



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 6, 2025 – 10:22 PM JST

PDB ID : 9W4B / pdb\_00009w4b  
Title : Crystal structure of beta-glucosidase CaBGL  
Authors : You, C.; Feng, Y.G.  
Deposited on : 2025-07-31  
Resolution : 2.34 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
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) : 2.46

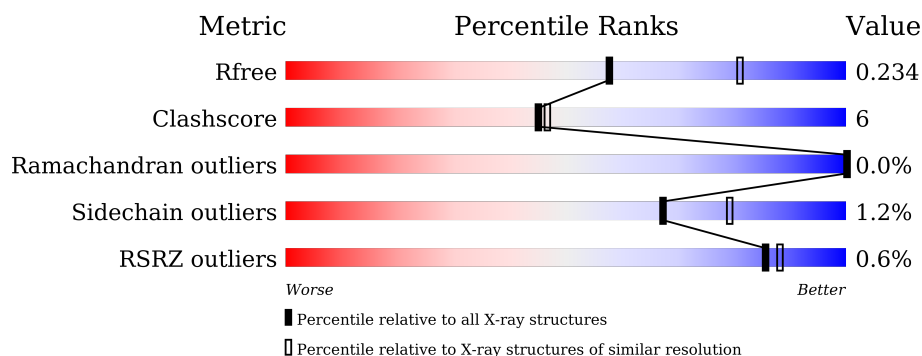
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.34 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))
$R_{free}$	164625	2747 (2.36-2.32)
Clashscore	180529	2936 (2.36-2.32)
Ramachandran outliers	177936	2912 (2.36-2.32)
Sidechain outliers	177891	2912 (2.36-2.32)
RSRZ outliers	164620	2747 (2.36-2.32)

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	454	<div> <div></div> <div>93%7%</div> </div>
1	B	454	<div> <div>%</div> <div>83%16%</div> </div>
1	C	454	<div> <div></div> <div>91%9%</div> </div>
1	D	454	<div> <div></div> <div>86%13%</div> </div>
1	E	454	<div> <div></div> <div>88%11%</div> </div>
1	F	454	<div> <div></div> <div>82%17%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	454	 87% 12%
1	H	454	 88% 11%
1	I	454	 84% 15%
1	J	454	 83% 16%
1	K	454	 72% 28%
1	L	454	 78% 21%

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	GOL	D	501	-	-	X	-
3	SO4	H	502	-	-	X	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 46367 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 beta-glucosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	453	Total	C	N	O	S	0	0	0
			3788	2468	621	687	12			
1	B	453	Total	C	N	O	S	0	0	0
			3788	2468	621	687	12			
1	C	453	Total	C	N	O	S	0	0	0
			3788	2468	621	687	12			
1	D	453	Total	C	N	O	S	0	0	0
			3788	2468	621	687	12			
1	E	453	Total	C	N	O	S	0	0	0
			3788	2468	621	687	12			
1	F	452	Total	C	N	O	S	0	0	0
			3780	2463	620	686	11			
1	G	453	Total	C	N	O	S	0	0	0
			3788	2468	621	687	12			
1	H	452	Total	C	N	O	S	0	0	0
			3780	2463	620	686	11			
1	I	452	Total	C	N	O	S	0	0	0
			3780	2463	620	686	11			
1	J	453	Total	C	N	O	S	0	0	0
			3788	2468	621	687	12			
1	K	452	Total	C	N	O	S	0	0	0
			3780	2463	620	686	11			
1	L	453	Total	C	N	O	S	0	0	0
			3788	2468	621	687	12			

There are 12 discrepancies between the modelled and reference sequences:

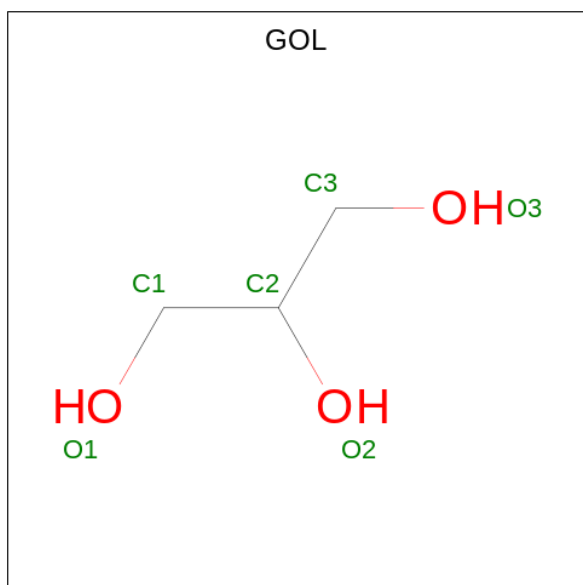
Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP I7DLX2
B	0	SER	-	expression tag	UNP I7DLX2
C	0	SER	-	expression tag	UNP I7DLX2
D	0	SER	-	expression tag	UNP I7DLX2
E	0	SER	-	expression tag	UNP I7DLX2

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	0	SER	-	expression tag	UNP I7DLX2
G	0	SER	-	expression tag	UNP I7DLX2
H	0	SER	-	expression tag	UNP I7DLX2
I	0	SER	-	expression tag	UNP I7DLX2
J	0	SER	-	expression tag	UNP I7DLX2
K	0	SER	-	expression tag	UNP I7DLX2
L	0	SER	-	expression tag	UNP I7DLX2

- Molecule 2 is GLYCEROL (CCD ID: GOL) (formula:  $C_3H_8O_3$ ).



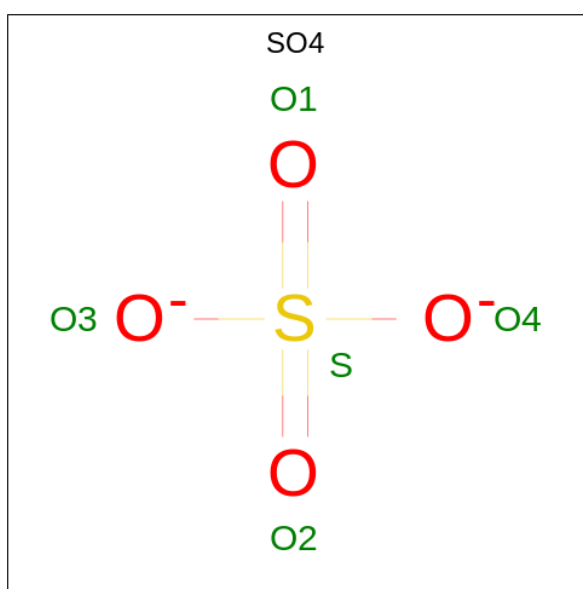
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		
2	C	1	Total	C	O	0	0
			6	3	3		
2	D	1	Total	C	O	0	0
			6	3	3		
2	E	1	Total	C	O	0	0
			6	3	3		
2	F	1	Total	C	O	0	0
			6	3	3		
2	G	1	Total	C	O	0	0
			6	3	3		
2	H	1	Total	C	O	0	0
			6	3	3		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	I	1	Total	C	O	0	0
			6	3	3		
2	J	1	Total	C	O	0	0
			6	3	3		
2	K	1	Total	C	O	0	0
			6	3	3		
2	L	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is SULFATE ION (CCD ID: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	D	1	Total	O	S	0	0
			5	4	1		
3	E	1	Total	O	S	0	0
			5	4	1		
3	E	1	Total	O	S	0	0
			5	4	1		
3	F	1	Total	O	S	0	0
			5	4	1		
3	F	1	Total	O	S	0	0
			5	4	1		
3	G	1	Total	O	S	0	0
			5	4	1		
3	G	1	Total	O	S	0	0
			5	4	1		
3	H	1	Total	O	S	0	0
			5	4	1		
3	H	1	Total	O	S	0	0
			5	4	1		
3	I	1	Total	O	S	0	0
			5	4	1		
3	I	1	Total	O	S	0	0
			5	4	1		
3	J	1	Total	O	S	0	0
			5	4	1		
3	J	1	Total	O	S	0	0
			5	4	1		
3	K	1	Total	O	S	0	0
			5	4	1		
3	K	1	Total	O	S	0	0
			5	4	1		
3	L	1	Total	O	S	0	0
			5	4	1		
3	L	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	120	Total	O	0	0
			120	120		
4	B	29	Total	O	0	0
			29	29		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	136	Total 136	O 136	0	0
4	D	91	Total 91	O 91	0	0
4	E	68	Total 68	O 68	0	0
4	F	66	Total 66	O 66	0	0
4	G	44	Total 44	O 44	0	0
4	H	61	Total 61	O 61	0	0
4	I	50	Total 50	O 50	0	0
4	J	35	Total 35	O 35	0	0
4	K	26	Total 26	O 26	0	0
4	L	25	Total 25	O 25	0	0

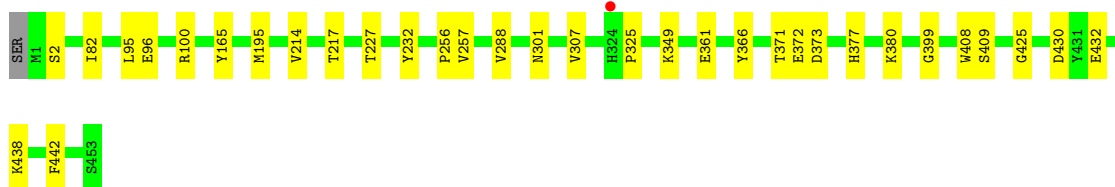


### 3 Residue-property plots


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.

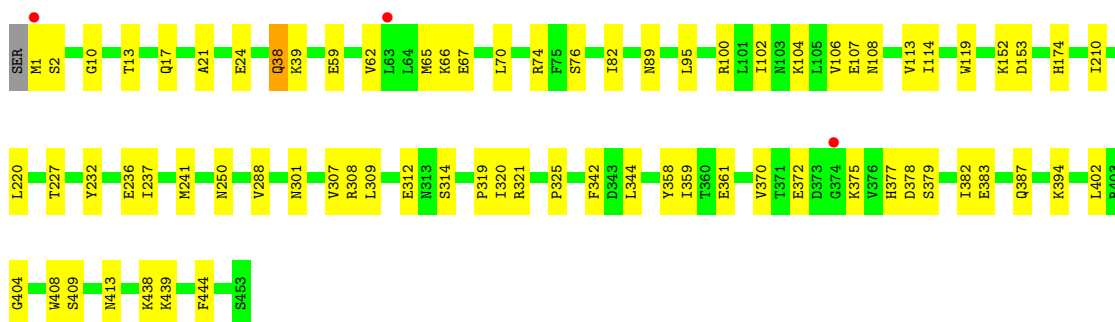
- Molecule 1: beta-glucosidase

Chain A: 



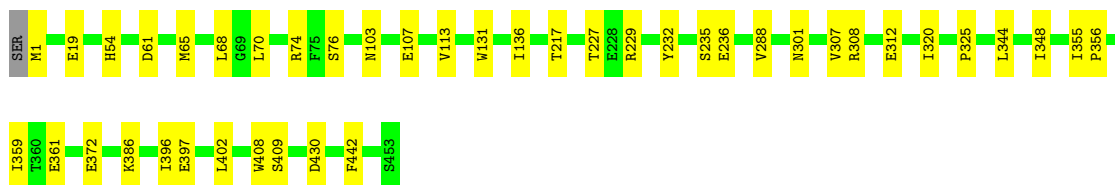
- Molecule 1: beta-glucosidase

Chain B: 




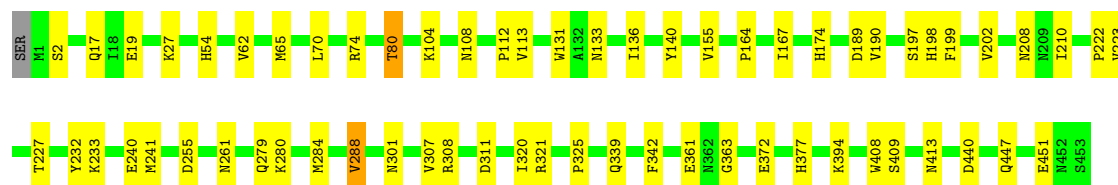
- Molecule 1: beta-glucosidase

Chain C: 



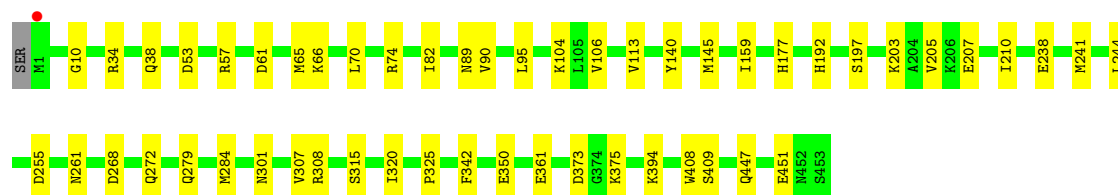
- Molecule 1: beta-glucosidase

Chain D: 



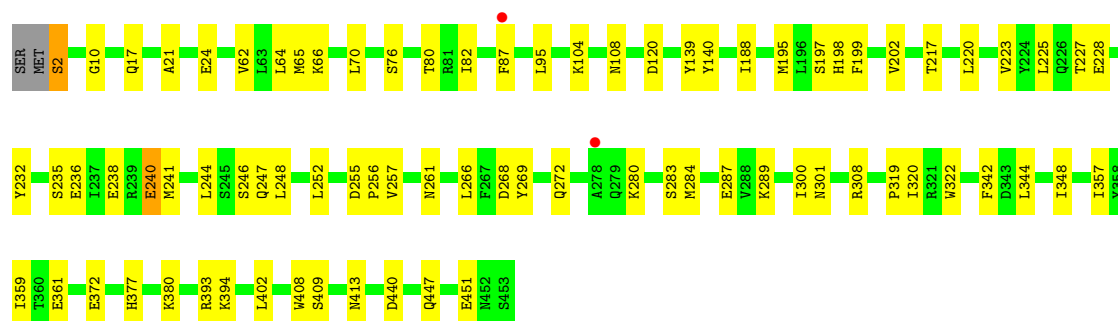
- Molecule 1: beta-glucosidase

Chain E: 88% 11%



- Molecule 1: beta-glucosidase

Chain F: 82% 17%



- Molecule 1: beta-glucosidase

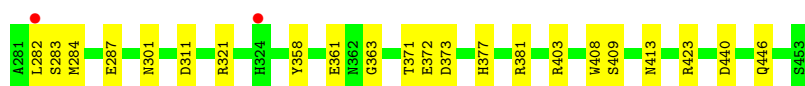
Chain G: 87% 12%



- Molecule 1: beta-glucosidase

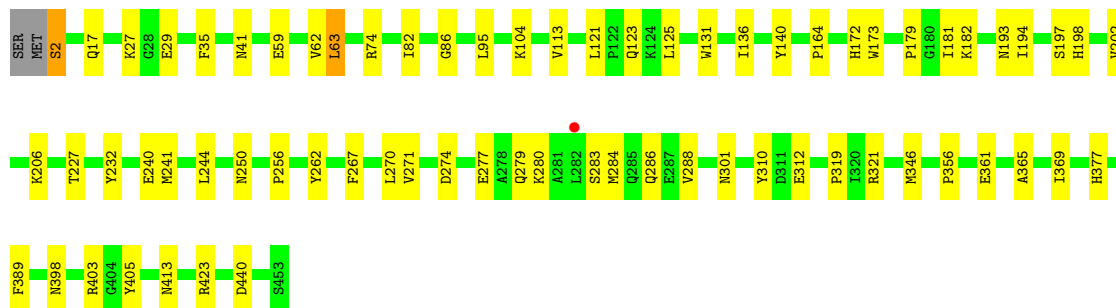
Chain H: 88% 11%





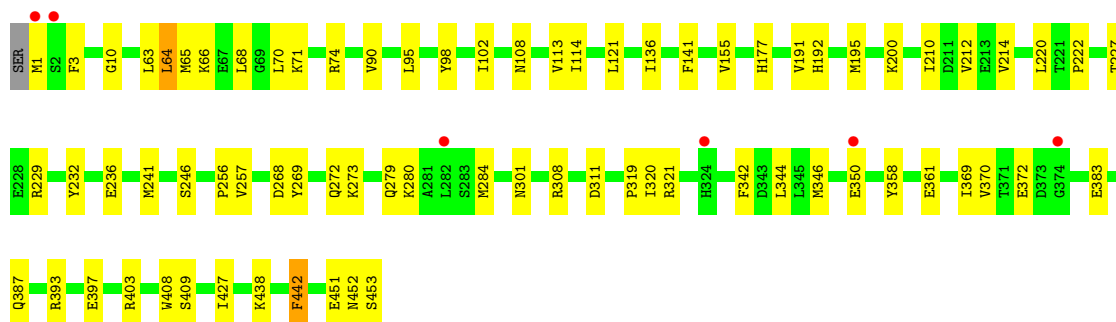
- Molecule 1: beta-glucosidase

Chain I:   84% 15%



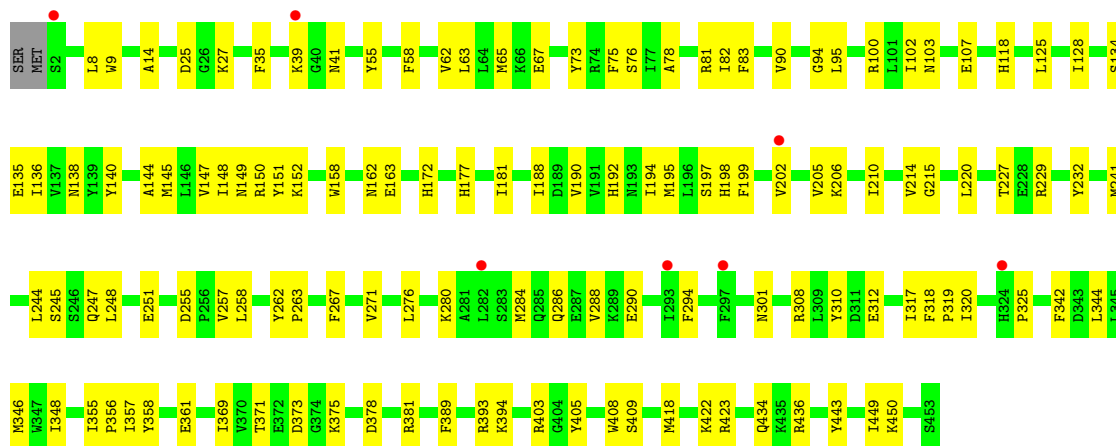
- Molecule 1: beta-glucosidase

Chain J:   83% 16%

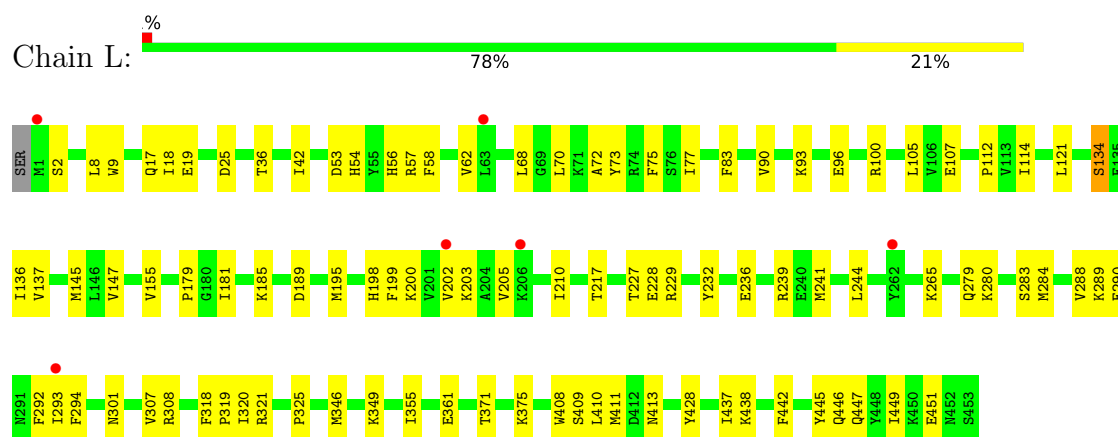


- Molecule 1: beta-glucosidase

Chain K:   72% 28%



- Molecule 1: beta-glucosidase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	138.84Å 221.66Å 236.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.74 – 2.34 28.74 – 2.34	Depositor EDS
% Data completeness (in resolution range)	99.8 (28.74-2.34) 99.7 (28.74-2.34)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.94 (at 2.34Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487, PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.188 , 0.233 0.189 , 0.234	Depositor DCC
$R_{free}$ test set	14848 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	51.6	Xtriage
Anisotropy	0.250	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 42.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	46367	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.47% 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: GOL, SO4

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	0.42	0/3900	0.58	0/5285
1	B	0.32	0/3900	0.49	0/5285
1	C	0.42	0/3900	0.59	0/5285
1	D	0.39	0/3900	0.56	0/5285
1	E	0.36	0/3900	0.54	0/5285
1	F	0.37	0/3892	0.54	0/5275
1	G	0.35	0/3900	0.53	0/5285
1	H	0.34	0/3892	0.51	0/5275
1	I	0.33	0/3892	0.49	0/5275
1	J	0.31	0/3900	0.49	0/5285
1	K	0.32	0/3892	0.53	0/5275
1	L	0.28	0/3900	0.47	0/5285
All	All	0.35	0/46768	0.53	0/63380

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	3788	0	3654	18	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	3788	0	3654	42	0
1	C	3788	0	3654	26	0
1	D	3788	0	3654	43	0
1	E	3788	0	3654	29	0
1	F	3780	0	3642	48	0
1	G	3788	0	3654	30	0
1	H	3780	0	3642	34	0
1	I	3780	0	3642	47	0
1	J	3788	0	3654	55	0
1	K	3780	0	3642	88	0
1	L	3788	0	3654	60	0
2	A	6	0	7	0	0
2	B	6	0	8	1	0
2	C	6	0	8	0	0
2	D	6	0	8	4	0
2	E	6	0	8	0	0
2	F	6	0	8	2	0
2	G	6	0	8	0	0
2	H	6	0	8	0	0
2	I	6	0	8	1	0
2	J	6	0	8	2	0
2	K	6	0	8	0	0
2	L	6	0	8	0	0
3	A	10	0	0	0	0
3	B	10	0	0	0	0
3	C	10	0	0	0	0
3	D	10	0	0	0	0
3	E	10	0	0	0	0
3	F	10	0	0	0	0
3	G	10	0	0	1	0
3	H	10	0	0	4	0
3	I	10	0	0	0	0
3	J	10	0	0	1	0
3	K	10	0	0	0	0
3	L	10	0	0	2	0
4	A	120	0	0	0	0
4	B	29	0	0	1	0
4	C	136	0	0	1	0
4	D	91	0	0	6	0
4	E	68	0	0	1	0
4	F	66	0	0	0	0
4	G	44	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	H	61	0	0	0	0
4	I	50	0	0	1	0
4	J	35	0	0	1	0
4	K	26	0	0	0	0
4	L	25	0	0	0	0
All	All	46367	0	43895	520	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.

The worst 5 of 520 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:D:223:VAL:H	2:D:501:GOL:H11	1.33	0.91
1:A:195:MET:HE2	1:A:256:PRO:HB2	1.60	0.82
1:K:27:LYS:HD2	1:K:81:ARG:HG2	1.60	0.82
1:J:195:MET:HE2	1:J:256:PRO:HB2	1.63	0.81
1:K:103:ASN:O	1:K:107:GLU:HG3	1.82	0.79

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	451/454 (99%)	436 (97%)	15 (3%)	0	100	100
1	B	451/454 (99%)	431 (96%)	19 (4%)	1 (0%)	44	51
1	C	451/454 (99%)	436 (97%)	15 (3%)	0	100	100
1	D	451/454 (99%)	439 (97%)	12 (3%)	0	100	100
1	E	451/454 (99%)	436 (97%)	15 (3%)	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	450/454 (99%)	435 (97%)	15 (3%)	0	100	100
1	G	451/454 (99%)	439 (97%)	12 (3%)	0	100	100
1	H	450/454 (99%)	435 (97%)	15 (3%)	0	100	100
1	I	450/454 (99%)	435 (97%)	15 (3%)	0	100	100
1	J	451/454 (99%)	434 (96%)	17 (4%)	0	100	100
1	K	450/454 (99%)	432 (96%)	18 (4%)	0	100	100
1	L	451/454 (99%)	439 (97%)	12 (3%)	0	100	100
All	All	5408/5448 (99%)	5227 (97%)	180 (3%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	119	TRP

### 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	400/401 (100%)	395 (99%)	5 (1%)	65	77
1	B	400/401 (100%)	393 (98%)	7 (2%)	54	66
1	C	400/401 (100%)	396 (99%)	4 (1%)	73	83
1	D	400/401 (100%)	397 (99%)	3 (1%)	79	87
1	E	400/401 (100%)	395 (99%)	5 (1%)	65	77
1	F	399/401 (100%)	391 (98%)	8 (2%)	50	62
1	G	400/401 (100%)	398 (100%)	2 (0%)	86	92
1	H	399/401 (100%)	398 (100%)	1 (0%)	91	95
1	I	399/401 (100%)	394 (99%)	5 (1%)	65	77
1	J	400/401 (100%)	395 (99%)	5 (1%)	65	77
1	K	399/401 (100%)	395 (99%)	4 (1%)	73	83

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	L	400/401 (100%)	391 (98%)	9 (2%)	45 56
All	All	4796/4812 (100%)	4738 (99%)	58 (1%)	67 79

5 of 58 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	217	THR
1	L	217	THR
1	I	2	SER
1	L	185	LYS
1	L	2	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 59 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	123	GLN
1	L	126	GLN
1	H	126	GLN
1	L	118	HIS
1	J	452	ASN

### 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](#)

36 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	$\# Z  > 2$
3	SO4	E	502	-	4,4,4	0.18	0	6,6,6	0.31	0
3	SO4	G	502	-	4,4,4	0.12	0	6,6,6	0.56	0
3	SO4	F	503	-	4,4,4	0.15	0	6,6,6	0.26	0
2	GOL	K	501	-	5,5,5	0.92	0	5,5,5	1.11	1 (20%)
2	GOL	L	501	-	5,5,5	1.03	0	5,5,5	0.96	0
2	GOL	F	501	-	5,5,5	1.47	0	5,5,5	0.66	0
3	SO4	G	503	-	4,4,4	0.17	0	6,6,6	0.29	0
3	SO4	H	503	-	4,4,4	0.17	0	6,6,6	0.14	0
2	GOL	I	501	-	5,5,5	1.04	0	5,5,5	0.87	0
3	SO4	E	503	-	4,4,4	0.15	0	6,6,6	0.20	0
2	GOL	C	501	-	5,5,5	1.28	0	5,5,5	0.93	0
3	SO4	K	503	-	4,4,4	0.15	0	6,6,6	0.49	0
2	GOL	J	501	-	5,5,5	0.72	0	5,5,5	1.04	0
3	SO4	D	503	-	4,4,4	0.10	0	6,6,6	0.27	0
3	SO4	A	502	-	4,4,4	0.14	0	6,6,6	0.76	0
3	SO4	I	502	-	4,4,4	0.16	0	6,6,6	0.42	0
3	SO4	F	502	-	4,4,4	0.18	0	6,6,6	0.30	0
2	GOL	H	501	-	5,5,5	0.93	0	5,5,5	1.03	0
3	SO4	C	502	-	4,4,4	0.16	0	6,6,6	0.57	0
3	SO4	L	502	-	4,4,4	0.16	0	6,6,6	0.30	0
2	GOL	D	501	-	5,5,5	1.06	1 (20%)	5,5,5	0.69	0
2	GOL	E	501	-	5,5,5	0.85	0	5,5,5	0.98	0
3	SO4	B	503	-	4,4,4	0.17	0	6,6,6	0.33	0
3	SO4	K	502	-	4,4,4	0.12	0	6,6,6	0.21	0
3	SO4	L	503	-	4,4,4	0.13	0	6,6,6	0.26	0
2	GOL	B	501	-	5,5,5	0.79	0	5,5,5	1.06	0
3	SO4	J	503	-	4,4,4	0.20	0	6,6,6	0.34	0
3	SO4	B	502	-	4,4,4	0.18	0	6,6,6	0.35	0
3	SO4	I	503	-	4,4,4	0.11	0	6,6,6	0.29	0
3	SO4	C	503	-	4,4,4	0.22	0	6,6,6	0.13	0
2	GOL	A	501	-	5,5,5	1.55	1 (20%)	5,5,5	0.68	0
2	GOL	G	501	-	5,5,5	0.92	0	5,5,5	1.09	1 (20%)
3	SO4	H	502	-	4,4,4	0.16	0	6,6,6	0.47	0
3	SO4	A	503	-	4,4,4	0.23	0	6,6,6	0.23	0
3	SO4	J	502	-	4,4,4	0.18	0	6,6,6	0.16	0
3	SO4	D	502	-	4,4,4	0.17	0	6,6,6	0.29	0

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	GOL	E	501	-	-	1/4/4/4	-
2	GOL	J	501	-	-	2/4/4/4	-
2	GOL	A	501	-	-	2/4/4/4	-
2	GOL	G	501	-	-	0/4/4/4	-
2	GOL	K	501	-	-	4/4/4/4	-
2	GOL	L	501	-	-	2/4/4/4	-
2	GOL	H	501	-	-	0/4/4/4	-
2	GOL	F	501	-	-	1/4/4/4	-
2	GOL	B	501	-	-	2/4/4/4	-
2	GOL	I	501	-	-	0/4/4/4	-
2	GOL	D	501	-	-	0/4/4/4	-
2	GOL	C	501	-	-	2/4/4/4	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	GOL	O2-C2	-2.82	1.35	1.43
2	D	501	GOL	O2-C2	-2.03	1.37	1.43

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	K	501	GOL	C3-C2-C1	-2.26	102.92	111.70
2	G	501	GOL	C3-C2-C1	-2.06	103.71	111.70

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	501	GOL	C1-C2-C3-O3
2	J	501	GOL	C1-C2-C3-O3
2	K	501	GOL	C1-C2-C3-O3
2	K	501	GOL	O2-C2-C3-O3
2	L	501	GOL	O1-C1-C2-C3

There are no ring outliers.

11 monomers are involved in 18 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	501	GOL	2	0
3	G	503	SO4	1	0
3	H	503	SO4	1	0
2	I	501	GOL	1	0
2	J	501	GOL	2	0
3	L	502	SO4	1	0
2	D	501	GOL	4	0
3	L	503	SO4	1	0
2	B	501	GOL	1	0
3	H	502	SO4	3	0
3	J	502	SO4	1	0

## 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	453/454 (99%)	-0.42	1 (0%) 92 93	34, 45, 65, 92	0
1	B	453/454 (99%)	0.13	3 (0%) 84 86	46, 63, 89, 123	0
1	C	453/454 (99%)	-0.38	0 100 100	35, 45, 67, 105	0
1	D	453/454 (99%)	-0.26	0 100 100	41, 50, 74, 92	0
1	E	453/454 (99%)	-0.17	1 (0%) 92 93	43, 54, 76, 103	0
1	F	452/454 (99%)	-0.11	2 (0%) 89 90	40, 56, 80, 113	0
1	G	453/454 (99%)	-0.12	2 (0%) 89 90	44, 55, 80, 110	0
1	H	452/454 (99%)	-0.10	2 (0%) 89 90	43, 57, 79, 114	0
1	I	452/454 (99%)	-0.10	1 (0%) 92 93	46, 61, 88, 111	0
1	J	453/454 (99%)	0.17	6 (1%) 74 78	45, 65, 92, 123	0
1	K	452/454 (99%)	0.39	7 (1%) 71 75	46, 68, 90, 116	0
1	L	453/454 (99%)	0.31	6 (1%) 74 78	50, 75, 98, 119	0
All	All	5432/5448 (99%)	-0.06	31 (0%) 85 88	34, 57, 86, 123	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	324	HIS	3.3
1	B	1	MET	3.3
1	L	202	VAL	3.1
1	I	282	LEU	3.0
1	J	282	LEU	2.9

### 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 ⓘ

There are no oligosaccharides in this entry.

## 6.4 Ligands ⓘ

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	GOL	F	501	6/6	0.80	0.15	63,65,71,76	0
2	GOL	H	501	6/6	0.88	0.12	53,54,64,79	0
3	SO4	B	502	5/5	0.89	0.10	67,75,84,84	0
3	SO4	K	503	5/5	0.89	0.10	73,78,82,83	0
2	GOL	B	501	6/6	0.90	0.12	69,79,81,84	0
2	GOL	D	501	6/6	0.90	0.14	58,69,69,70	0
3	SO4	L	503	5/5	0.90	0.08	85,87,90,94	0
3	SO4	I	503	5/5	0.91	0.12	76,83,85,87	0
3	SO4	J	503	5/5	0.91	0.09	70,71,79,88	0
2	GOL	I	501	6/6	0.91	0.10	62,65,67,68	0
3	SO4	F	502	5/5	0.91	0.08	68,72,81,96	0
2	GOL	G	501	6/6	0.92	0.11	56,63,66,69	0
2	GOL	K	501	6/6	0.92	0.09	67,74,81,82	0
2	GOL	L	501	6/6	0.92	0.11	75,79,85,88	0
2	GOL	E	501	6/6	0.92	0.10	60,65,68,68	0
3	SO4	D	502	5/5	0.92	0.08	59,63,72,82	0
2	GOL	C	501	6/6	0.93	0.09	48,51,55,61	0
3	SO4	J	502	5/5	0.93	0.07	74,78,88,93	0
2	GOL	A	501	6/6	0.93	0.11	47,57,60,71	0
3	SO4	F	503	5/5	0.93	0.09	61,70,75,79	0
3	SO4	L	502	5/5	0.93	0.09	65,65,80,86	0
3	SO4	I	502	5/5	0.93	0.06	75,75,87,89	0
3	SO4	G	502	5/5	0.94	0.07	58,66,70,71	0
3	SO4	G	503	5/5	0.94	0.07	57,65,69,71	0
2	GOL	J	501	6/6	0.94	0.09	71,74,79,81	0
3	SO4	C	502	5/5	0.95	0.07	60,60,65,75	0
3	SO4	B	503	5/5	0.95	0.07	58,64,67,70	0
3	SO4	K	502	5/5	0.95	0.08	58,63,73,83	0
3	SO4	A	502	5/5	0.96	0.07	58,61,64,76	0
3	SO4	E	503	5/5	0.96	0.06	68,70,75,76	0
3	SO4	H	502	5/5	0.96	0.07	48,59,74,74	0
3	SO4	H	503	5/5	0.96	0.06	62,69,72,73	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	SO4	C	503	5/5	0.97	0.07	52,54,59,62	0
3	SO4	D	503	5/5	0.97	0.06	52,54,60,63	0
3	SO4	E	502	5/5	0.97	0.04	60,64,70,74	0
3	SO4	A	503	5/5	0.98	0.04	42,49,54,57	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.