



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 5, 2025 – 08:12 PM JST

PDB ID : 8J5L  
Title : Structure of GH1 Br2 beta-glucosidase E163Q mutant from bovine rumen metagenome  
Authors : Kaenying, W.; Kongsaree, P.T.; Tagami, T.  
Deposited on : 2023-04-23  
Resolution : 2.10 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](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 : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.21  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.004 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.41.2

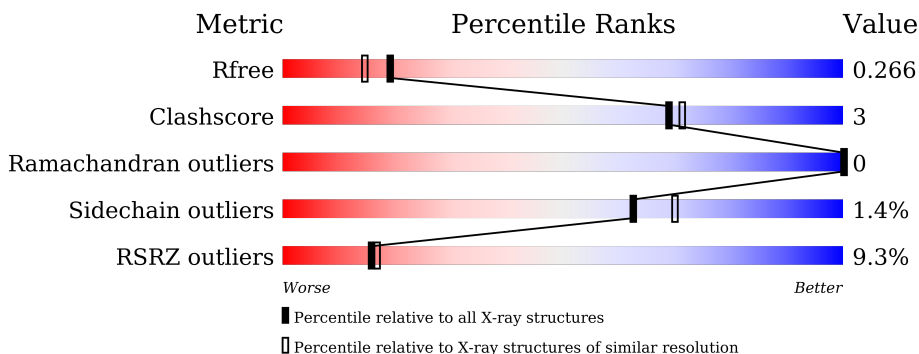
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	6234 (2.10-2.10)
Clashscore	180529	6893 (2.10-2.10)
Ramachandran outliers	177936	6839 (2.10-2.10)
Sidechain outliers	177891	6840 (2.10-2.10)
RSRZ outliers	164620	6234 (2.10-2.10)

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	465	<div> <div>2%</div> <div> <div></div> <div>89%</div> <div>8%</div> <div>•</div> </div> </div>
1	B	465	<div> <div>2%</div> <div> <div></div> <div>87%</div> <div>10%</div> <div>•</div> </div> </div>
1	C	465	<div> <div>19%</div> <div> <div></div> <div>87%</div> <div>8%</div> <div>5%</div> </div> </div>
1	D	465	<div> <div>13%</div> <div> <div></div> <div>88%</div> <div>7%</div> <div>5%</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 14876 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	452	Total	C	N	O	S	0	0	0
			3642	2349	609	668	16			
1	B	452	Total	C	N	O	S	0	0	0
			3642	2349	609	668	16			
1	C	443	Total	C	N	O	S	0	0	0
			3581	2312	596	658	15			
1	D	443	Total	C	N	O	S	0	0	0
			3574	2306	595	658	15			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP A0A1S5SJM8
A	-18	GLY	-	expression tag	UNP A0A1S5SJM8
A	-17	SER	-	expression tag	UNP A0A1S5SJM8
A	-16	SER	-	expression tag	UNP A0A1S5SJM8
A	-15	HIS	-	expression tag	UNP A0A1S5SJM8
A	-14	HIS	-	expression tag	UNP A0A1S5SJM8
A	-13	HIS	-	expression tag	UNP A0A1S5SJM8
A	-12	HIS	-	expression tag	UNP A0A1S5SJM8
A	-11	HIS	-	expression tag	UNP A0A1S5SJM8
A	-10	HIS	-	expression tag	UNP A0A1S5SJM8
A	-9	SER	-	expression tag	UNP A0A1S5SJM8
A	-8	SER	-	expression tag	UNP A0A1S5SJM8
A	-7	GLY	-	expression tag	UNP A0A1S5SJM8
A	-6	LEU	-	expression tag	UNP A0A1S5SJM8
A	-5	VAL	-	expression tag	UNP A0A1S5SJM8
A	-4	PRO	-	expression tag	UNP A0A1S5SJM8
A	-3	ARG	-	expression tag	UNP A0A1S5SJM8
A	-2	GLY	-	expression tag	UNP A0A1S5SJM8
A	-1	SER	-	expression tag	UNP A0A1S5SJM8
A	0	HIS	-	expression tag	UNP A0A1S5SJM8
A	163	GLN	GLU	engineered mutation	UNP A0A1S5SJM8

*Continued on next page...*

*Continued from previous page...*

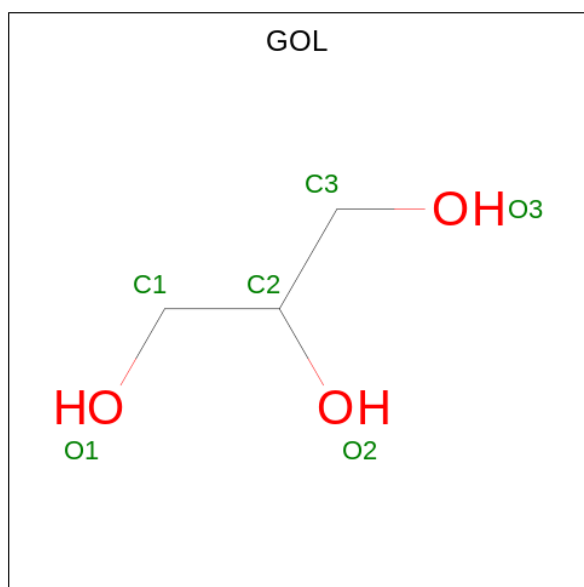
Chain	Residue	Modelled	Actual	Comment	Reference
B	-19	MET	-	initiating methionine	UNP A0A1S5SJM8
B	-18	GLY	-	expression tag	UNP A0A1S5SJM8
B	-17	SER	-	expression tag	UNP A0A1S5SJM8
B	-16	SER	-	expression tag	UNP A0A1S5SJM8
B	-15	HIS	-	expression tag	UNP A0A1S5SJM8
B	-14	HIS	-	expression tag	UNP A0A1S5SJM8
B	-13	HIS	-	expression tag	UNP A0A1S5SJM8
B	-12	HIS	-	expression tag	UNP A0A1S5SJM8
B	-11	HIS	-	expression tag	UNP A0A1S5SJM8
B	-10	HIS	-	expression tag	UNP A0A1S5SJM8
B	-9	SER	-	expression tag	UNP A0A1S5SJM8
B	-8	SER	-	expression tag	UNP A0A1S5SJM8
B	-7	GLY	-	expression tag	UNP A0A1S5SJM8
B	-6	LEU	-	expression tag	UNP A0A1S5SJM8
B	-5	VAL	-	expression tag	UNP A0A1S5SJM8
B	-4	PRO	-	expression tag	UNP A0A1S5SJM8
B	-3	ARG	-	expression tag	UNP A0A1S5SJM8
B	-2	GLY	-	expression tag	UNP A0A1S5SJM8
B	-1	SER	-	expression tag	UNP A0A1S5SJM8
B	0	HIS	-	expression tag	UNP A0A1S5SJM8
B	163	GLN	GLU	engineered mutation	UNP A0A1S5SJM8
C	-19	MET	-	initiating methionine	UNP A0A1S5SJM8
C	-18	GLY	-	expression tag	UNP A0A1S5SJM8
C	-17	SER	-	expression tag	UNP A0A1S5SJM8
C	-16	SER	-	expression tag	UNP A0A1S5SJM8
C	-15	HIS	-	expression tag	UNP A0A1S5SJM8
C	-14	HIS	-	expression tag	UNP A0A1S5SJM8
C	-13	HIS	-	expression tag	UNP A0A1S5SJM8
C	-12	HIS	-	expression tag	UNP A0A1S5SJM8
C	-11	HIS	-	expression tag	UNP A0A1S5SJM8
C	-10	HIS	-	expression tag	UNP A0A1S5SJM8
C	-9	SER	-	expression tag	UNP A0A1S5SJM8
C	-8	SER	-	expression tag	UNP A0A1S5SJM8
C	-7	GLY	-	expression tag	UNP A0A1S5SJM8
C	-6	LEU	-	expression tag	UNP A0A1S5SJM8
C	-5	VAL	-	expression tag	UNP A0A1S5SJM8
C	-4	PRO	-	expression tag	UNP A0A1S5SJM8
C	-3	ARG	-	expression tag	UNP A0A1S5SJM8
C	-2	GLY	-	expression tag	UNP A0A1S5SJM8
C	-1	SER	-	expression tag	UNP A0A1S5SJM8
C	0	HIS	-	expression tag	UNP A0A1S5SJM8
C	163	GLN	GLU	engineered mutation	UNP A0A1S5SJM8

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-19	MET	-	initiating methionine	UNP A0A1S5SJM8
D	-18	GLY	-	expression tag	UNP A0A1S5SJM8
D	-17	SER	-	expression tag	UNP A0A1S5SJM8
D	-16	SER	-	expression tag	UNP A0A1S5SJM8
D	-15	HIS	-	expression tag	UNP A0A1S5SJM8
D	-14	HIS	-	expression tag	UNP A0A1S5SJM8
D	-13	HIS	-	expression tag	UNP A0A1S5SJM8
D	-12	HIS	-	expression tag	UNP A0A1S5SJM8
D	-11	HIS	-	expression tag	UNP A0A1S5SJM8
D	-10	HIS	-	expression tag	UNP A0A1S5SJM8
D	-9	SER	-	expression tag	UNP A0A1S5SJM8
D	-8	SER	-	expression tag	UNP A0A1S5SJM8
D	-7	GLY	-	expression tag	UNP A0A1S5SJM8
D	-6	LEU	-	expression tag	UNP A0A1S5SJM8
D	-5	VAL	-	expression tag	UNP A0A1S5SJM8
D	-4	PRO	-	expression tag	UNP A0A1S5SJM8
D	-3	ARG	-	expression tag	UNP A0A1S5SJM8
D	-2	GLY	-	expression tag	UNP A0A1S5SJM8
D	-1	SER	-	expression tag	UNP A0A1S5SJM8
D	0	HIS	-	expression tag	UNP A0A1S5SJM8
D	163	GLN	GLU	engineered mutation	UNP A0A1S5SJM8

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



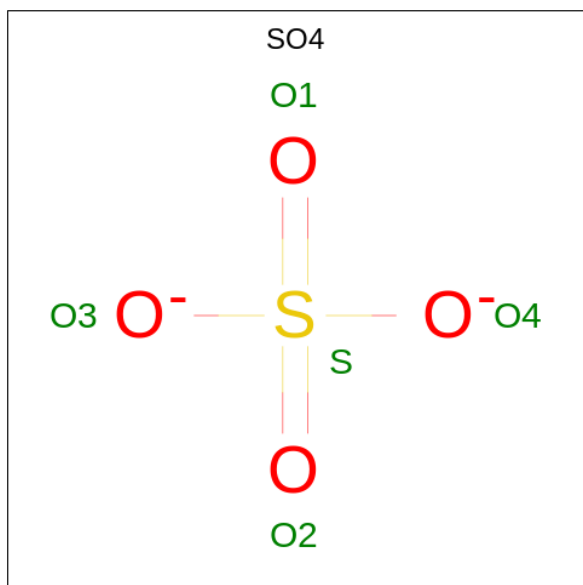
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	3	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
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		

- Molecule 3 is SULFATE ION (three-letter code: 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	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	B	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

*Continued from previous page...*

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

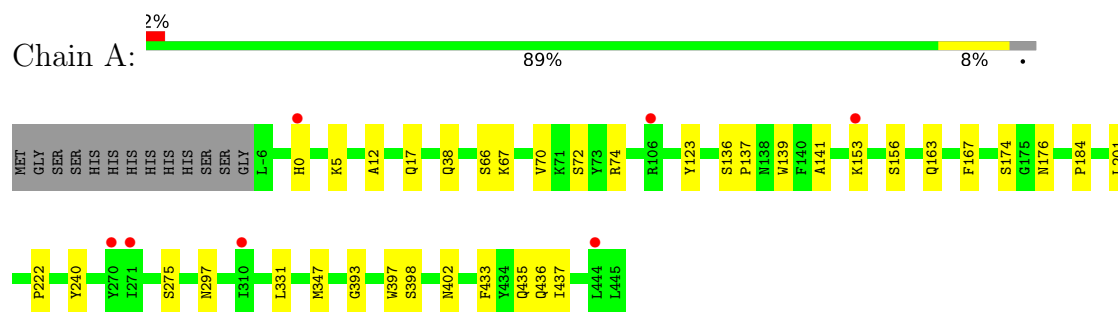
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	123	Total	O	0	0
			123	123		
4	B	116	Total	O	0	0
			116	116		
4	C	46	Total	O	0	0
			46	46		
4	D	53	Total	O	0	0
			53	53		

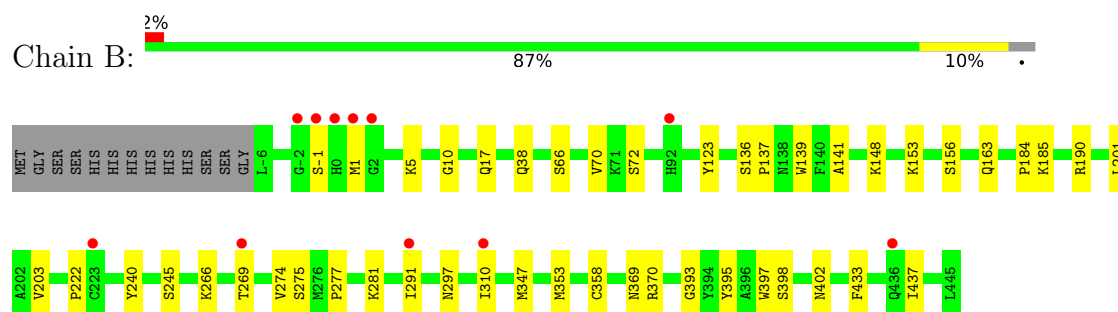
### 3 Residue-property plots [i](#)

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.

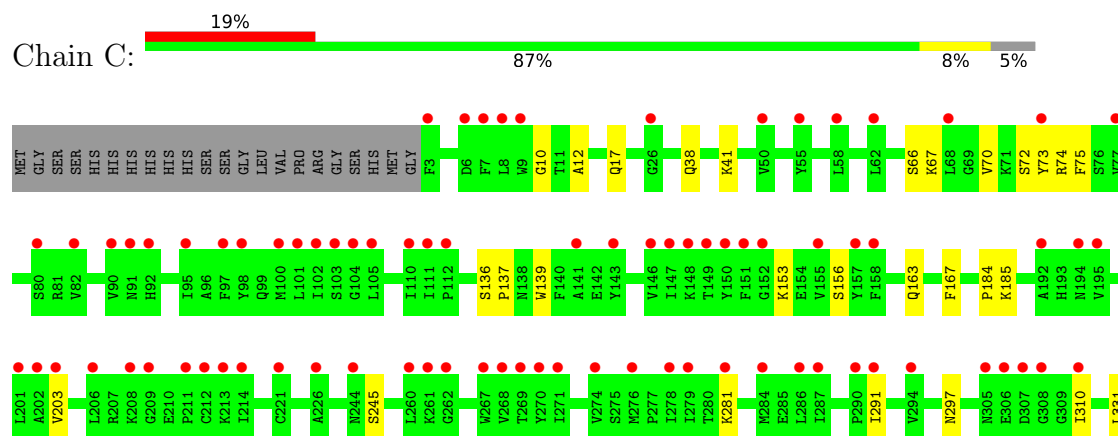
#### • Molecule 1: Beta-glucosidase



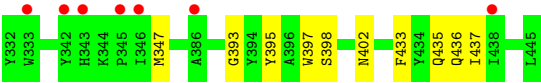
#### • Molecule 1: Beta-glucosidase



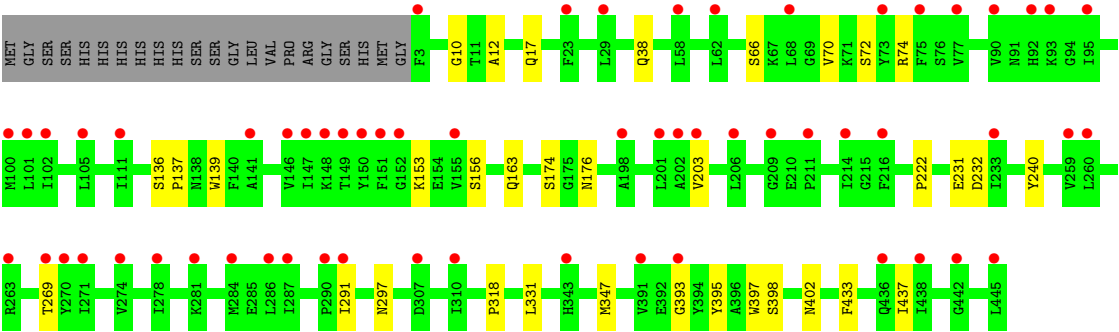
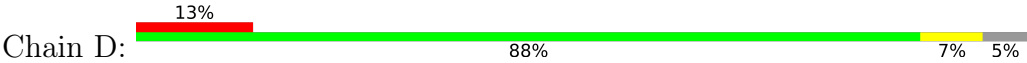
#### • Molecule 1: Beta-glucosidase







● Molecule 1: Beta-glucosidase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	105.49Å 113.42Å 180.49Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.14 – 2.10 48.14 – 2.10	Depositor EDS
% Data completeness (in resolution range)	90.3 (48.14-2.10) 90.3 (48.14-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.37 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.227 , 0.261 0.231 , 0.266	Depositor DCC
$R_{free}$ test set	113115 reflections (1.73%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	39.6	Xtriage
Anisotropy	0.127	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 34.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	14876	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.17% 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	0.62	0/3750	0.72	0/5089
1	B	0.63	0/3750	0.73	0/5089
1	C	0.62	0/3687	0.72	0/5003
1	D	0.62	0/3680	0.71	0/4995
All	All	0.62	0/14867	0.72	0/20176

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	3642	0	3487	22	0
1	B	3642	0	3487	34	0
1	C	3581	0	3428	22	0
1	D	3574	0	3408	23	0
2	A	6	0	8	0	0
2	B	6	0	8	0	0
2	C	6	0	8	0	0
2	D	6	0	8	0	0
3	A	25	0	0	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	20	0	0	1	0
3	C	10	0	0	0	0
3	D	20	0	0	0	0
4	A	123	0	0	1	0
4	B	116	0	0	1	0
4	C	46	0	0	1	0
4	D	53	0	0	0	0
All	All	14876	0	13842	95	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 (95) 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:353:MET:CE	1:B:370:ARG:HA	1.56	1.33
1:B:353:MET:HE1	1:B:370:ARG:HA	1.30	1.06
1:B:353:MET:HE3	1:B:370:ARG:HA	1.44	0.99
1:B:353:MET:CE	1:B:370:ARG:CA	2.48	0.88
1:B:353:MET:HE1	1:B:370:ARG:CA	2.07	0.82
1:B:353:MET:HE3	1:B:370:ARG:CA	2.17	0.71
1:B:433:PHE:CE2	1:B:437:ILE:HD11	2.28	0.69
1:A:433:PHE:CE2	1:A:437:ILE:HD11	2.29	0.68
1:B:185:LYS:HD2	1:B:245:SER:O	1.93	0.68
1:C:433:PHE:CE2	1:C:437:ILE:HD11	2.30	0.67
1:D:433:PHE:CE2	1:D:437:ILE:HD11	2.30	0.67
1:C:185:LYS:HD2	1:C:245:SER:O	1.98	0.63
1:B:148:LYS:HE3	4:B:612:HOH:O	1.99	0.60
1:B:353:MET:HE1	1:B:369:ASN:O	2.02	0.60
1:D:231:GLU:HG2	1:D:232:ASP:H	1.67	0.60
1:D:231:GLU:HG2	1:D:232:ASP:N	2.21	0.56
1:B:266:LYS:O	1:B:269:THR:HG22	2.08	0.54
1:A:184:PRO:HB3	1:B:38:GLN:HG2	1.90	0.54
1:A:275:SER:HB2	1:B:123:TYR:CE2	2.42	0.53
1:A:123:TYR:CE2	1:B:275:SER:HB2	2.44	0.52
1:B:347:MET:CE	1:B:393:GLY:HA3	2.40	0.52
1:A:163:GLN:HE21	1:A:297:ASN:HB2	1.76	0.51
1:A:347:MET:CE	1:A:393:GLY:HA3	2.41	0.50
1:A:433:PHE:CZ	1:A:437:ILE:HD11	2.47	0.50
1:C:433:PHE:CZ	1:C:437:ILE:HD11	2.47	0.49
1:D:17:GLN:O	1:D:402:ASN:HB2	2.13	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:17:GLN:O	1:B:402:ASN:HB2	2.12	0.49
1:D:163:GLN:HE21	1:D:297:ASN:HB2	1.78	0.49
1:D:203:VAL:HG21	1:D:291:ILE:HD11	1.95	0.49
1:B:163:GLN:HE21	1:B:297:ASN:HB2	1.77	0.49
1:C:17:GLN:O	1:C:402:ASN:HB2	2.13	0.49
1:A:17:GLN:O	1:A:402:ASN:HB2	2.13	0.48
1:B:433:PHE:CZ	1:B:437:ILE:HD11	2.48	0.48
1:C:184:PRO:HB3	1:D:38:GLN:HG2	1.94	0.48
1:D:433:PHE:CZ	1:D:437:ILE:HD11	2.48	0.48
1:D:203:VAL:CG2	1:D:291:ILE:HD11	2.44	0.48
1:C:163:GLN:HE21	1:C:297:ASN:HB2	1.78	0.48
1:C:347:MET:CE	1:C:393:GLY:HA3	2.44	0.47
1:C:203:VAL:CG2	1:C:291:ILE:HD11	2.45	0.47
1:D:347:MET:CE	1:D:393:GLY:HA3	2.44	0.47
1:B:203:VAL:CG2	1:B:291:ILE:HD11	2.45	0.47
1:B:203:VAL:HG21	1:B:291:ILE:HD11	1.96	0.46
1:B:353:MET:HE1	1:B:369:ASN:C	2.36	0.46
1:C:203:VAL:HG21	1:C:291:ILE:HD11	1.98	0.46
1:B:136:SER:HA	1:B:139:TRP:CE3	2.51	0.46
1:B:163:GLN:HE21	1:B:297:ASN:CB	2.29	0.46
1:A:38:GLN:HG2	1:B:184:PRO:HB3	1.97	0.45
1:A:136:SER:HA	1:A:139:TRP:CE3	2.51	0.45
1:D:136:SER:HA	1:D:139:TRP:CE3	2.52	0.45
1:C:136:SER:HA	1:C:139:TRP:CE3	2.52	0.45
1:C:310:ILE:HD12	1:C:310:ILE:O	2.17	0.44
1:B:310:ILE:HD12	1:B:310:ILE:O	2.17	0.44
1:A:163:GLN:HE21	1:A:297:ASN:CB	2.31	0.44
1:C:167:PHE:HA	4:C:620:HOH:O	2.17	0.44
1:D:231:GLU:HG3	1:D:232:ASP:OD1	2.18	0.44
1:C:163:GLN:HE21	1:C:297:ASN:CB	2.31	0.43
1:B:397:TRP:HA	1:B:398:SER:HA	1.78	0.43
1:A:331:LEU:HD12	1:A:331:LEU:HA	1.93	0.43
1:C:70:VAL:CG1	1:C:72:SER:O	2.67	0.43
1:B:136:SER:N	1:B:137:PRO:CD	2.82	0.42
1:D:163:GLN:HE21	1:D:297:ASN:CB	2.31	0.42
1:D:331:LEU:HD12	1:D:331:LEU:HA	1.92	0.42
1:A:70:VAL:CG1	1:A:72:SER:O	2.66	0.42
1:A:136:SER:N	1:A:137:PRO:CD	2.82	0.42
1:B:70:VAL:CG1	1:B:72:SER:O	2.68	0.42
1:C:136:SER:N	1:C:137:PRO:CD	2.82	0.42
1:D:231:GLU:CG	1:D:232:ASP:N	2.83	0.42

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:38:GLN:HB3	1:C:41:LYS:HG2	2.01	0.42
1:B:358:CYS:HB3	1:D:318:PRO:HB3	2.01	0.42
1:B:274:VAL:HG23	3:B:503:SO4:O3	2.20	0.42
1:D:174:SER:OG	1:D:176:ASN:OD1	2.37	0.42
1:A:141:ALA:HB2	1:A:201:LEU:HB3	2.01	0.42
1:B:190:ARG:HA	1:B:277:PRO:HG3	2.02	0.41
1:D:136:SER:N	1:D:137:PRO:CD	2.83	0.41
1:A:397:TRP:HA	1:A:398:SER:HA	1.79	0.41
1:D:70:VAL:CG1	1:D:72:SER:O	2.68	0.41
1:D:397:TRP:HA	1:D:398:SER:HA	1.79	0.41
1:A:167:PHE:HA	4:A:699:HOH:O	2.19	0.41
1:A:222:PRO:HA	1:A:240:TYR:CE1	2.56	0.41
1:A:174:SER:OG	1:A:176:ASN:OD1	2.36	0.41
1:B:10:GLY:HA3	1:B:395:TYR:CD2	2.56	0.41
1:B:141:ALA:HB2	1:B:201:LEU:HB3	2.02	0.41
1:D:222:PRO:HA	1:D:240:TYR:CE1	2.56	0.41
1:C:10:GLY:HA3	1:C:395:TYR:CD2	2.56	0.40
1:A:12:ALA:HA	1:A:74:ARG:O	2.21	0.40
1:B:222:PRO:HA	1:B:240:TYR:CE1	2.57	0.40
1:C:67:LYS:O	1:C:435:GLN:NE2	2.54	0.40
1:C:73:TYR:CE2	1:C:75:PHE:HB3	2.56	0.40
1:D:12:ALA:HA	1:D:74:ARG:O	2.22	0.40
1:A:0:HIS:NE2	3:A:506:SO4:O2	2.53	0.40
1:A:67:LYS:O	1:A:435:GLN:NE2	2.54	0.40
1:C:397:TRP:HA	1:C:398:SER:HA	1.78	0.40
1:C:12:ALA:HA	1:C:74:ARG:O	2.21	0.40
1:C:331:LEU:HD12	1:C:331:LEU:HA	1.92	0.40
1:D:10:GLY:HA3	1:D:395:TYR:CD2	2.57	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	450/465 (97%)	433 (96%)	17 (4%)	0	100	100
1	B	450/465 (97%)	432 (96%)	18 (4%)	0	100	100
1	C	441/465 (95%)	424 (96%)	17 (4%)	0	100	100
1	D	441/465 (95%)	423 (96%)	18 (4%)	0	100	100
All	All	1782/1860 (96%)	1712 (96%)	70 (4%)	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	383/396 (97%)	378 (99%)	5 (1%)	65	72
1	B	383/396 (97%)	376 (98%)	7 (2%)	54	61
1	C	377/396 (95%)	372 (99%)	5 (1%)	65	72
1	D	375/396 (95%)	371 (99%)	4 (1%)	70	77
All	All	1518/1584 (96%)	1497 (99%)	21 (1%)	62	70

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	5	LYS
1	A	66	SER
1	A	153	LYS
1	A	156	SER
1	A	436	GLN
1	B	-1	SER
1	B	1	MET
1	B	5	LYS
1	B	66	SER
1	B	153	LYS
1	B	156	SER
1	B	281	LYS
1	C	66	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	153	LYS
1	C	156	SER
1	C	281	LYS
1	C	436	GLN
1	D	66	SER
1	D	153	LYS
1	D	156	SER
1	D	269	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	163	GLN
1	B	163	GLN
1	B	380	GLN
1	C	163	GLN
1	C	369	ASN
1	C	380	GLN
1	D	163	GLN
1	D	435	GLN
1	D	436	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](#)

19 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	GOL	C	501	-	5,5,5	0.10	0	5,5,5	0.28	0
2	GOL	A	501	-	5,5,5	0.07	0	5,5,5	0.14	0
3	SO4	D	505	-	4,4,4	0.37	0	6,6,6	0.06	0
3	SO4	D	502	-	4,4,4	0.38	0	6,6,6	0.08	0
3	SO4	B	504	-	4,4,4	0.34	0	6,6,6	0.06	0
2	GOL	B	501	-	5,5,5	0.06	0	5,5,5	0.26	0
3	SO4	C	503	-	4,4,4	0.36	0	6,6,6	0.04	0
3	SO4	B	505	-	4,4,4	0.35	0	6,6,6	0.04	0
3	SO4	B	503	-	4,4,4	0.39	0	6,6,6	0.08	0
3	SO4	B	502	-	4,4,4	0.36	0	6,6,6	0.06	0
2	GOL	D	501	-	5,5,5	0.12	0	5,5,5	0.29	0
3	SO4	A	505	-	4,4,4	0.37	0	6,6,6	0.07	0
3	SO4	A	506	-	4,4,4	0.32	0	6,6,6	0.05	0
3	SO4	D	503	-	4,4,4	0.37	0	6,6,6	0.10	0
3	SO4	C	502	-	4,4,4	0.37	0	6,6,6	0.12	0
3	SO4	D	504	-	4,4,4	0.36	0	6,6,6	0.05	0
3	SO4	A	504	-	4,4,4	0.34	0	6,6,6	0.11	0
3	SO4	A	503	-	4,4,4	0.37	0	6,6,6	0.11	0
3	SO4	A	502	-	4,4,4	0.35	0	6,6,6	0.07	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	D	501	-	-	2/4/4/4	-
2	GOL	B	501	-	-	0/4/4/4	-
2	GOL	C	501	-	-	2/4/4/4	-
2	GOL	A	501	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	501	GOL	C1-C2-C3-O3
2	D	501	GOL	C1-C2-C3-O3
2	C	501	GOL	O2-C2-C3-O3
2	D	501	GOL	O2-C2-C3-O3

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	503	SO4	1	0
3	A	506	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

### 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, 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	452/465 (97%)	0.09	7 (1%) 71 73	24, 37, 56, 80	0
1	B	452/465 (97%)	0.36	11 (2%) 59 61	25, 41, 62, 92	0
1	C	443/465 (95%)	1.24	89 (20%) 3 4	31, 57, 80, 104	0
1	D	443/465 (95%)	1.06	60 (13%) 8 9	32, 55, 75, 96	0
All	All	1790/1860 (96%)	0.68	167 (9%) 16 17	24, 46, 74, 104	0

All (167) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	201	LEU	4.8
1	B	2	GLY	4.4
1	C	308	GLY	4.2
1	C	279	ILE	3.9
1	C	101	LEU	3.8
1	C	214	ILE	3.8
1	C	3	PHE	3.7
1	C	111	ILE	3.6
1	B	0	HIS	3.6
1	C	278	ILE	3.6
1	C	211	PRO	3.5
1	C	270	TYR	3.4
1	A	270	TYR	3.4
1	C	310	ILE	3.4
1	D	310	ILE	3.3
1	C	284	MET	3.3
1	C	269	THR	3.3
1	C	202	ALA	3.3
1	C	9	TRP	3.3
1	D	152	GLY	3.3
1	D	202	ALA	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	152	GLY	3.2
1	C	291	ILE	3.2
1	B	310	ILE	3.1
1	D	105	LEU	3.1
1	D	260	LEU	3.1
1	B	1	MET	3.1
1	C	260	LEU	3.1
1	B	291	ILE	3.1
1	C	92	HIS	3.0
1	D	201	LEU	3.0
1	C	271	ILE	3.0
1	C	143	TYR	3.0
1	C	148	LYS	3.0
1	D	93	LYS	3.0
1	C	80	SER	3.0
1	D	269	THR	3.0
1	C	386	ALA	3.0
1	D	198	ALA	3.0
1	D	291	ILE	2.9
1	B	269	THR	2.9
1	D	281	LYS	2.9
1	C	73	TYR	2.9
1	C	294	VAL	2.9
1	C	77	VAL	2.8
1	C	206	LEU	2.8
1	C	286	LEU	2.8
1	D	58	LEU	2.8
1	D	278	ILE	2.8
1	C	147	ILE	2.8
1	D	95	ILE	2.8
1	C	98	TYR	2.8
1	D	90	VAL	2.8
1	C	141	ALA	2.7
1	C	346	ILE	2.7
1	C	100	MET	2.7
1	C	438	ILE	2.7
1	D	111	ILE	2.7
1	D	147	ILE	2.7
1	C	55	TYR	2.7
1	C	267	TRP	2.7
1	D	290	PRO	2.6
1	B	-1	SER	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	442	GLY	2.6
1	C	68	LEU	2.6
1	C	343	HIS	2.6
1	C	151	PHE	2.6
1	C	95	ILE	2.6
1	D	141	ALA	2.6
1	C	203	VAL	2.6
1	D	274	VAL	2.6
1	A	153	LYS	2.5
1	C	192	ALA	2.5
1	C	155	VAL	2.5
1	D	436	GLN	2.5
1	A	0	HIS	2.5
1	A	444	LEU	2.5
1	B	-2	GLY	2.5
1	C	26	GLY	2.5
1	C	281	LYS	2.5
1	C	307	ASP	2.5
1	D	146	VAL	2.5
1	C	150	TYR	2.5
1	D	149	THR	2.5
1	C	287	ILE	2.5
1	D	101	LEU	2.5
1	D	206	LEU	2.5
1	C	104	GLY	2.5
1	D	286	LEU	2.5
1	C	146	VAL	2.5
1	D	214	ILE	2.4
1	C	276	MET	2.4
1	D	68	LEU	2.4
1	D	263	ARG	2.4
1	C	50	VAL	2.4
1	D	259	VAL	2.4
1	D	211	PRO	2.4
1	D	3	PHE	2.4
1	C	8	LEU	2.4
1	D	271	ILE	2.4
1	D	287	ILE	2.4
1	C	82	VAL	2.4
1	C	90	VAL	2.4
1	C	62	LEU	2.3
1	D	29	LEU	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	77	VAL	2.3
1	D	155	VAL	2.3
1	D	391	VAL	2.3
1	D	23	PHE	2.3
1	C	157	TYR	2.3
1	C	105	LEU	2.3
1	D	100	MET	2.3
1	C	208	LYS	2.3
1	C	213	LYS	2.3
1	C	58	LEU	2.3
1	C	212	CYS	2.3
1	C	112	PRO	2.3
1	C	262	GLY	2.3
1	D	209	GLY	2.3
1	D	150	TYR	2.3
1	D	270	TYR	2.3
1	D	203	VAL	2.2
1	D	393	GLY	2.2
1	B	92	HIS	2.2
1	C	110	ILE	2.2
1	C	290	PRO	2.2
1	C	209	GLY	2.2
1	C	274	VAL	2.2
1	D	92	HIS	2.2
1	C	7	PHE	2.2
1	D	216	PHE	2.2
1	A	271	ILE	2.2
1	C	306	GLU	2.2
1	C	333	TRP	2.2
1	D	148	LYS	2.2
1	D	307	ASP	2.2
1	C	194	ASN	2.2
1	C	195	VAL	2.2
1	C	158	PHE	2.1
1	C	6	ASP	2.1
1	C	221	CYS	2.1
1	C	97	PHE	2.1
1	A	310	ILE	2.1
1	C	261	LYS	2.1
1	D	62	LEU	2.1
1	D	151	PHE	2.1
1	D	233	ILE	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	244	ASN	2.1
1	C	102	ILE	2.1
1	C	103	SER	2.1
1	C	149	THR	2.1
1	D	343	HIS	2.1
1	C	345	PRO	2.1
1	B	436	GLN	2.1
1	C	91	ASN	2.1
1	C	305	ASN	2.1
1	B	223	CYS	2.1
1	C	268	VAL	2.1
1	D	75	PHE	2.1
1	C	226	ALA	2.0
1	A	106	ARG	2.0
1	D	445	LEU	2.0
1	D	102	ILE	2.0
1	D	438	ILE	2.0
1	C	342	TYR	2.0
1	D	73	TYR	2.0
1	D	284	MET	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 monosaccharides 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
3	SO4	B	505	5/5	0.68	0.10	92,98,103,106	0
3	SO4	D	505	5/5	0.71	0.11	91,92,95,104	0
3	SO4	C	502	5/5	0.75	0.11	73,73,78,80	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	B	504	5/5	0.78	0.10	69,78,80,86	0
3	SO4	B	503	5/5	0.80	0.10	73,77,79,81	0
3	SO4	A	506	5/5	0.81	0.10	71,74,78,81	0
3	SO4	A	505	5/5	0.81	0.09	79,81,85,88	0
3	SO4	D	503	5/5	0.83	0.09	74,79,82,87	0
3	SO4	A	503	5/5	0.84	0.09	64,73,74,80	0
3	SO4	A	504	5/5	0.87	0.08	63,71,77,80	0
3	SO4	C	503	5/5	0.88	0.08	83,88,92,94	0
3	SO4	D	502	5/5	0.89	0.10	74,77,77,86	0
3	SO4	B	502	5/5	0.89	0.08	74,74,75,75	0
2	GOL	B	501	6/6	0.89	0.12	35,38,40,48	0
3	SO4	D	504	5/5	0.90	0.08	66,76,81,88	0
3	SO4	A	502	5/5	0.90	0.08	55,65,69,72	0
2	GOL	C	501	6/6	0.93	0.09	44,46,51,52	0
2	GOL	A	501	6/6	0.94	0.10	33,34,35,38	0
2	GOL	D	501	6/6	0.95	0.10	43,47,49,49	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.