



Full wwPDB X-ray Structure Validation Report ⓘ

Jun 14, 2020 – 02:30 pm BST

PDB ID : 2JIW
Title : Bacteroides thetaiotaomicron GH84 O-GlcNAcase in complex with 2- Acetylamino-2-deoxy-1-epivalienamine
Authors : Dennis, R.J.; Davies, G.J.
Deposited on : 2007-07-02
Resolution : 1.95 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
Mogul : 1.8.5 (274361), CSD as541be (2020)
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

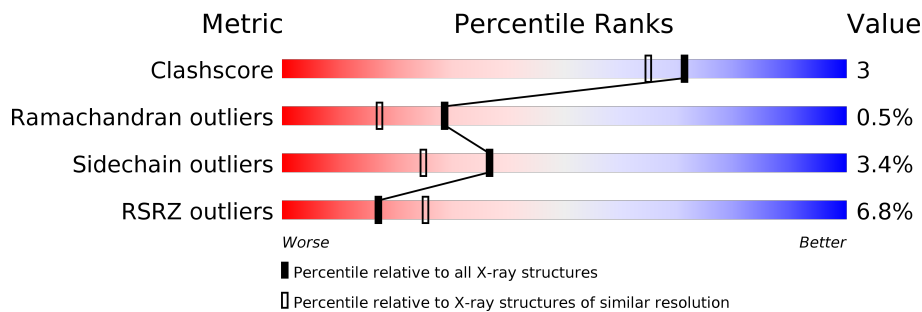
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.95 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))
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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	716	 6% 75% 7% 18%
1	B	716	 5% 72% 9% 18%

2 Entry composition [i](#)

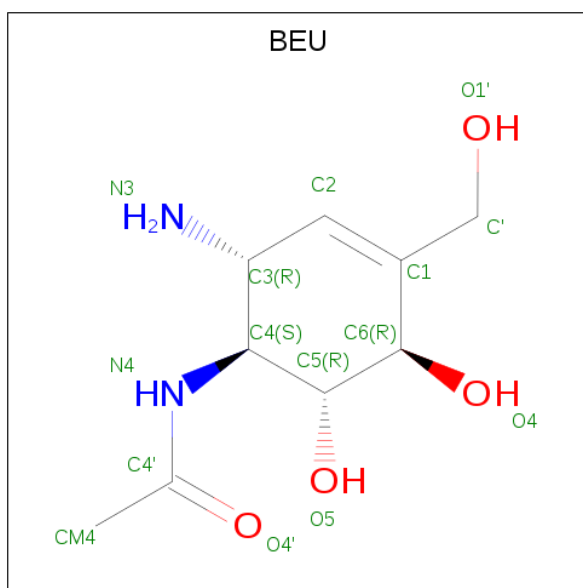
There are 3 unique types of molecules in this entry. The entry contains 10470 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 O-GLCNACASE BT_4395.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	590	Total 4821	C 3087	N 812	O 904	S 18	0	4	0
1	B	590	Total 4853	C 3109	N 820	O 905	S 19	0	9	0

- Molecule 2 is N-[(1S,2R,5R,6R)-2-AMINO-5,6-DIHYDROXY-4-(HYDROXYMETHYL)CYCLOHEX-3-EN-1-YL]ACETAMIDE (three-letter code: BEU) (formula: C₉H₁₆N₂O₄).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	Total 15	C 9	N 2	O 4	0	0
2	B	1	Total 15	C 9	N 2	O 4	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	363	Total 363	O 363	0	0
3	B	403	Total 403	O 403	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	51.40Å 94.13Å 99.81Å 104.97° 94.09° 102.48°	Depositor
Resolution (Å)	95.35 – 1.95 25.46 – 1.95	Depositor EDS
% Data completeness (in resolution range)	97.0 (95.35-1.95) 97.0 (25.46-1.95)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.82 (at 1.95Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.206 , 0.248 0.207 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	22.1	Xtrriage
Anisotropy	0.028	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 49.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	10470	wwPDB-VP
Average B, all atoms (Å ²)	23.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.58% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: BEU

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	0.64	0/4961	0.68	0/6726
1	B	0.65	0/5008	0.67	1/6786 (0.0%)
All	All	0.64	0/9969	0.67	1/13512 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	2
All	All	0	3

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	B	451	LEU	CA-CB-CG	6.91	131.20	115.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	22	PRO	Peptide
1	B	22	PRO	Peptide
1	B	4	SER	Peptide

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	4821	0	4721	28	0
1	B	4853	0	4775	39	0
2	A	15	0	16	0	0
2	B	15	0	16	0	0
3	A	363	0	0	3	0
3	B	403	0	0	1	0
All	All	10470	0	9528	67	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 3.

All (67) 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:458:LYS:HB3	1:B:459:ASN:HA	1.25	1.12
1:B:22:PRO:HA	1:B:23:ALA:HB3	1.35	1.08
1:B:22:PRO:HA	1:B:23:ALA:CB	1.85	1.06
1:B:219[A]:ARG:HG3	1:B:219[A]:ARG:HH11	1.32	0.92
1:A:22:PRO:HA	1:A:23:ALA:CB	2.03	0.89
1:A:22:PRO:HA	1:A:23:ALA:HB3	1.55	0.89
1:B:458:LYS:HB3	1:B:459:ASN:CA	2.12	0.77
1:B:22:PRO:CA	1:B:23:ALA:HB3	2.17	0.73
1:B:219[A]:ARG:HG3	1:B:219[A]:ARG:NH1	2.03	0.72
1:B:453:ALA:O	1:B:458:LYS:N	2.26	0.63
1:A:532:LYS:O	1:A:536:GLN:HG3	1.99	0.60
1:B:355:VAL:O	1:B:399:THR:HG23	2.01	0.60
1:A:39:LYS:HG2	3:A:2026:HOH:O	2.01	0.60
1:B:536:GLN:HG2	1:B:590:TYR:CD1	2.38	0.58
1:B:458:LYS:CB	1:B:459:ASN:HA	2.14	0.57
1:A:22:PRO:CA	1:A:23:ALA:HB3	2.33	0.56
1:A:70:TYR:OH	1:A:89:GLU:OE1	2.22	0.56
1:B:22:PRO:HA	1:B:23:ALA:HB2	1.82	0.56
1:B:462:LYS:HD3	1:B:466:GLU:OE2	2.06	0.56
1:A:262:ASP:HA	1:A:266:ALA:HB3	1.88	0.55
1:B:21:LEU:HD23	1:B:48:LYS:HE3	1.89	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:238:ALA:HA	1:A:276:VAL:O	2.07	0.55
1:A:585:ASP:OD1	1:A:587:THR:HG23	2.08	0.54
1:B:219[A]:ARG:HH12	1:B:257:LEU:HA	1.73	0.53
1:B:28:ASN:HD22	1:B:56:LEU:HD11	1.75	0.51
1:A:22:PRO:CA	1:A:23:ALA:CB	2.85	0.51
1:A:81:TYR:CZ	1:A:123:ASP:HB3	2.46	0.50
1:A:22:PRO:HA	1:A:23:ALA:HB2	1.93	0.49
1:A:455:LYS:HB3	3:A:2302:HOH:O	2.13	0.49
1:B:549:PRO:HG2	1:B:550:TYR:CE2	2.48	0.49
1:A:70:TYR:O	1:A:73:GLN:HG2	2.14	0.48
1:B:451:LEU:HD13	1:B:564:LEU:HD12	1.96	0.48
1:A:451:LEU:HD13	1:A:564:LEU:HD12	1.97	0.47
1:B:534:LEU:O	1:B:538:MET:HG3	2.16	0.46
1:B:356:TYR:HB3	1:B:399:THR:HG21	1.97	0.46
1:B:308:MET:HA	1:B:335:TYR:O	2.16	0.45
1:B:83:LEU:HD22	1:B:121:ILE:HD12	1.98	0.45
1:B:173:ALA:HA	1:B:174:PRO:HA	1.67	0.45
1:A:228[A]:LYS:HD2	1:A:228[A]:LYS:HA	1.83	0.45
1:B:585:ASP:OD1	1:B:587:THR:HG23	2.17	0.45
1:A:293:TYR:CD2	1:A:293:TYR:C	2.91	0.44
1:B:238:ALA:HA	1:B:276:VAL:O	2.18	0.44
1:A:341:PRO:O	1:A:342:VAL:C	2.56	0.44
1:B:310:THR:HA	1:B:337:TRP:O	2.17	0.44
1:B:540[A]:TYR:CE1	1:B:544:THR:HG21	2.53	0.43
1:A:261:ILE:O	1:A:265:PHE:HB3	2.18	0.43
1:B:236:SER:OG	1:B:273:ASN:HB2	2.19	0.43
1:A:536:GLN:HG2	1:A:590:TYR:CD1	2.53	0.43
1:A:308:MET:HA	1:A:335:TYR:O	2.18	0.43
1:B:81:TYR:CZ	1:B:123:ASP:HB3	2.52	0.43
1:B:475:MET:HG2	1:B:502:PHE:CZ	2.53	0.43
1:B:516:VAL:HG22	1:B:572:VAL:HG11	2.02	0.42
1:A:378:GLU:HG3	1:A:490:PRO:HB2	2.01	0.42
1:B:320:ARG:O	1:B:324:SER:OG	2.32	0.42
1:B:173:ALA:HB3	3:B:2049:HOH:O	2.19	0.42
1:A:303:PRO:HD2	3:A:2208:HOH:O	2.20	0.42
1:A:557:ALA:HB1	1:A:561:ILE:HB	2.01	0.42
1:B:314:VAL:HG12	1:B:315:ILE:HD13	2.02	0.41
1:B:302:ASN:HA	1:B:303:PRO:HD3	1.93	0.41
1:B:219[A]:ARG:NH1	1:B:257:LEU:HD12	2.35	0.41
1:A:73:GLN:HG3	1:A:82:TYR:CE2	2.56	0.40
1:B:531:VAL:HG11	1:B:569:PHE:CE1	2.56	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:534:LEU:O	1:A:538:MET:HG3	2.21	0.40
1:A:562:LYS:HB3	1:A:563:PRO:HD3	2.04	0.40
1:B:454:PHE:HZ	1:B:516:VAL:HG13	1.86	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	592/716 (83%)	570 (96%)	18 (3%)	4 (1%)	22 11
1	B	597/716 (83%)	577 (97%)	18 (3%)	2 (0%)	41 30
All	All	1189/1432 (83%)	1147 (96%)	36 (3%)	6 (0%)	29 17

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	458	LYS
1	A	23	ALA
1	A	518	GLY
1	B	23	ALA
1	A	49	GLN
1	B	49	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	522/630 (83%)	506 (97%)	16 (3%)	40	28
1	B	527/630 (84%)	506 (96%)	21 (4%)	31	19
All	All	1049/1260 (83%)	1012 (96%)	37 (4%)	37	24

All (37) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	4	SER
1	A	27	LEU
1	A	32	GLU
1	A	43	GLU
1	A	45	LEU
1	A	46	SER
1	A	53	LYS
1	A	73	GLN
1	A	88	LYS
1	A	219	ARG
1	A	221	LEU
1	A	371	THR
1	A	408	ARG
1	A	455	LYS
1	A	456	GLU
1	A	459	ASN
1	B	6	GLN
1	B	24	VAL
1	B	26	GLN
1	B	27	LEU
1	B	45	LEU
1	B	49	GLN
1	B	52	LYS
1	B	87	GLU
1	B	120	GLU
1	B	219[A]	ARG
1	B	219[B]	ARG
1	B	221	LEU
1	B	324	SER
1	B	371	THR
1	B	408	ARG
1	B	416	GLU
1	B	441[A]	MET
1	B	441[B]	MET
1	B	455	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	462	LYS
1	B	519	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	11	GLN
1	A	189	GLN
1	A	273	ASN
1	A	459	ASN
1	B	6	GLN
1	B	11	GLN
1	B	28	ASN
1	B	254	GLN
1	B	274	GLN
1	B	306	GLN
1	B	543	GLN

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](#)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	BEU	A	1594	-	15,15,15	0.76	0	13,21,21	0.91	1 (7%)
2	BEU	B	1594	-	15,15,15	1.00	1 (6%)	13,21,21	1.31	1 (7%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	BEU	A	1594	-	-	2/6/26/26	0/1/1/1
2	BEU	B	1594	-	-	2/6/26/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1594	BEU	C3-C4	3.38	1.57	1.54

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1594	BEU	C4-C3-N3	2.54	115.87	111.52
2	A	1594	BEU	C6-C1-C2	-2.25	118.53	122.23

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1594	BEU	C3-C4-N4-C4'
2	B	1594	BEU	C3-C4-N4-C4'
2	A	1594	BEU	C5-C4-N4-C4'
2	B	1594	BEU	C5-C4-N4-C4'

There are no ring outliers.

No monomer is involved in short contacts.

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	590/716 (82%)	0.20	42 (7%) 16 24	11, 21, 47, 64	0
1	B	590/716 (82%)	0.19	38 (6%) 19 28	11, 20, 45, 62	0
All	All	1180/1432 (82%)	0.20	80 (6%) 17 25	11, 20, 47, 64	0

All (80) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	47	GLY	6.7
1	B	16	ASN	6.4
1	B	455	LYS	6.3
1	A	47	GLY	6.3
1	B	52	LYS	5.6
1	A	113	ASP	5.3
1	A	16	ASN	5.3
1	A	54	GLY	5.0
1	A	92	LEU	4.8
1	A	457	GLY	4.6
1	B	53	LYS	4.5
1	A	53	LYS	4.5
1	A	51	SER	4.4
1	B	54	GLY	4.3
1	B	15	GLN	4.2
1	B	92	LEU	4.1
1	B	18	THR	4.0
1	A	15	GLN	3.9
1	A	455	LYS	3.9
1	B	581	ASN	3.9
1	A	544	THR	3.8
1	A	456	GLU	3.8
1	B	112	LYS	3.7
1	A	583	HIS	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	593	HIS	3.5
1	A	112	LYS	3.5
1	A	46	SER	3.5
1	A	18	THR	3.5
1	A	519	ARG	3.4
1	B	583	HIS	3.4
1	A	49	GLN	3.3
1	B	456	GLU	3.3
1	B	49	GLN	3.2
1	B	457	GLY	3.1
1	A	582	ALA	3.0
1	B	458	LYS	3.0
1	B	107	PHE	2.9
1	B	24	VAL	2.9
1	B	51	SER	2.9
1	A	26	GLN	2.9
1	B	46	SER	2.9
1	B	45	LEU	2.8
1	A	52	LYS	2.8
1	B	452	LYS	2.7
1	A	461	ASP	2.7
1	A	87	GLU	2.6
1	B	336	ILE	2.6
1	A	57	ILE	2.5
1	A	17	LYS	2.5
1	B	462	LYS	2.5
1	A	115	LYS	2.5
1	A	578	GLN	2.4
1	B	593	HIS	2.4
1	A	451	LEU	2.4
1	B	113	ASP	2.4
1	A	4	SER	2.4
1	B	87	GLU	2.4
1	B	41	LEU	2.4
1	A	581	ASN	2.4
1	A	276	VAL	2.4
1	B	582	ALA	2.3
1	B	540[A]	TYR	2.3
1	A	14	VAL	2.3
1	A	462	LYS	2.3
1	B	592	PRO	2.3
1	A	41	LEU	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	458	LYS	2.2
1	A	107	PHE	2.2
1	A	580	PHE	2.2
1	B	132	VAL	2.2
1	A	83	LEU	2.1
1	A	459	ASN	2.1
1	A	118	GLU	2.1
1	B	106	THR	2.1
1	B	17	LYS	2.1
1	B	519	ARG	2.1
1	B	104	LEU	2.1
1	B	103	ALA	2.0
1	A	591	MET	2.0
1	B	115	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 carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q<0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	BEU	A	1594	15/15	0.96	0.09	13,16,18,19	0
2	BEU	B	1594	15/15	0.96	0.08	11,14,15,16	0

6.5 Other polymers [i](#)

There are no such residues in this entry.