



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 6, 2026 – 12:45 PM UTC

PDB ID : 6D5B / pdb\_00006d5b  
Title : Structure of Caldicellulosiruptor danielii CBM3 module of glycoside hydrolase  
WP\_045175321  
Authors : Alahuhta, P.M.; Lunin, V.V.  
Deposited on : 2018-04-19  
Resolution : 2.00 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 : 4-5-2 with Phenix2.0  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

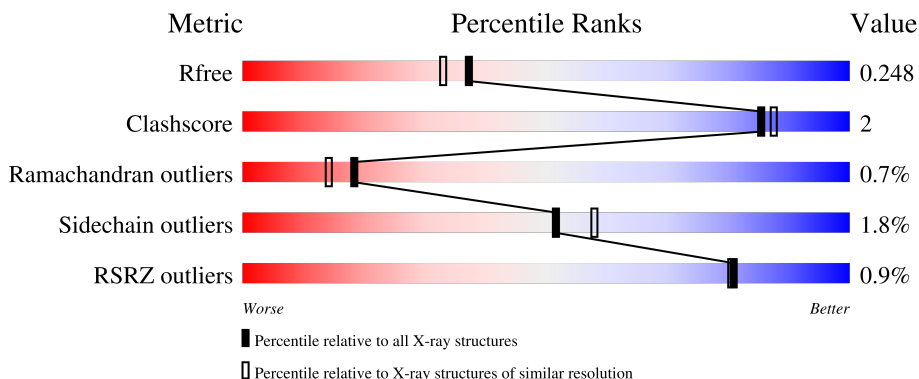
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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.




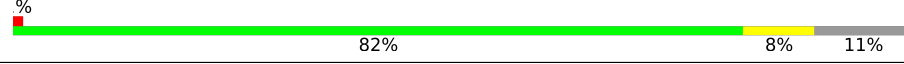

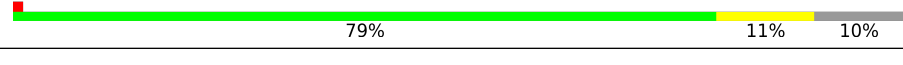
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	180053	10052 (2.00-2.00)
Clashscore	190562	11152 (2.00-2.00)
Ramachandran outliers	187476	11031 (2.00-2.00)
Sidechain outliers	187428	11029 (2.00-2.00)
RSRZ outliers	180081	10067 (2.00-2.00)

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  $\geq 3$ , 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	170	 82% 8% 10%
1	B	170	 85% 5% 9%
1	C	170	 85% 5% 10%
1	D	170	 81% 9% 10%
1	E	170	 82% 7% 10%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	170	 <p>% 81% 9% 11%</p>
1	G	170	 <p>% 82% 8% 10%</p>
1	H	170	 <p>% 82% 8% 11%</p>
1	I	170	 <p>2% 84% 6% 10%</p>
1	J	170	 <p>% 85% 5% 10%</p>
1	K	170	 <p>% 85% 5% 10%</p>
1	L	170	 <p>% 79% 11% 10%</p>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 16774 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 glycoside hydrolase WP\_045175321.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	153	1208	762	204	241	1	0	1	0
1	B	154	1209	763	204	241	1	0	0	0
1	C	153	1208	762	204	241	1	0	1	0
1	D	153	1202	759	203	239	1	0	0	0
1	E	153	1214	765	205	243	1	0	2	0
1	F	152	1196	756	202	237	1	0	0	0
1	G	153	1208	762	204	241	1	0	1	0
1	H	152	1196	756	202	237	1	0	0	0
1	I	153	1202	759	203	239	1	0	0	0
1	J	153	1223	771	207	244	1	0	3	0
1	K	153	1202	759	203	239	1	0	0	0
1	L	153	1216	766	205	244	1	0	2	0

- Molecule 2 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Ca 1	0	0
2	B	1	Total 1	Ca 1	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	1	Total Ca 1 1	0	0
2	D	1	Total Ca 1 1	0	0
2	E	1	Total Ca 1 1	0	0
2	F	1	Total Ca 1 1	0	0
2	G	1	Total Ca 1 1	0	0
2	H	1	Total Ca 1 1	0	0
2	I	1	Total Ca 1 1	0	0
2	J	1	Total Ca 1 1	0	0
2	K	1	Total Ca 1 1	0	0
2	L	1	Total Ca 1 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	182	Total O 182 182	0	0
3	B	186	Total O 186 186	0	0
3	C	172	Total O 172 172	0	0
3	D	189	Total O 189 189	0	0
3	E	223	Total O 223 223	0	0
3	F	209	Total O 209 209	0	0
3	G	181	Total O 181 181	0	0
3	H	198	Total O 198 198	0	0
3	I	189	Total O 189 189	0	0

*Continued on next page...*


*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
3	J	146	Total 146	O 146	0	0
3	K	215	Total 215	O 215	0	0
3	L	188	Total 188	O 188	0	0

### 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.

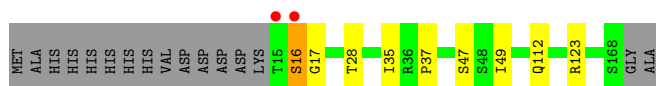
- Molecule 1: glycoside hydrolase WP\_045175321

Chain A: 




- Molecule 1: glycoside hydrolase WP\_045175321

Chain B: 




- Molecule 1: glycoside hydrolase WP\_045175321

Chain C: 




- Molecule 1: glycoside hydrolase WP\_045175321

Chain D: 

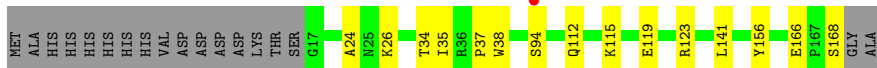
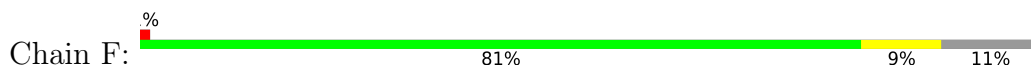


- Molecule 1: glycoside hydrolase WP\_045175321

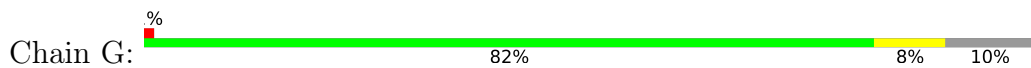
Chain E: 



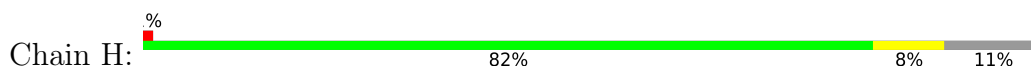
- Molecule 1: glycoside hydrolase WP\_045175321



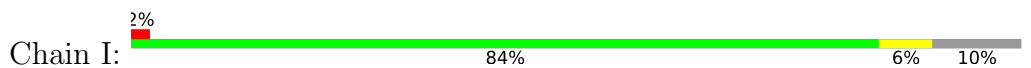
- Molecule 1: glycoside hydrolase WP\_045175321



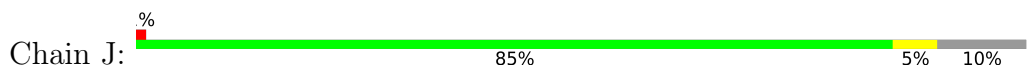
- Molecule 1: glycoside hydrolase WP\_045175321



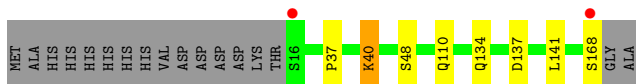
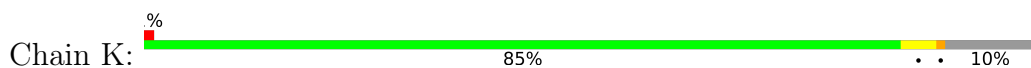
- Molecule 1: glycoside hydrolase WP\_045175321



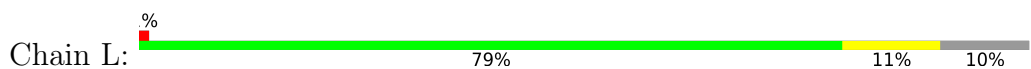
- Molecule 1: glycoside hydrolase WP\_045175321



- Molecule 1: glycoside hydrolase WP\_045175321



- Molecule 1: glycoside hydrolase WP\_045175321



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.17Å 86.20Å 86.57Å 83.87° 83.93° 60.55°	Depositor
Resolution (Å)	85.91 – 2.00 85.91 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.5 (85.91-2.00) 98.5 (85.91-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.12	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.71 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.178 , 0.243 0.185 , 0.248	Depositor DCC
$R_{free}$ test set	7221 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	24.2	Xtrriage
Anisotropy	0.249	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 34.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.034 for -k,-h,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	16774	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.92% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: CA

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	1.02	0/1236	0.95	1/1680 (0.1%)
1	B	1.08	0/1237	0.98	0/1682
1	C	0.98	0/1236	0.99	0/1680
1	D	1.07	1/1230 (0.1%)	0.96	0/1672
1	E	1.20	2/1242 (0.2%)	0.97	0/1688
1	F	1.11	1/1224 (0.1%)	1.01	0/1664
1	G	1.07	1/1236 (0.1%)	0.96	0/1680
1	H	1.17	1/1224 (0.1%)	0.99	0/1664
1	I	1.05	0/1230	0.96	0/1672
1	J	1.08	0/1251	0.94	0/1699
1	K	1.12	0/1230	1.00	0/1672
1	L	1.13	0/1244	1.02	1/1691 (0.1%)
All	All	1.09	6/14820 (0.0%)	0.98	2/20144 (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	E	0	1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	160	VAL	C-O	-5.96	1.17	1.24
1	H	96	ALA	CA-C	-5.89	1.45	1.52
1	G	86	PHE	N-CA	-5.48	1.39	1.46
1	E	50	ASP	C-O	-5.33	1.17	1.24
1	E	41	VAL	C-O	-5.26	1.18	1.24

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	L	17	GLY	N-CA-C	6.27	119.14	111.93
1	A	123	ARG	NE-CZ-NH2	5.00	123.70	119.20

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	98	TYR	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	1208	0	1155	6	0
1	B	1209	0	1158	3	0
1	C	1208	0	1155	3	0
1	D	1202	0	1151	7	0
1	E	1214	0	1159	7	0
1	F	1196	0	1146	6	0
1	G	1208	0	1155	6	0
1	H	1196	0	1146	4	0
1	I	1202	0	1151	5	0
1	J	1223	0	1171	5	0
1	K	1202	0	1151	3	0
1	L	1216	0	1158	10	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
2	I	1	0	0	0	0
2	J	1	0	0	0	0
2	K	1	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	L	1	0	0	0	0
3	A	182	0	0	0	0
3	B	186	0	0	1	0
3	C	172	0	0	1	0
3	D	189	0	0	0	0
3	E	223	0	0	2	0
3	F	209	0	0	1	0
3	G	181	0	0	1	0
3	H	198	0	0	1	0
3	I	189	0	0	1	0
3	J	146	0	0	1	0
3	K	215	0	0	2	0
3	L	188	0	0	1	0
All	All	16774	0	13856	62	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 2.

The worst 5 of 62 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:C:97:ASP:OD2	3:C:301:HOH:O	2.05	0.73
1:F:112:GLN:HB2	1:F:115:LYS:HG3	1.80	0.63
1:K:110:GLN:NE2	3:K:303:HOH:O	2.35	0.60
1:D:155:ALA:HB3	1:D:163:TRP:HB3	1.85	0.57
1:E:47:SER:HB3	3:E:403:HOH:O	2.04	0.56

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	152/170 (89%)	146 (96%)	5 (3%)	1 (1%)	18	14
1	B	152/170 (89%)	145 (95%)	5 (3%)	2 (1%)	9	5
1	C	152/170 (89%)	146 (96%)	5 (3%)	1 (1%)	18	14
1	D	151/170 (89%)	146 (97%)	4 (3%)	1 (1%)	18	14
1	E	153/170 (90%)	146 (95%)	7 (5%)	0	100	100
1	F	150/170 (88%)	147 (98%)	1 (1%)	2 (1%)	9	5
1	G	152/170 (89%)	147 (97%)	4 (3%)	1 (1%)	18	14
1	H	150/170 (88%)	145 (97%)	4 (3%)	1 (1%)	18	14
1	I	151/170 (89%)	148 (98%)	2 (1%)	1 (1%)	18	14
1	J	154/170 (91%)	149 (97%)	5 (3%)	0	100	100
1	K	151/170 (89%)	149 (99%)	1 (1%)	1 (1%)	18	14
1	L	153/170 (90%)	149 (97%)	3 (2%)	1 (1%)	18	14
All	All	1821/2040 (89%)	1763 (97%)	46 (2%)	12 (1%)	18	14

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	16	SER
1	F	94	SER
1	A	37	PRO
1	B	37	PRO
1	G	37	PRO

### 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	132/145 (91%)	131 (99%)	1 (1%)	73	80
1	B	132/145 (91%)	129 (98%)	3 (2%)	44	49
1	C	132/145 (91%)	129 (98%)	3 (2%)	44	49
1	D	131/145 (90%)	128 (98%)	3 (2%)	44	49

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	133/145 (92%)	131 (98%)	2 (2%)	57	64
1	F	130/145 (90%)	128 (98%)	2 (2%)	57	64
1	G	132/145 (91%)	131 (99%)	1 (1%)	73	80
1	H	130/145 (90%)	125 (96%)	5 (4%)	29	29
1	I	131/145 (90%)	129 (98%)	2 (2%)	57	64
1	J	134/145 (92%)	133 (99%)	1 (1%)	76	82
1	K	131/145 (90%)	127 (97%)	4 (3%)	35	37
1	L	133/145 (92%)	131 (98%)	2 (2%)	57	64
All	All	1581/1740 (91%)	1552 (98%)	29 (2%)	51	58

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	141	LEU
1	L	141	LEU
1	H	115	LYS
1	K	48	SER
1	H	69	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	133	ASN
1	L	76	GLN
1	L	131	ASN
1	L	81	ASN
1	D	68	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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 12 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	153/170 (90%)	-0.30	0 <b>100</b> <b>100</b>	17, 27, 39, 71	1 (0%)
1	B	154/170 (90%)	-0.23	2 (1%) 75 74	19, 27, 40, 71	0
1	C	153/170 (90%)	-0.13	1 (0%) 84 84	15, 30, 44, 79	1 (0%)
1	D	153/170 (90%)	-0.28	1 (0%) 84 84	16, 26, 38, 64	0
1	E	153/170 (90%)	-0.37	1 (0%) 84 84	11, 23, 35, 58	2 (1%)
1	F	152/170 (89%)	-0.35	1 (0%) 84 84	17, 25, 36, 50	0
1	G	153/170 (90%)	-0.11	1 (0%) 84 84	14, 29, 46, 66	1 (0%)
1	H	152/170 (89%)	-0.30	1 (0%) 84 84	17, 24, 39, 50	0
1	I	153/170 (90%)	-0.20	3 (1%) 65 64	20, 27, 43, 67	0
1	J	153/170 (90%)	-0.18	2 (1%) 75 74	14, 29, 49, 64	3 (1%)
1	K	153/170 (90%)	-0.34	2 (1%) 75 74	16, 23, 39, 67	0
1	L	153/170 (90%)	-0.24	1 (0%) 84 84	14, 25, 39, 60	2 (1%)
All	All	1835/2040 (89%)	-0.25	16 (0%) <b>81</b> <b>80</b>	11, 26, 41, 79	10 (0%)

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	168	SER	3.3
1	C	16	SER	3.0
1	J	168	SER	2.7
1	G	16	SER	2.7
1	B	16	SER	2.5

### 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 oligosaccharides 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, 95<sup>th</sup> 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(Å <sup>2</sup> )	Q<0.9
2	CA	G	201	1/1	0.97	0.04	32,32,32,32	0
2	CA	J	201	1/1	0.97	0.04	36,36,36,36	0
2	CA	D	201	1/1	0.98	0.04	25,25,25,25	0
2	CA	A	201	1/1	0.98	0.04	31,31,31,31	0
2	CA	C	201	1/1	0.98	0.04	33,33,33,33	0
2	CA	L	201	1/1	0.98	0.04	30,30,30,30	0
2	CA	H	201	1/1	0.99	0.02	23,23,23,23	0
2	CA	I	201	1/1	0.99	0.04	33,33,33,33	0
2	CA	F	201	1/1	0.99	0.04	22,22,22,22	0
2	CA	K	201	1/1	0.99	0.03	23,23,23,23	0
2	CA	B	201	1/1	0.99	0.03	29,29,29,29	0
2	CA	E	201	1/1	1.00	0.04	21,21,21,21	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.