2	1:A:324:LYS:HA	1.49	0.46
1:A:396:ARG:NH1	1:A:396:ARG:CG	2.72	0.46
1:A:43:HIS:CE1	1:A:47:ARG:CG	2.97	0.46
1:A:197:GLN:H	1:A:197:GLN:NE2	2.01	0.46
1:A:289:GLU:HG2	1:A:400:LYS:N	2.30	0.46
1:A:401:GLU:OE1	1:A:401:GLU:HA	2.15	0.45
1:A:416:ALA:O	1:A:420:LYS:HG3	2.16	0.45
1:A:43:HIS:CE1	1:A:47:ARG:HB2	2.51	0.45
1:A:177:PHE:HB3	1:A:370:TRP:CE3	2.52	0.45
1:A:401:GLU:O	1:A:405:SER:HB3	2.18	0.44
1:A:151:LYS:HD2	1:A:154:PRO:HA	2.00	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:10:VAL:HG23	1:A:38:LEU:HD12	1.99	0.44
1:A:201:LEU:HD11	1:A:203:PHE:CZ	2.53	0.44
1:A:47:ARG:CZ	1:A:47:ARG:CB	2.94	0.43
1:A:395:GLN:HG2	1:A:395:GLN:H	1.47	0.43
1:A:230:PRO:HD2	1:A:434:VAL:HG13	2.01	0.43
1:A:32:LYS:HD3	1:A:445:GLU:OE2	2.19	0.42
1:A:262:TRP:CE2	1:A:326:LEU:HB2	2.55	0.42
1:A:215:ILE:O	1:A:219:GLU:HG3	2.19	0.42
1:A:271:VAL:HG13	1:A:341:SER:O	2.19	0.42
1:A:76:LEU:HD23	1:A:76:LEU:HA	1.86	0.41
1:A:368:PRO:O	1:A:392:ARG:NH1	2.40	0.41
1:A:409:GLU:OE1	1:A:409:GLU:HA	2.19	0.41
1:A:348:HIS:CD2	2:A:523:HOH:O	2.58	0.41
1:A:392:ARG:CG	1:A:398:ALA:HB2	2.49	0.41
1:A:49:GLN:N	1:A:49:GLN:CD	2.72	0.40
1:A:348:HIS:H	1:A:348:HIS:CD2	2.40	0.40
1:A:70:ARG:NH2	1:A:70:ARG:CG	2.82	0.40

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	434/454 (96%)	418 (96%)	15 (4%)	1 (0%)	47 39

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	153	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	390/406 (96%)	349 (90%)	41 (10%)	7 2

All (41) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	7	HIS
1	A	16	GLN
1	A	22	LEU
1	A	38	LEU
1	A	47	ARG
1	A	49	GLN
1	A	50	LEU
1	A	51	GLN
1	A	76	LEU
1	A	77	ARG
1	A	85	GLN
1	A	113	THR
1	A	128	ARG
1	A	138	LEU
1	A	149	GLN
1	A	151	LYS
1	A	152	LEU
1	A	168	ARG
1	A	177	PHE
1	A	194	SER
1	A	197	GLN
1	A	251	LEU
1	A	253	LYS
1	A	260	LEU
1	A	271	VAL
1	A	287	LEU
1	A	288	LYS
1	A	289	GLU
1	A	299	LYS
1	A	307	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	312	LYS
1	A	318	VAL
1	A	324	LYS
1	A	335	GLU
1	A	345	PHE
1	A	370	TRP
1	A	396	ARG
1	A	405	SER
1	A	425	GLU
1	A	431	LYS
1	A	442	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	18	HIS
1	A	43	HIS
1	A	46	ASN
1	A	51	GLN
1	A	90	ASN
1	A	96	GLN
1	A	197	GLN
1	A	200	ASN
1	A	255	ASN
1	A	286	GLN
1	A	316	ASN
1	A	348	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	438/454 (96%)	0.85	62 (14%) 2 2	30, 55, 99, 134	0

All (62) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	394	GLU	9.1
1	A	395	GLN	8.4
1	A	49	GLN	6.9
1	A	248	GLY	6.6
1	A	396	ARG	6.1
1	A	393	ASN	5.6
1	A	274	VAL	5.5
1	A	346	PHE	5.0
1	A	272	LEU	5.0
1	A	155	GLU	5.0
1	A	414	GLU	4.9
1	A	47	ARG	4.8
1	A	102	SER	4.6
1	A	177	PHE	4.4
1	A	311	GLU	4.4
1	A	397	LEU	4.3
1	A	152	LEU	4.2
1	A	50	LEU	4.0
1	A	256	GLU	3.9
1	A	11	PHE	3.9
1	A	303	TRP	3.8
1	A	392	ARG	3.7
1	A	365	VAL	3.7
1	A	398	ALA	3.7
1	A	16	GLN	3.5
1	A	301	PHE	3.3
1	A	302	LEU	3.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	391	LYS	3.2
1	A	408	TRP	3.2
1	A	319	GLU	3.0
1	A	236	TYR	3.0
1	A	308	THR	3.0
1	A	307	ASP	2.9
1	A	255	ASN	2.8
1	A	100	VAL	2.7
1	A	253	LYS	2.7
1	A	138	LEU	2.7
1	A	294	ILE	2.6
1	A	322	ALA	2.6
1	A	345	PHE	2.6
1	A	39	VAL	2.5
1	A	347	THR	2.5
1	A	407	ILE	2.5
1	A	364	VAL	2.4
1	A	273	TYR	2.4
1	A	281	GLU	2.4
1	A	330	TRP	2.4
1	A	10	VAL	2.4
1	A	366	ALA	2.3
1	A	141	ILE	2.3
1	A	136	CYS	2.3
1	A	109	ILE	2.3
1	A	153	PRO	2.2
1	A	291	ALA	2.2
1	A	316	ASN	2.2
1	A	117	VAL	2.1
1	A	114	MET	2.1
1	A	271	VAL	2.1
1	A	235	ALA	2.1
1	A	108	ILE	2.1
1	A	370	TRP	2.1
1	A	327	VAL	2.1

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 [i](#)

There are no carbohydrates 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.