



Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 07:39 PM GMT

PDB ID : 1GPI
Title : CELLOBIOHYDROLASE CEL7D (CBH 58) FROM PHANEROCHAETE
CHRYSPORIUM. CATALYTIC MODULE AT 1.32 ANG RESOLUTION
Authors : Munoz, I.G.; Mowbray, S.L.; Stahlberg, J.
Deposited on : 2001-11-05
Resolution : 1.32 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
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

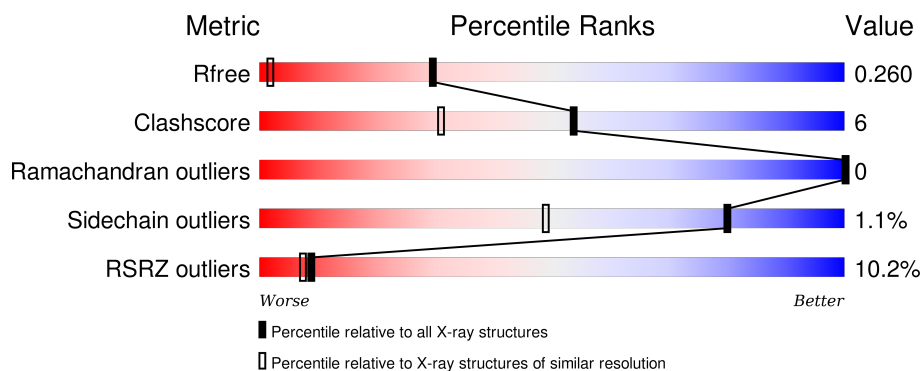
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.32 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}	91344	1004 (1.34-1.30)
Clashscore	102246	1059 (1.34-1.30)
Ramachandran outliers	100387	1015 (1.34-1.30)
Sidechain outliers	100360	1015 (1.34-1.30)
RSRZ outliers	91569	1002 (1.34-1.30)

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. 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	431	<div> <div>10%</div> <div>86%</div> <div>13%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NAG	A	435	-	-	-	X

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 3659 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.

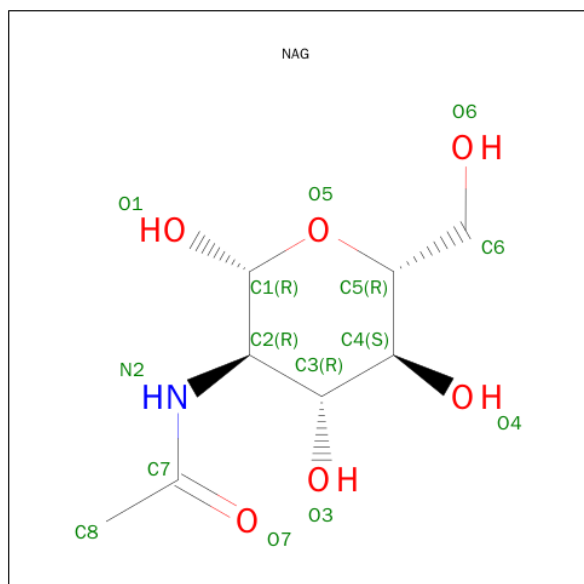
- Molecule 1 is a protein called EXOGLUCANASE I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	431	Total	C	N	O	S	0	0	0
			3198	1962	541	667	28			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	57	ASP	ASN	CONFLICT	UNP Q09431
A	277	VAL	ASP	CONFLICT	UNP Q09431
A	.	-	HIS	DELETION	UNP Q09431

- Molecule 2 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		

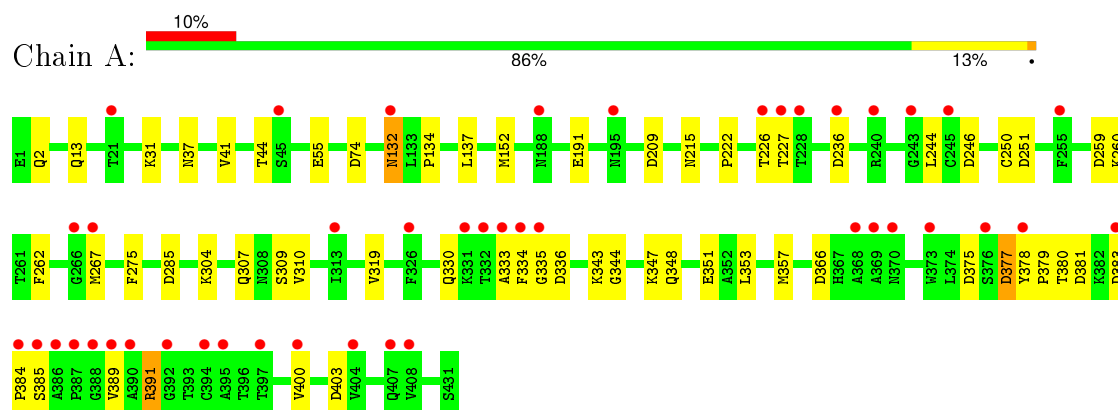
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	447	Total	O	0	0
			447	447		

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: EXOGLUCANASE I



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	86.31Å 46.62Å 99.08Å 90.00° 103.19° 90.00°	Depositor
Resolution (Å)	13.00 – 1.32 12.83 – 1.32	Depositor EDS
% Data completeness (in resolution range)	99.0 (13.00-1.32) 98.9 (12.83-1.32)	Depositor EDS
R_{merge}	0.03	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.05 (at 1.32Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.213 , 0.242 0.231 , 0.260	Depositor DCC
R_{free} test set	8967 reflections (10.06%)	DCC
Wilson B-factor (Å ²)	12.9	Xtriage
Anisotropy	0.588	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.46 , 58.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 89527 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	3659	wwPDB-VP
Average B, all atoms (Å ²)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.86% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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

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	1.07	5/3256 (0.2%)	1.09	11/4429 (0.2%)

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

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	41	VAL	CB-CG1	-6.65	1.38	1.52
1	A	152	MET	SD-CE	-6.23	1.43	1.77
1	A	310	VAL	CB-CG2	-6.02	1.40	1.52
1	A	391	ARG	NE-CZ	-5.85	1.25	1.33
1	A	275	PHE	CD2-CE2	-5.68	1.27	1.39

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	259	ASP	CB-CG-OD2	7.40	124.96	118.30
1	A	246	ASP	CB-CG-OD2	7.28	124.85	118.30
1	A	366	ASP	CB-CG-OD2	7.07	124.66	118.30
1	A	336	ASP	CB-CG-OD2	6.95	124.56	118.30
1	A	209	ASP	CB-CG-OD2	6.06	123.75	118.30
1	A	236	ASP	CB-CG-OD2	5.86	123.58	118.30
1	A	377	ASP	CB-CG-OD2	5.72	123.44	118.30
1	A	285	ASP	CB-CG-OD2	5.54	123.28	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	259	ASP	OD1-CG-OD2	-5.36	113.12	123.30
1	A	251	ASP	CB-CG-OD2	-5.26	113.56	118.30
1	A	375	ASP	CB-CG-OD2	5.11	122.90	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	215	ASN	Mainchain

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	3198	0	2972	40	0
2	A	14	0	13	0	0
3	A	447	0	0	15	3
All	All	3659	0	2985	40	3

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 (40) 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:377:ASP:OD2	3:A:2408:HOH:O	1.58	1.20
1:A:334:PHE:O	3:A:2383:HOH:O	1.92	0.87
1:A:132:ASN:ND2	3:A:2213:HOH:O	2.03	0.85
1:A:384:PRO:O	3:A:2415:HOH:O	1.93	0.84
1:A:384:PRO:CB	3:A:2415:HOH:O	2.28	0.82
1:A:333:ALA:HB2	3:A:2359:HOH:O	1.88	0.74
1:A:400:VAL:HG21	1:A:403:ASP:OD2	1.88	0.74
1:A:400:VAL:HG22	3:A:2427:HOH:O	1.89	0.73
1:A:134:PRO:HD2	1:A:137:LEU:HD12	1.70	0.73
1:A:307:GLN:NE2	3:A:2352:HOH:O	2.24	0.71

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:400:VAL:CG2	1:A:403:ASP:OD2	2.42	0.68
1:A:384:PRO:HB2	3:A:2415:HOH:O	1.91	0.66
1:A:377:ASP:CG	3:A:2408:HOH:O	2.19	0.66
1:A:267:MET:HA	1:A:267:MET:HE2	1.83	0.60
1:A:222:PRO:HD2	1:A:250:CYS:O	2.02	0.59
1:A:384:PRO:C	3:A:2415:HOH:O	2.37	0.59
1:A:55:GLU:OE2	1:A:191:GLU:OE2	2.21	0.58
1:A:226:THR:HG23	1:A:343:LYS:HD2	1.86	0.58
1:A:379:PRO:HB2	1:A:381:ASP:OD1	2.05	0.56
1:A:267:MET:CE	1:A:267:MET:HA	2.37	0.55
1:A:380:THR:HG22	3:A:2407:HOH:O	2.07	0.54
1:A:13:GLN:HE22	1:A:31:LYS:HE2	1.73	0.53
1:A:383:ASP:OD2	1:A:385:SER:OG	2.27	0.53
1:A:378:TYR:HA	1:A:379:PRO:C	2.32	0.49
1:A:389:VAL:O	1:A:391:ARG:HD2	2.13	0.48
1:A:227:THR:HG21	1:A:244:LEU:HD13	1.96	0.48
1:A:347:LYS:O	1:A:351:GLU:HG3	2.15	0.47
1:A:2:GLN:HG3	1:A:44:THR:HG21	1.96	0.47
1:A:304:LYS:NZ	3:A:2346:HOH:O	2.48	0.46
1:A:262:PHE:CG	1:A:309:SER:HB3	2.51	0.45
1:A:353:LEU:HD23	1:A:357:MET:SD	2.58	0.44
1:A:400:VAL:O	1:A:400:VAL:HG23	2.18	0.44
1:A:400:VAL:CG2	1:A:403:ASP:CG	2.86	0.43
1:A:344:GLY:HA3	1:A:348:GLN:HB2	2.00	0.43
1:A:343:LYS:O	1:A:348:GLN:HG2	2.19	0.43
1:A:335:GLY:C	3:A:2385:HOH:O	2.57	0.42
1:A:319:VAL:HB	3:A:2366:HOH:O	2.20	0.41
1:A:267:MET:CE	1:A:267:MET:CA	2.99	0.41
1:A:260:LYS:HE3	1:A:260:LYS:HB3	1.27	0.41
1:A:132:ASN:HA	1:A:132:ASN:HD22	1.65	0.40

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:2039:HOH:O	3:A:2390:HOH:O[1_545]	0.49	1.71
3:A:2220:HOH:O	3:A:2238:HOH:O[3_455]	0.58	1.62
3:A:2443:HOH:O	3:A:2443:HOH:O[2_655]	2.11	0.09

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	429/431 (100%)	421 (98%)	8 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	351/351 (100%)	347 (99%)	4 (1%)	80	49

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	37	ASN
1	A	74	ASP
1	A	132	ASN
1	A	330	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	13	GLN
1	A	17	GLN
1	A	18	GLN
1	A	115	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	132	ASN
1	A	307	GLN
1	A	410	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

1 non-standard protein/DNA/RNA residue is 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	$\# Z > 2$
1	PCA	A	1	1	7,8,9	1.34	2 (28%)	9,10,12	1.20	2 (22%)

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
1	PCA	A	1	1	-	0/0/11/13	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1	PCA	CA-N	2.13	1.49	1.46
1	A	1	PCA	CD-N	2.63	1.42	1.33

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1	PCA	O-C-CA	-2.19	119.66	125.44
1	A	1	PCA	CB-CA-C	-2.11	109.88	112.76

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

1 ligand is 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	NAG	A	435	1	14,14,15	1.04	1 (7%)	15,19,21	1.58	4 (26%)

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	NAG	A	435	1	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	435	NAG	C1-C2	2.18	1.55	1.52

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	435	NAG	C1-O5-C5	-2.83	108.66	112.25
2	A	435	NAG	C6-C5-C4	-2.58	106.66	113.02
2	A	435	NAG	O7-C7-C8	-2.23	117.97	122.06
2	A	435	NAG	C4-C3-C2	2.52	115.15	111.23

There are no chirality outliers.

There are no torsion outliers.

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	430/431 (99%)	0.72	44 (10%) 9 7	8, 17, 25, 32	0

All (44) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	228	THR	4.7
1	A	394	CYS	4.7
1	A	335	GLY	4.5
1	A	400	VAL	4.4
1	A	334	PHE	4.3
1	A	45	SER	4.3
1	A	227	THR	4.1
1	A	386	ALA	4.0
1	A	332	THR	3.8
1	A	408	VAL	3.8
1	A	395	ALA	3.7
1	A	188	ASN	3.5
1	A	385	SER	3.4
1	A	378	TYR	3.2
1	A	407	GLN	3.1
1	A	313	ILE	3.1
1	A	236	ASP	3.1
1	A	387	PRO	3.1
1	A	389	VAL	3.0
1	A	195	ASN	2.9
1	A	384	PRO	2.9
1	A	243	GLY	2.8
1	A	390	ALA	2.8
1	A	397	THR	2.7
1	A	226	THR	2.6
1	A	369	ALA	2.6
1	A	21	THR	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	404	VAL	2.5
1	A	383	ASP	2.4
1	A	376	SER	2.4
1	A	333	ALA	2.4
1	A	132	ASN	2.3
1	A	267	MET	2.2
1	A	368	ALA	2.2
1	A	240	ARG	2.2
1	A	266	GLY	2.2
1	A	392	GLY	2.1
1	A	331	LYS	2.1
1	A	326	PHE	2.1
1	A	245	CYS	2.1
1	A	370	ASN	2.1
1	A	255	PHE	2.1
1	A	388	GLY	2.0
1	A	373	TRP	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
1	PCA	A	1	8/9	0.96	0.09	-	12,12,13,13	0

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
2	NAG	A	435	14/15	0.78	0.18	2.54	24,25,32,32	0

6.5 Other polymers [i](#)

There are no such residues in this entry.