



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 23, 2023 – 01:31 PM EDT

PDB ID : 3E1F  
Title : E.Coli (lacZ) beta-galactosidase (H418E) in complex with galactose  
Authors : Huber, R.E.; Dugdale, M.L.  
Deposited on : 2008-08-04  
Resolution : 3.00 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 : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35

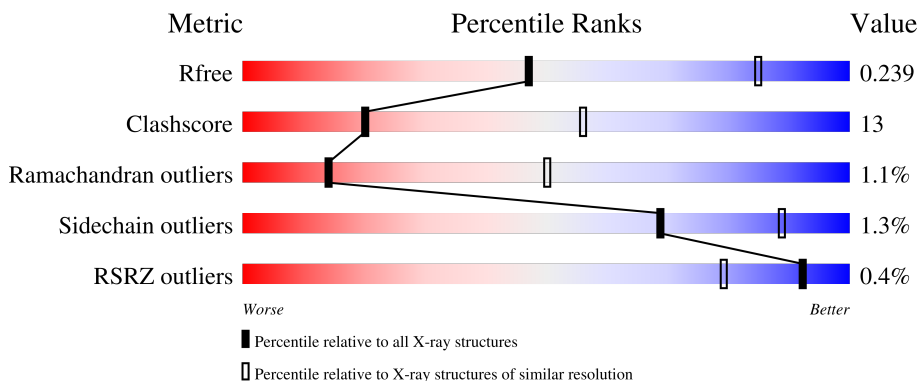
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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))
$R_{free}$	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	1	1023	
1	2	1023	
1	3	1023	
1	4	1023	

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
4	DMS	1	8416	-	-	-	X
4	DMS	1	8425	-	-	-	X
4	DMS	1	8503	-	-	-	X
4	DMS	2	8001	-	-	X	-
4	DMS	3	8414	-	-	-	X
4	DMS	3	8415	-	-	-	X
4	DMS	3	8416	-	-	-	X
4	DMS	3	8421	-	-	-	X

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 33235 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-galactosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	1	1011	Total	C	N	O	S	0	0	0
			8124	5137	1438	1511	38			
1	2	1011	Total	C	N	O	S	0	0	0
			8124	5137	1438	1511	38			
1	3	1011	Total	C	N	O	S	0	0	0
			8124	5137	1438	1511	38			
1	4	1011	Total	C	N	O	S	0	0	0
			8124	5137	1438	1511	38			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
1	1	GLY	-	expression tag	UNP P00722
1	2	SER	-	expression tag	UNP P00722
1	3	HIS	-	expression tag	UNP P00722
1	4	MET	-	expression tag	UNP P00722
1	5	LEU	-	expression tag	UNP P00722
1	6	GLU	-	expression tag	UNP P00722
1	7	ASP	-	expression tag	UNP P00722
1	8	PRO	-	expression tag	UNP P00722
1	418	GLU	HIS	engineered mutation	UNP P00722
2	1	GLY	-	expression tag	UNP P00722
2	2	SER	-	expression tag	UNP P00722
2	3	HIS	-	expression tag	UNP P00722
2	4	MET	-	expression tag	UNP P00722
2	5	LEU	-	expression tag	UNP P00722
2	6	GLU	-	expression tag	UNP P00722
2	7	ASP	-	expression tag	UNP P00722
2	8	PRO	-	expression tag	UNP P00722
2	418	GLU	HIS	engineered mutation	UNP P00722
3	1	GLY	-	expression tag	UNP P00722
3	2	SER	-	expression tag	UNP P00722
3	3	HIS	-	expression tag	UNP P00722

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
3	4	MET	-	expression tag	UNP P00722
3	5	LEU	-	expression tag	UNP P00722
3	6	GLU	-	expression tag	UNP P00722
3	7	ASP	-	expression tag	UNP P00722
3	8	PRO	-	expression tag	UNP P00722
3	418	GLU	HIS	engineered mutation	UNP P00722
4	1	GLY	-	expression tag	UNP P00722
4	2	SER	-	expression tag	UNP P00722
4	3	HIS	-	expression tag	UNP P00722
4	4	MET	-	expression tag	UNP P00722
4	5	LEU	-	expression tag	UNP P00722
4	6	GLU	-	expression tag	UNP P00722
4	7	ASP	-	expression tag	UNP P00722
4	8	PRO	-	expression tag	UNP P00722
4	418	GLU	HIS	engineered mutation	UNP P00722

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	1	3	Total Mg 3 3	0	0
2	2	3	Total Mg 3 3	0	0
2	3	2	Total Mg 2 2	0	0
2	4	2	Total Mg 2 2	0	0

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	1	3	Total Na 3 3	0	0
3	2	4	Total Na 4 4	0	0
3	3	4	Total Na 4 4	0	0
3	4	2	Total Na 2 2	0	0

- Molecule 4 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C<sub>2</sub>H<sub>6</sub>OS).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	1	1	Total	C	O	S	0	0
			4	2	1	1		
4	1	1	Total	C	O	S	0	0
			4	2	1	1		
4	1	1	Total	C	O	S	0	0
			4	2	1	1		
4	1	1	Total	C	O	S	0	0
			4	2	1	1		
4	1	1	Total	C	O	S	0	0
			4	2	1	1		
4	1	1	Total	C	O	S	0	0
			4	2	1	1		
4	1	1	Total	C	O	S	0	0
			4	2	1	1		
4	1	1	Total	C	O	S	0	0
			4	2	1	1		
4	1	1	Total	C	O	S	0	0
			4	2	1	1		
4	1	1	Total	C	O	S	0	0
			4	2	1	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	1	1	Total 4	C 2	O 1	S 1	0	0
4	1	1	Total 4	C 2	O 1	S 1	0	0
4	1	1	Total 4	C 2	O 1	S 1	0	0
4	1	1	Total 4	C 2	O 1	S 1	0	0
4	1	1	Total 4	C 2	O 1	S 1	0	0
4	1	1	Total 4	C 2	O 1	S 1	0	0
4	1	1	Total 4	C 2	O 1	S 1	0	0
4	1	1	Total 4	C 2	O 1	S 1	0	0
4	1	1	Total 4	C 2	O 1	S 1	0	0
4	1	1	Total 4	C 2	O 1	S 1	0	0
4	2	1	Total 4	C 2	O 1	S 1	0	0
4	2	1	Total 4	C 2	O 1	S 1	0	0
4	2	1	Total 4	C 2	O 1	S 1	0	0
4	2	1	Total 4	C 2	O 1	S 1	0	0
4	2	1	Total 4	C 2	O 1	S 1	0	0
4	2	1	Total 4	C 2	O 1	S 1	0	0
4	2	1	Total 4	C 2	O 1	S 1	0	0
4	2	1	Total 4	C 2	O 1	S 1	0	0
4	2	1	Total 4	C 2	O 1	S 1	0	0
4	2	1	Total 4	C 2	O 1	S 1	0	0
4	2	1	Total 4	C 2	O 1	S 1	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	2	1	Total	C	O	S	0	0
			4	2	1	1		
4	2	1	Total	C	O	S	0	0
			4	2	1	1		
4	2	1	Total	C	O	S	0	0
			4	2	1	1		
4	2	1	Total	C	O	S	0	0
			4	2	1	1		
4	2	1	Total	C	O	S	0	0
			4	2	1	1		
4	2	1	Total	C	O	S	0	0
			4	2	1	1		
4	2	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		

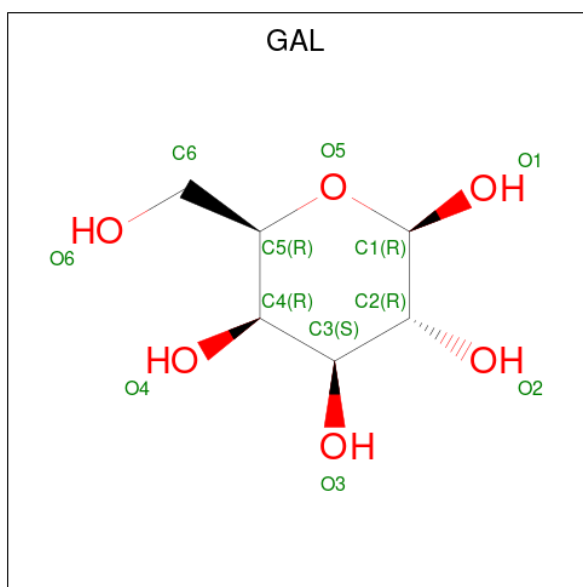
*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	3	1	Total	C	O	S	0	0
			4	2	1	1		
4	4	1	Total	C	O	S	0	0
			4	2	1	1		
4	4	1	Total	C	O	S	0	0
			4	2	1	1		
4	4	1	Total	C	O	S	0	0
			4	2	1	1		
4	4	1	Total	C	O	S	0	0
			4	2	1	1		

- Molecule 5 is beta-D-galactopyranose (three-letter code: GAL) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	1	1	Total C O 12 6 6	0	0
5	1	1	Total C O 12 6 6	0	0
5	2	1	Total C O 12 6 6	0	0
5	2	1	Total C O 12 6 6	0	0
5	3	1	Total C O 12 6 6	0	0
5	3	1	Total C O 12 6 6	0	0
5	4	1	Total C O 12 6 6	0	0
5	4	1	Total C O 12 6 6	0	0

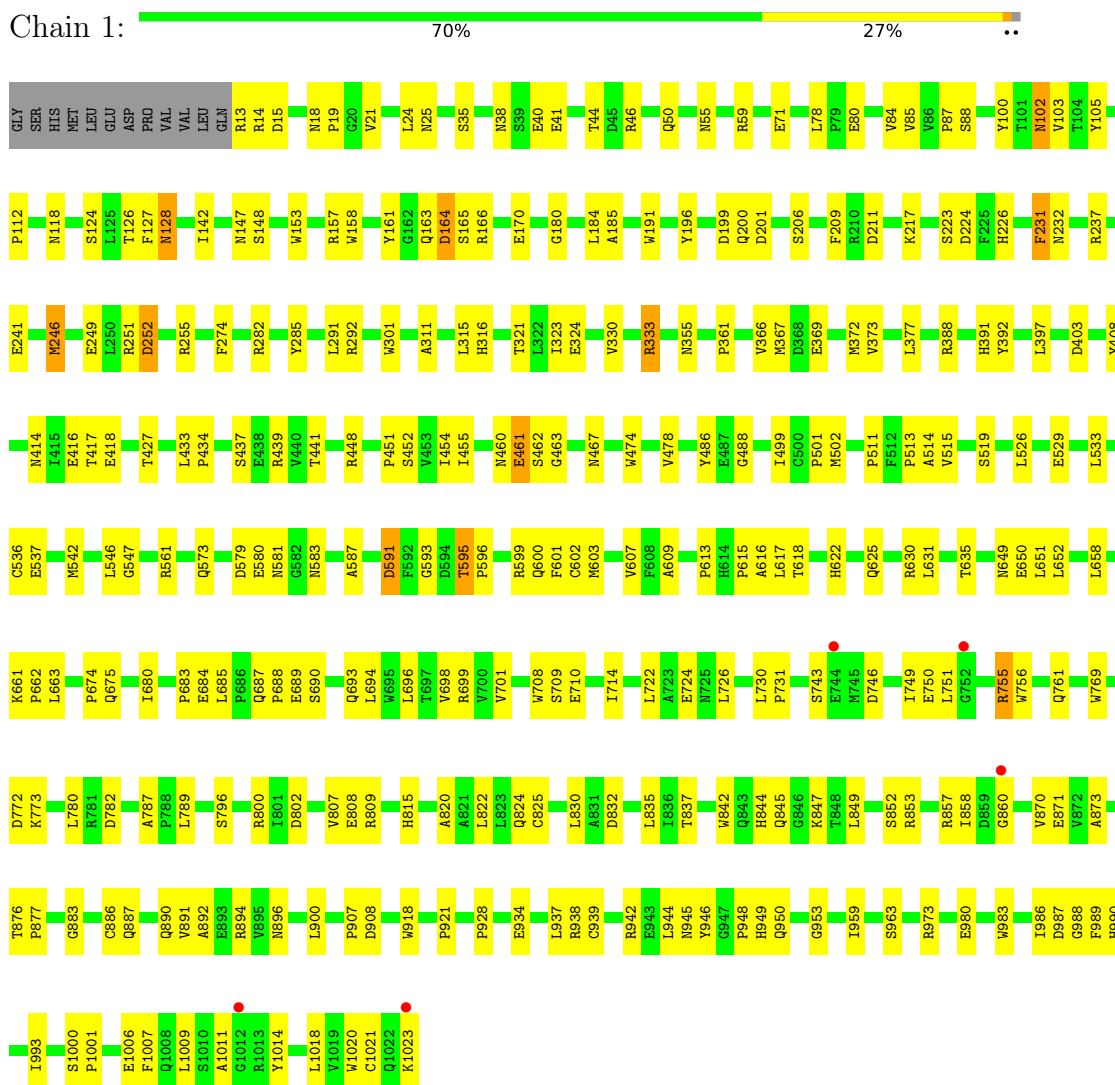
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	1	56	Total O 56 56	0	0
6	2	59	Total O 59 59	0	0
6	3	109	Total O 109 109	0	0
6	4	104	Total O 104 104	0	0

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

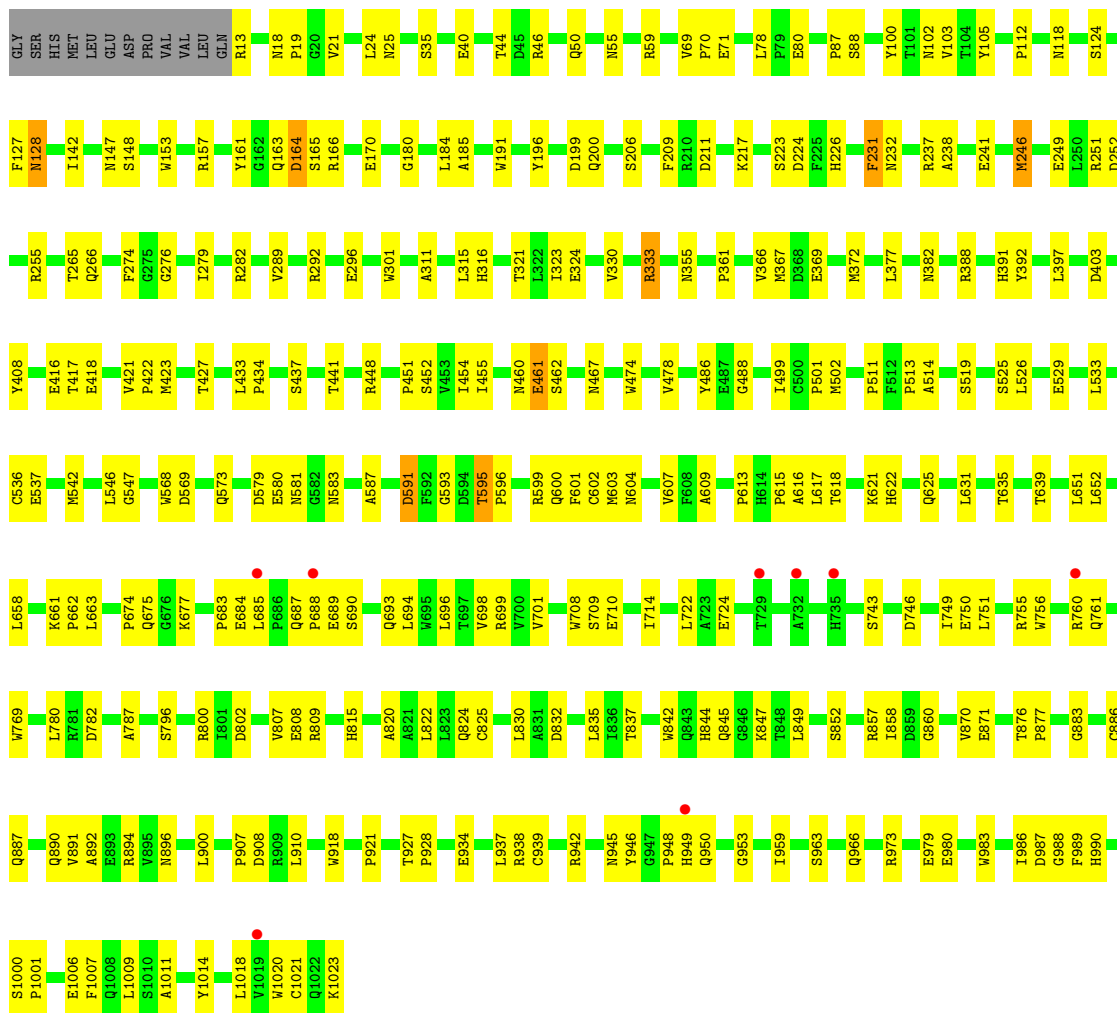
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Beta-galactosidase

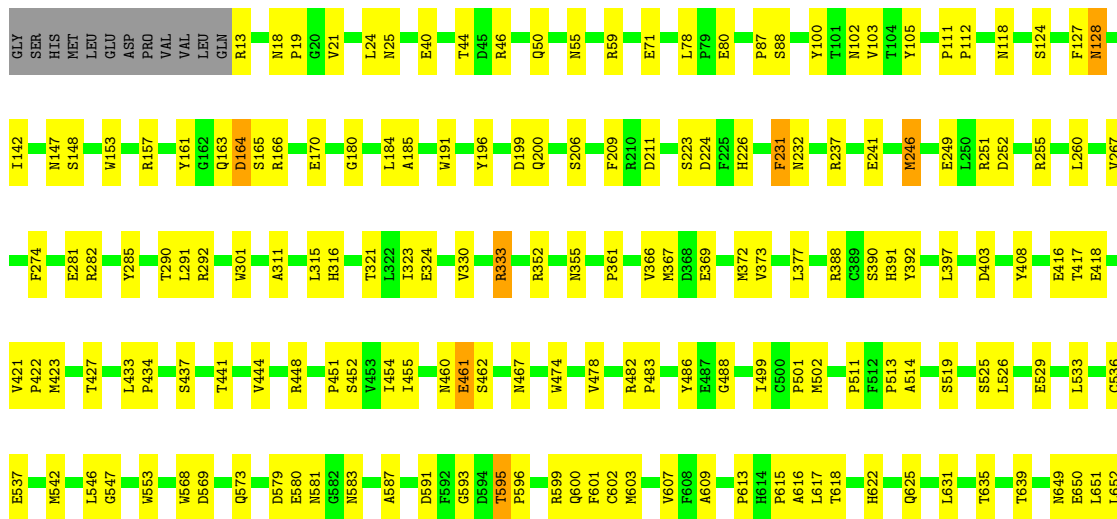


- Molecule 1: Beta-galactosidase



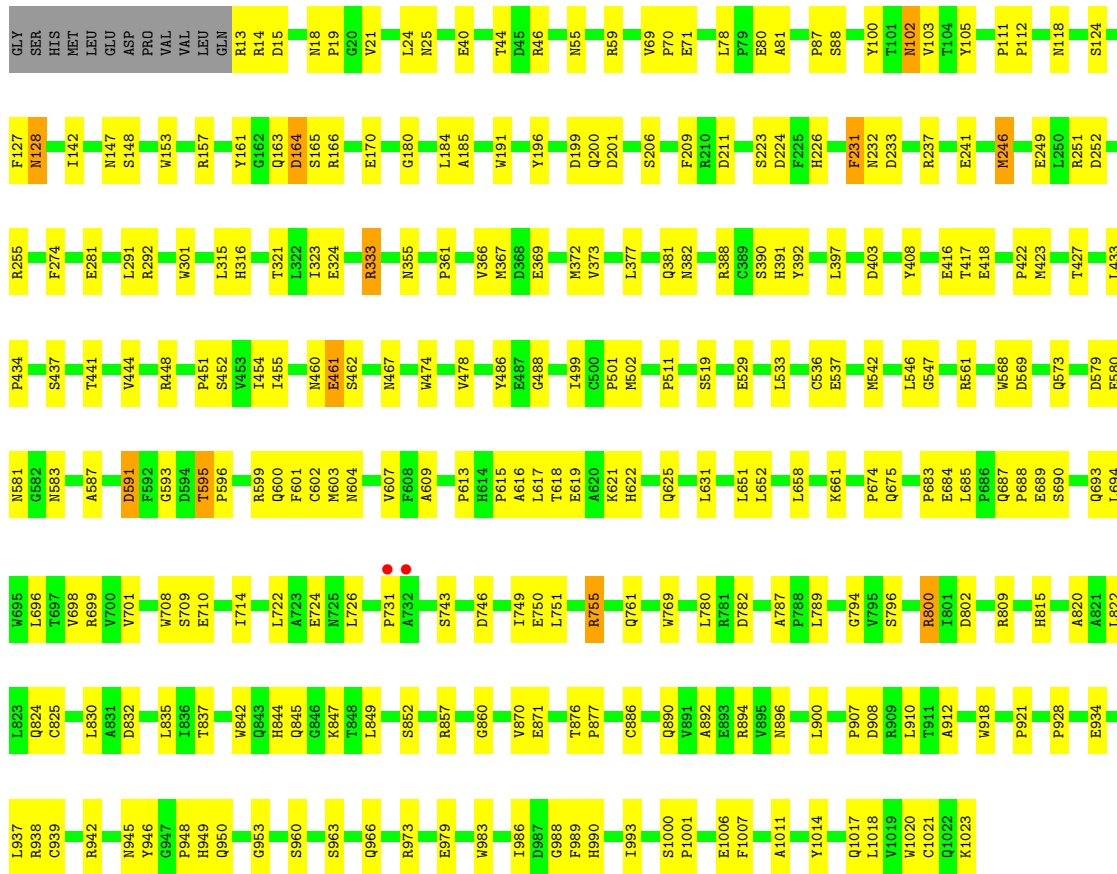


● Molecule 1: Beta-galactosidase





● Molecule 1: Beta-galactosidase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	128.37Å 152.87Å 132.12Å 90.00° 102.68° 90.00°	Depositor
Resolution (Å)	98.54 – 3.00 98.54 – 3.00	Depositor EDS
% Data completeness (in resolution range)	91.1 (98.54-3.00) 91.1 (98.54-3.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.81 (at 3.01Å)	Xtrriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.218 , 0.249 0.210 , 0.239	Depositor DCC
$R_{free}$ test set	4529 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.5	Xtrriage
Anisotropy	0.423	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 53.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.022 for l,-k,h	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	33235	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.63% 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: MG, DMS, NA, GAL

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	1	0.35	0/8365	0.69	7/11412 (0.1%)
1	2	0.37	0/8365	0.69	7/11412 (0.1%)
1	3	0.37	0/8365	0.64	6/11412 (0.1%)
1	4	0.38	0/8365	0.69	7/11412 (0.1%)
All	All	0.37	0/33460	0.68	27/45648 (0.1%)

There are no bond length outliers.

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	1	157	ARG	NE-CZ-NH1	-20.01	110.30	120.30
1	1	157	ARG	NE-CZ-NH2	19.62	130.11	120.30
1	4	800	ARG	NE-CZ-NH1	-19.58	110.51	120.30
1	4	800	ARG	NE-CZ-NH2	19.27	129.93	120.30
1	2	809	ARG	NE-CZ-NH2	18.68	129.64	120.30

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	1	8124	0	7714	222	0
1	2	8124	0	7715	212	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	3	8124	0	7714	209	0
1	4	8124	0	7714	201	0
2	1	3	0	0	0	0
2	2	3	0	0	0	0
2	3	2	0	0	0	0
2	4	2	0	0	0	0
3	1	3	0	0	0	0
3	2	4	0	0	0	0
3	3	4	0	0	0	0
3	4	2	0	0	0	0
4	1	96	0	144	8	0
4	2	80	0	120	5	0
4	3	92	0	138	3	0
4	4	24	0	36	1	0
5	1	24	0	24	1	0
5	2	24	0	24	1	0
5	3	24	0	23	0	0
5	4	24	0	24	0	0
6	1	56	0	0	2	0
6	2	59	0	0	0	0
6	3	109	0	0	2	0
6	4	104	0	0	4	0
All	All	33235	0	31390	805	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 13.

The worst 5 of 805 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)
1:2:142:ILE:HG12	1:2:170:GLU:HG2	1.53	0.91
1:1:652:LEU:HD11	1:1:698:VAL:HB	1.51	0.90
1:3:142:ILE:HG12	1:3:170:GLU:HG2	1.52	0.90
1:3:652:LEU:HD11	1:3:698:VAL:HB	1.52	0.90
1:1:142:ILE:HG12	1:1:170:GLU:HG2	1.54	0.90

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	1	1009/1023 (99%)	912 (90%)	85 (8%)	12 (1%)	13	48
1	2	1009/1023 (99%)	907 (90%)	91 (9%)	11 (1%)	14	50
1	3	1009/1023 (99%)	911 (90%)	88 (9%)	10 (1%)	15	53
1	4	1009/1023 (99%)	911 (90%)	87 (9%)	11 (1%)	14	50
All	All	4036/4092 (99%)	3641 (90%)	351 (9%)	44 (1%)	14	50

5 of 44 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	1	722	LEU
1	2	722	LEU
1	3	722	LEU
1	4	722	LEU
1	1	164	ASP

### 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	1	864/875 (99%)	853 (99%)	11 (1%)	69	89
1	2	864/875 (99%)	854 (99%)	10 (1%)	71	90
1	3	864/875 (99%)	852 (99%)	12 (1%)	67	88
1	4	864/875 (99%)	851 (98%)	13 (2%)	65	87
All	All	3456/3500 (99%)	3410 (99%)	46 (1%)	69	89

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	3	519	SER
1	4	231	PHE
1	3	595	THR
1	3	1017	GLN
1	4	252	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	3	510	GLN
1	3	878	HIS
1	3	757	GLN
1	3	887	GLN
1	2	38	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 104 ligands modelled in this entry, 23 are monoatomic - leaving 81 for Mogul analysis.

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
4	DMS	4	8003	-	3,3,3	0.31	0	3,3,3	0.68	0
5	GAL	4	2001	3	12,12,12	1.14	1 (8%)	17,17,17	0.78	0
4	DMS	2	8001	-	3,3,3	0.28	0	3,3,3	0.67	0
4	DMS	1	8421	-	3,3,3	0.27	0	3,3,3	0.67	0
4	DMS	4	8005	-	3,3,3	0.29	0	3,3,3	0.70	0
4	DMS	2	8401	-	3,3,3	0.26	0	3,3,3	0.59	0
4	DMS	3	8420	-	3,3,3	0.26	0	3,3,3	0.61	0
4	DMS	3	8406	-	3,3,3	0.26	0	3,3,3	0.65	0
4	DMS	1	8003	-	3,3,3	0.26	0	3,3,3	0.65	0
4	DMS	2	8404	-	3,3,3	0.23	0	3,3,3	0.62	0
4	DMS	2	8002	-	3,3,3	0.19	0	3,3,3	0.61	0
4	DMS	1	8602	-	3,3,3	0.26	0	3,3,3	0.65	0
4	DMS	2	8000	-	3,3,3	0.31	0	3,3,3	0.65	0
4	DMS	2	8414	-	3,3,3	0.24	0	3,3,3	0.64	0
4	DMS	3	8414	-	3,3,3	0.24	0	3,3,3	0.65	0
4	DMS	1	8002	-	3,3,3	0.22	0	3,3,3	0.63	0
4	DMS	2	8423	-	3,3,3	0.26	0	3,3,3	0.69	0
4	DMS	1	8410	-	3,3,3	0.23	0	3,3,3	0.65	0
4	DMS	2	8402	-	3,3,3	0.24	0	3,3,3	0.62	0
4	DMS	1	8417	-	3,3,3	0.23	0	3,3,3	0.65	0
4	DMS	3	8503	-	3,3,3	0.27	0	3,3,3	0.67	0
4	DMS	2	8421	-	3,3,3	0.25	0	3,3,3	0.65	0
4	DMS	1	8503	-	3,3,3	0.28	0	3,3,3	0.65	0
4	DMS	1	8403	-	3,3,3	0.28	0	3,3,3	0.68	0
4	DMS	3	8411	-	3,3,3	0.26	0	3,3,3	0.66	0
4	DMS	3	8404	-	3,3,3	0.27	0	3,3,3	0.71	0
5	GAL	4	2002	-	12,12,12	0.81	0	17,17,17	0.83	0
4	DMS	1	8501	-	3,3,3	0.27	0	3,3,3	0.59	0
5	GAL	2	2001	3	12,12,12	0.82	0	17,17,17	0.81	0
5	GAL	3	2002	-	12,12,12	0.87	0	17,17,17	0.85	0
4	DMS	2	8409	-	3,3,3	0.24	0	3,3,3	0.65	0
4	DMS	1	8001	-	3,3,3	0.23	0	3,3,3	0.66	0
4	DMS	1	8406	-	3,3,3	0.20	0	3,3,3	0.60	0
4	DMS	3	8425	-	3,3,3	0.21	0	3,3,3	0.66	0
4	DMS	1	8409	-	3,3,3	0.25	0	3,3,3	0.67	0
4	DMS	2	8501	-	3,3,3	0.28	0	3,3,3	0.63	0
4	DMS	3	8419	-	3,3,3	0.28	0	3,3,3	0.67	0
4	DMS	2	8411	-	3,3,3	0.25	0	3,3,3	0.62	0
4	DMS	3	8415	-	3,3,3	0.27	0	3,3,3	0.63	0
4	DMS	3	8421	-	3,3,3	0.24	0	3,3,3	0.64	0
4	DMS	1	8502	-	3,3,3	0.25	0	3,3,3	0.59	0
4	DMS	2	8427	-	3,3,3	0.21	0	3,3,3	0.65	0
4	DMS	2	8419	-	3,3,3	0.27	0	3,3,3	0.66	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	DMS	4	8004	-	3,3,3	0.27	0	3,3,3	0.62	0
4	DMS	3	8402	-	3,3,3	0.27	0	3,3,3	0.65	0
4	DMS	3	8407	-	3,3,3	0.25	0	3,3,3	0.70	0
4	DMS	3	8001	-	3,3,3	0.23	0	3,3,3	0.64	0
4	DMS	1	8405	-	3,3,3	0.31	0	3,3,3	0.66	0
4	DMS	3	8401	-	3,3,3	0.25	0	3,3,3	0.62	0
4	DMS	1	8416	-	3,3,3	0.25	0	3,3,3	0.66	0
4	DMS	3	8405	-	3,3,3	0.25	0	3,3,3	0.68	0
4	DMS	2	8003	-	3,3,3	0.24	0	3,3,3	0.66	0
4	DMS	2	8504	-	3,3,3	0.26	0	3,3,3	0.58	0
4	DMS	4	8002	-	3,3,3	0.29	0	3,3,3	0.65	0
5	GAL	3	2001	3	12,12,12	0.73	0	17,17,17	0.79	0
4	DMS	2	8503	-	3,3,3	0.26	0	3,3,3	0.67	0
5	GAL	1	2001	2,3	12,12,12	0.77	0	17,17,17	0.87	0
4	DMS	2	8405	-	3,3,3	0.27	0	3,3,3	0.68	0
4	DMS	1	8423	-	3,3,3	0.25	0	3,3,3	0.64	0
4	DMS	1	8425	-	3,3,3	0.25	0	3,3,3	0.64	0
4	DMS	3	8403	-	3,3,3	0.18	0	3,3,3	0.61	0
4	DMS	1	8402	-	3,3,3	0.24	0	3,3,3	0.63	0
4	DMS	3	8000	-	3,3,3	0.26	0	3,3,3	0.67	0
4	DMS	1	8411	-	3,3,3	0.21	0	3,3,3	0.69	0
4	DMS	2	8601	-	3,3,3	0.30	0	3,3,3	0.65	0
4	DMS	1	8404	-	3,3,3	0.22	0	3,3,3	0.69	0
4	DMS	3	8409	-	3,3,3	0.29	0	3,3,3	0.68	0
4	DMS	3	8705	-	3,3,3	0.23	0	3,3,3	0.65	0
4	DMS	1	8401	-	3,3,3	0.19	0	3,3,3	0.55	0
4	DMS	3	8408	-	3,3,3	0.24	0	3,3,3	0.69	0
5	GAL	2	2002	-	12,12,12	0.99	0	17,17,17	0.81	0
4	DMS	1	8414	-	3,3,3	0.31	0	3,3,3	0.66	0
4	DMS	4	8006	-	3,3,3	0.28	0	3,3,3	0.72	0
4	DMS	3	8416	-	3,3,3	0.24	0	3,3,3	0.68	0
4	DMS	1	8415	-	3,3,3	0.29	0	3,3,3	0.62	0
4	DMS	2	8403	-	3,3,3	0.31	0	3,3,3	0.67	0
4	DMS	1	8408	-	3,3,3	0.31	0	3,3,3	0.65	0
4	DMS	3	8501	-	3,3,3	0.24	0	3,3,3	0.64	0
5	GAL	1	2002	-	12,12,12	0.99	0	17,17,17	0.83	0
4	DMS	4	8001	-	3,3,3	0.21	0	3,3,3	0.58	0
4	DMS	3	8417	-	3,3,3	0.26	0	3,3,3	0.69	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
5	GAL	2	2001	3	-	1/2/22/22	0/1/1/1
5	GAL	4	2001	3	-	1/2/22/22	0/1/1/1
5	GAL	1	2002	-	-	2/2/22/22	0/1/1/1
5	GAL	1	2001	2,3	-	2/2/22/22	0/1/1/1
5	GAL	3	2002	-	-	2/2/22/22	0/1/1/1
5	GAL	2	2002	-	-	2/2/22/22	0/1/1/1
5	GAL	4	2002	-	-	2/2/22/22	0/1/1/1
5	GAL	3	2001	3	-	1/2/22/22	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	4	2001	GAL	O5-C1	2.31	1.48	1.42

There are no bond angle outliers.

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	2	2002	GAL	O5-C5-C6-O6
5	2	2002	GAL	C4-C5-C6-O6
5	1	2001	GAL	O5-C5-C6-O6
5	1	2001	GAL	C4-C5-C6-O6
5	1	2002	GAL	O5-C5-C6-O6

There are no ring outliers.

11 monomers are involved in 19 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	2	8001	DMS	4	0
4	1	8503	DMS	1	0
5	2	2001	GAL	1	0
4	3	8415	DMS	1	0
4	1	8502	DMS	3	0
4	4	8004	DMS	1	0
4	3	8001	DMS	2	0
4	1	8416	DMS	1	0
5	1	2001	GAL	1	0
4	2	8601	DMS	1	0
4	1	8414	DMS	3	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	1	1011/1023 (98%)	-0.14	5 (0%) 91 75	18, 51, 89, 108	0
1	2	1011/1023 (98%)	-0.28	8 (0%) 86 65	15, 39, 91, 114	0
1	3	1011/1023 (98%)	-0.42	2 (0%) 95 87	16, 35, 57, 89	0
1	4	1011/1023 (98%)	-0.40	2 (0%) 95 87	13, 36, 59, 86	0
All	All	4044/4092 (98%)	-0.31	17 (0%) 92 79	13, 39, 79, 114	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	4	731	PRO	2.7
1	2	729	THR	2.6
1	1	1023	LYS	2.5
1	1	744	GLU	2.4
1	1	1012	GLY	2.4

### 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( $\text{\AA}^2$ )	Q<0.9
4	DMS	3	8421	4/4	0.35	0.84	120,121,121,121	0
4	DMS	3	8416	4/4	0.44	0.59	106,106,107,108	0
3	NA	2	3102	1/1	0.47	0.26	49,49,49,49	0
4	DMS	1	8416	4/4	0.48	0.55	115,115,115,117	0
4	DMS	1	8417	4/4	0.49	0.35	106,106,106,108	0
2	MG	2	3002	1/1	0.57	0.29	26,26,26,26	1
4	DMS	3	8415	4/4	0.58	1.12	133,133,133,134	0
4	DMS	1	8425	4/4	0.65	0.63	124,125,125,126	0
4	DMS	1	8410	4/4	0.71	0.39	110,110,110,111	0
4	DMS	1	8503	4/4	0.71	0.47	115,116,116,116	0
4	DMS	3	8414	4/4	0.71	0.52	114,114,115,115	0
2	MG	1	3001	1/1	0.72	0.15	131,131,131,131	0
3	NA	3	3102	1/1	0.78	0.20	44,44,44,44	0
4	DMS	1	8602	4/4	0.79	0.27	115,115,115,115	0
4	DMS	2	8503	4/4	0.79	0.38	105,106,106,107	0
4	DMS	1	8421	4/4	0.79	0.31	88,88,89,90	0
2	MG	1	3002	1/1	0.82	0.23	63,63,63,63	0
3	NA	4	3100	1/1	0.83	0.21	74,74,74,74	0
4	DMS	2	8423	4/4	0.84	0.27	84,85,85,86	0
4	DMS	3	8503	4/4	0.84	0.39	105,105,106,106	0
4	DMS	2	8421	4/4	0.85	0.44	116,116,116,117	0
5	GAL	2	2002	12/12	0.86	0.31	81,82,83,84	0
4	DMS	1	8502	4/4	0.87	0.50	91,91,92,92	0
3	NA	2	3100	1/1	0.87	0.14	36,36,36,36	0
4	DMS	3	8417	4/4	0.88	0.18	82,82,83,84	0
4	DMS	1	8415	4/4	0.89	0.24	84,84,85,85	0
4	DMS	2	8427	4/4	0.89	0.21	78,79,79,80	0
5	GAL	3	2002	12/12	0.89	0.26	76,78,79,80	0
3	NA	2	3103	1/1	0.90	0.13	30,30,30,30	0
5	GAL	1	2002	12/12	0.90	0.17	76,79,80,80	0
4	DMS	2	8003	4/4	0.91	0.26	75,75,75,76	0
3	NA	1	3100	1/1	0.91	0.10	44,44,44,44	0
4	DMS	3	8411	4/4	0.92	0.29	67,67,68,69	0
4	DMS	1	8406	4/4	0.92	0.20	82,82,83,84	0
4	DMS	1	8409	4/4	0.92	0.20	65,66,67,67	0
3	NA	3	3103	1/1	0.92	0.18	48,48,48,48	0
3	NA	1	3103	1/1	0.92	0.14	43,43,43,43	0
4	DMS	1	8414	4/4	0.93	0.27	54,54,56,56	0
4	DMS	2	8000	4/4	0.93	0.24	56,57,57,58	0
4	DMS	3	8705	4/4	0.93	0.19	107,107,107,107	0
4	DMS	4	8004	4/4	0.93	0.19	80,81,81,81	0
2	MG	2	3000	1/1	0.93	0.10	30,30,30,30	0
4	DMS	3	8406	4/4	0.93	0.19	78,79,79,80	0

Continued on next page...



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	DMS	3	8407	4/4	0.93	0.20	79,79,80,81	0
3	NA	4	3101	1/1	0.94	0.14	31,31,31,31	0
4	DMS	3	8408	4/4	0.94	0.20	58,59,59,60	0
4	DMS	4	8003	4/4	0.94	0.14	56,56,57,58	0
4	DMS	1	8003	4/4	0.94	0.21	79,79,79,80	0
5	GAL	1	2001	12/12	0.94	0.14	37,40,43,44	0
4	DMS	3	8419	4/4	0.94	0.22	85,86,86,86	0
3	NA	2	3101	1/1	0.94	0.15	25,25,25,25	0
4	DMS	3	8425	4/4	0.94	0.20	5,6,6,7	4
4	DMS	1	8402	4/4	0.95	0.25	66,66,67,67	0
4	DMS	3	8000	4/4	0.95	0.23	55,55,56,57	0
4	DMS	2	8409	4/4	0.95	0.20	42,43,43,43	0
4	DMS	1	8405	4/4	0.95	0.24	62,62,63,64	0
4	DMS	4	8005	4/4	0.95	0.26	74,74,76,76	0
4	DMS	4	8006	4/4	0.95	0.14	57,59,59,60	0
2	MG	3	3001	1/1	0.95	0.08	48,48,48,48	0
4	DMS	1	8408	4/4	0.95	0.28	62,62,62,63	0
4	DMS	2	8501	4/4	0.95	0.18	62,62,62,63	0
2	MG	4	3000	1/1	0.95	0.06	40,40,40,40	0
5	GAL	4	2002	12/12	0.95	0.21	79,80,81,82	0
4	DMS	1	8001	4/4	0.96	0.19	73,73,74,75	0
4	DMS	1	8002	4/4	0.96	0.21	72,73,73,74	0
3	NA	3	3101	1/1	0.96	0.11	38,38,38,38	0
4	DMS	2	8504	4/4	0.96	0.28	53,53,54,54	0
4	DMS	2	8403	4/4	0.96	0.19	40,41,42,43	0
2	MG	2	3001	1/1	0.96	0.12	57,57,57,57	0
4	DMS	3	8403	4/4	0.96	0.22	48,48,50,50	0
5	GAL	2	2001	12/12	0.96	0.12	39,41,42,42	0
4	DMS	2	8414	4/4	0.96	0.29	74,74,74,74	0
5	GAL	3	2001	12/12	0.96	0.17	30,32,33,35	0
4	DMS	1	8411	4/4	0.96	0.21	47,48,48,49	0
5	GAL	4	2001	12/12	0.96	0.15	21,29,30,31	0
4	DMS	1	8404	4/4	0.96	0.15	50,50,50,51	0
4	DMS	3	8405	4/4	0.97	0.24	51,52,52,53	0
3	NA	1	3101	1/1	0.97	0.11	30,30,30,30	0
4	DMS	3	8001	4/4	0.97	0.22	55,55,55,56	0
4	DMS	4	8002	4/4	0.97	0.24	48,49,50,51	0
4	DMS	1	8501	4/4	0.97	0.20	58,58,58,59	0
2	MG	3	3000	1/1	0.97	0.10	9,9,9,9	0
4	DMS	3	8409	4/4	0.97	0.14	46,47,47,48	0
3	NA	3	3100	1/1	0.97	0.15	16,16,16,16	0
4	DMS	2	8411	4/4	0.97	0.25	53,53,54,54	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	DMS	1	8403	4/4	0.97	0.13	52,52,54,54	0
4	DMS	2	8001	4/4	0.97	0.21	46,46,46,47	0
4	DMS	2	8002	4/4	0.97	0.30	67,67,68,68	0
4	DMS	2	8419	4/4	0.97	0.21	69,69,70,70	0
4	DMS	3	8420	4/4	0.97	0.16	65,65,66,66	0
4	DMS	1	8423	4/4	0.97	0.23	82,83,84,84	0
4	DMS	3	8404	4/4	0.97	0.19	39,40,40,41	0
4	DMS	2	8401	4/4	0.98	0.21	45,45,46,46	0
4	DMS	3	8501	4/4	0.98	0.21	44,44,45,45	0
4	DMS	3	8402	4/4	0.98	0.18	45,46,47,47	0
4	DMS	1	8401	4/4	0.98	0.21	39,40,41,41	0
4	DMS	2	8601	4/4	0.98	0.20	63,63,63,64	0
4	DMS	2	8404	4/4	0.98	0.23	44,44,45,46	0
4	DMS	4	8001	4/4	0.98	0.19	30,30,31,32	0
4	DMS	2	8405	4/4	0.98	0.22	45,46,46,47	0
2	MG	1	3000	1/1	0.98	0.09	18,18,18,18	0
2	MG	4	3004	1/1	0.98	0.11	1,1,1,1	0
4	DMS	3	8401	4/4	0.99	0.17	36,38,38,38	0
4	DMS	2	8402	4/4	0.99	0.26	71,71,72,72	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.