



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 08:48 am BST

PDB ID : 1JZ6  
Title : E. COLI (lacZ) BETA-GALACTOSIDASE IN COMPLEX WITH GALACTO-TETRAZOLE  
Authors : Juers, D.H.; Heightman, T.D.; Vasella, A.; Matthews, B.W.  
Deposited on : 2001-09-13  
Resolution : 2.10 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 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 : 2.11  
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) : 2.11

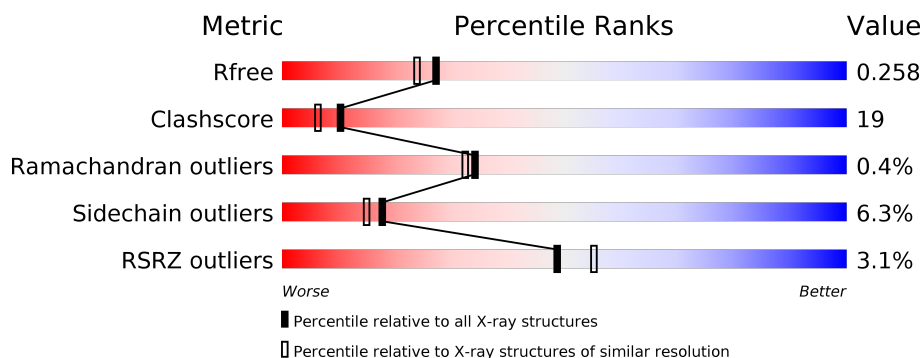
# 1 Overall quality at a glance ⓘ

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))
$R_{free}$	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (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 on 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	1023	<div> <div>3%</div> <div>61%</div> <div>31%</div> <div>7%</div> <div>..</div> </div>
1	B	1023	<div> <div>3%</div> <div>54%</div> <div>35%</div> <div>9%</div> <div>..</div> </div>
1	C	1023	<div> <div>3%</div> <div>60%</div> <div>31%</div> <div>7%</div> <div>..</div> </div>
1	D	1023	<div> <div>3%</div> <div>59%</div> <div>33%</div> <div>6%</div> <div>..</div> </div>

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

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	DMS	A	8404	-	-	X	-
5	DMS	B	8416	-	-	X	-
5	DMS	C	8423	-	-	X	-
5	DMS	C	8504	-	-	X	-
5	DMS	D	8403	-	-	X	-
5	DMS	D	8404	-	-	X	-

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 35881 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	A	1011	Total	C	N	O	S	0	0	0
			8125	5138	1440	1509	38			
1	B	1011	Total	C	N	O	S	0	0	0
			8125	5138	1440	1509	38			
1	C	1011	Total	C	N	O	S	0	0	0
			8125	5138	1440	1509	38			
1	D	1011	Total	C	N	O	S	0	0	0
			8125	5138	1440	1509	38			

There are 32 discrepancies between the modelled and reference sequences:

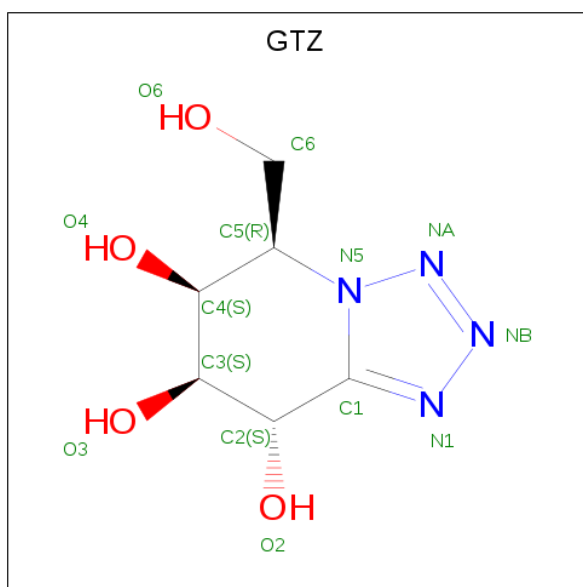
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	THR	CLONING ARTIFACT	? P00722
A	2	SER	MET	CLONING ARTIFACT	? P00722
A	3	HIS	ILE	CLONING ARTIFACT	? P00722
A	4	MET	THR	CLONING ARTIFACT	? P00722
A	5	LEU	ASP	CLONING ARTIFACT	? P00722
A	6	GLU	SER	CLONING ARTIFACT	? P00722
A	7	ASP	LEU	CLONING ARTIFACT	? P00722
A	8	PRO	ALA	CLONING ARTIFACT	? P00722
B	1	GLY	THR	CLONING ARTIFACT	? P00722
B	2	SER	MET	CLONING ARTIFACT	? P00722
B	3	HIS	ILE	CLONING ARTIFACT	? P00722
B	4	MET	THR	CLONING ARTIFACT	? P00722
B	5	LEU	ASP	CLONING ARTIFACT	? P00722
B	6	GLU	SER	CLONING ARTIFACT	? P00722
B	7	ASP	LEU	CLONING ARTIFACT	? P00722
B	8	PRO	ALA	CLONING ARTIFACT	? P00722
C	1	GLY	THR	CLONING ARTIFACT	? P00722
C	2	SER	MET	CLONING ARTIFACT	? P00722
C	3	HIS	ILE	CLONING ARTIFACT	? P00722
C	4	MET	THR	CLONING ARTIFACT	? P00722
C	5	LEU	ASP	CLONING ARTIFACT	? P00722

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	6	GLU	SER	CLONING ARTIFACT	? P00722
C	7	ASP	LEU	CLONING ARTIFACT	? P00722
C	8	PRO	ALA	CLONING ARTIFACT	? P00722
D	1	GLY	THR	CLONING ARTIFACT	? P00722
D	2	SER	MET	CLONING ARTIFACT	? P00722
D	3	HIS	ILE	CLONING ARTIFACT	? P00722
D	4	MET	THR	CLONING ARTIFACT	? P00722
D	5	LEU	ASP	CLONING ARTIFACT	? P00722
D	6	GLU	SER	CLONING ARTIFACT	? P00722
D	7	ASP	LEU	CLONING ARTIFACT	? P00722
D	8	PRO	ALA	CLONING ARTIFACT	? P00722

- Molecule 2 is (5R, 6S, 7S, 8S)-5-HYDROXYMETHYL-6,7,8-TRIHYDROXY-TETRAZOL O[1,5-A]PIPERIDINE (three-letter code: GTZ) (formula: C<sub>6</sub>H<sub>10</sub>N<sub>4</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	6	4	4		
2	B	1	Total	C	N	O	0	0
			14	6	4	4		
2	C	1	Total	C	N	O	0	0
			14	6	4	4		
2	D	1	Total	C	N	O	0	0
			14	6	4	4		

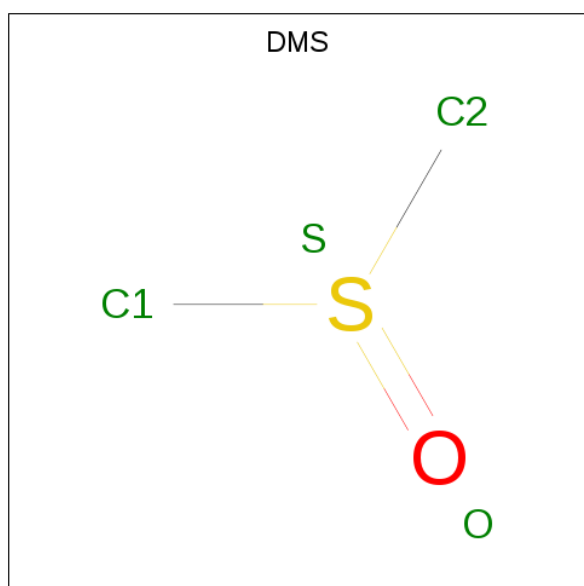
- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	3	Total Mg 3 3	0	0
3	A	4	Total Mg 4 4	0	0
3	D	3	Total Mg 3 3	0	0
3	C	3	Total Mg 3 3	0	0

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	3	Total Na 3 3	0	0
4	A	3	Total Na 3 3	0	0
4	D	3	Total Na 3 3	0	0
4	C	2	Total Na 2 2	0	0

- Molecule 5 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C<sub>2</sub>H<sub>6</sub>OS).



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

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	A	1	Total 4	C 2	O 1	S 1	0	0
5	B	1	Total 4	C 2	O 1	S 1	0	0
5	B	1	Total 4	C 2	O 1	S 1	0	0
5	B	1	Total 4	C 2	O 1	S 1	0	0
5	B	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
5	B	1	Total 4	C 2	O 1	S 1	0	0
5	B	1	Total 4	C 2	O 1	S 1	0	0
5	B	1	Total 4	C 2	O 1	S 1	0	0
5	B	1	Total 4	C 2	O 1	S 1	0	0
5	B	1	Total 4	C 2	O 1	S 1	0	0
5	B	1	Total 4	C 2	O 1	S 1	0	0
5	B	1	Total 4	C 2	O 1	S 1	0	0
5	B	1	Total 4	C 2	O 1	S 1	0	0
5	B	1	Total 4	C 2	O 1	S 1	0	0
5	B	1	Total 4	C 2	O 1	S 1	0	0
5	B	1	Total 4	C 2	O 1	S 1	0	0
5	B	1	Total 4	C 2	O 1	S 1	0	0
5	B	1	Total 4	C 2	O 1	S 1	0	0
5	B	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	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
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	C	1	Total 4	C 2	O 1	S 1	0	0
5	D	1	Total 4	C 2	O 1	S 1	0	0
5	D	1	Total 4	C 2	O 1	S 1	0	0
5	D	1	Total 4	C 2	O 1	S 1	0	0
5	D	1	Total 4	C 2	O 1	S 1	0	0
5	D	1	Total 4	C 2	O 1	S 1	0	0
5	D	1	Total 4	C 2	O 1	S 1	0	0
5	D	1	Total 4	C 2	O 1	S 1	0	0
5	D	1	Total 4	C 2	O 1	S 1	0	0
5	D	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
5	D	1	Total 4	C 2	O 1	S 1	0	0
5	D	1	Total 4	C 2	O 1	S 1	0	0
5	D	1	Total 4	C 2	O 1	S 1	0	0
5	D	1	Total 4	C 2	O 1	S 1	0	0
5	D	1	Total 4	C 2	O 1	S 1	0	0
5	D	1	Total 4	C 2	O 1	S 1	0	0
5	D	1	Total 4	C 2	O 1	S 1	0	0
5	D	1	Total 4	C 2	O 1	S 1	0	0
5	D	1	Total 4	C 2	O 1	S 1	0	0

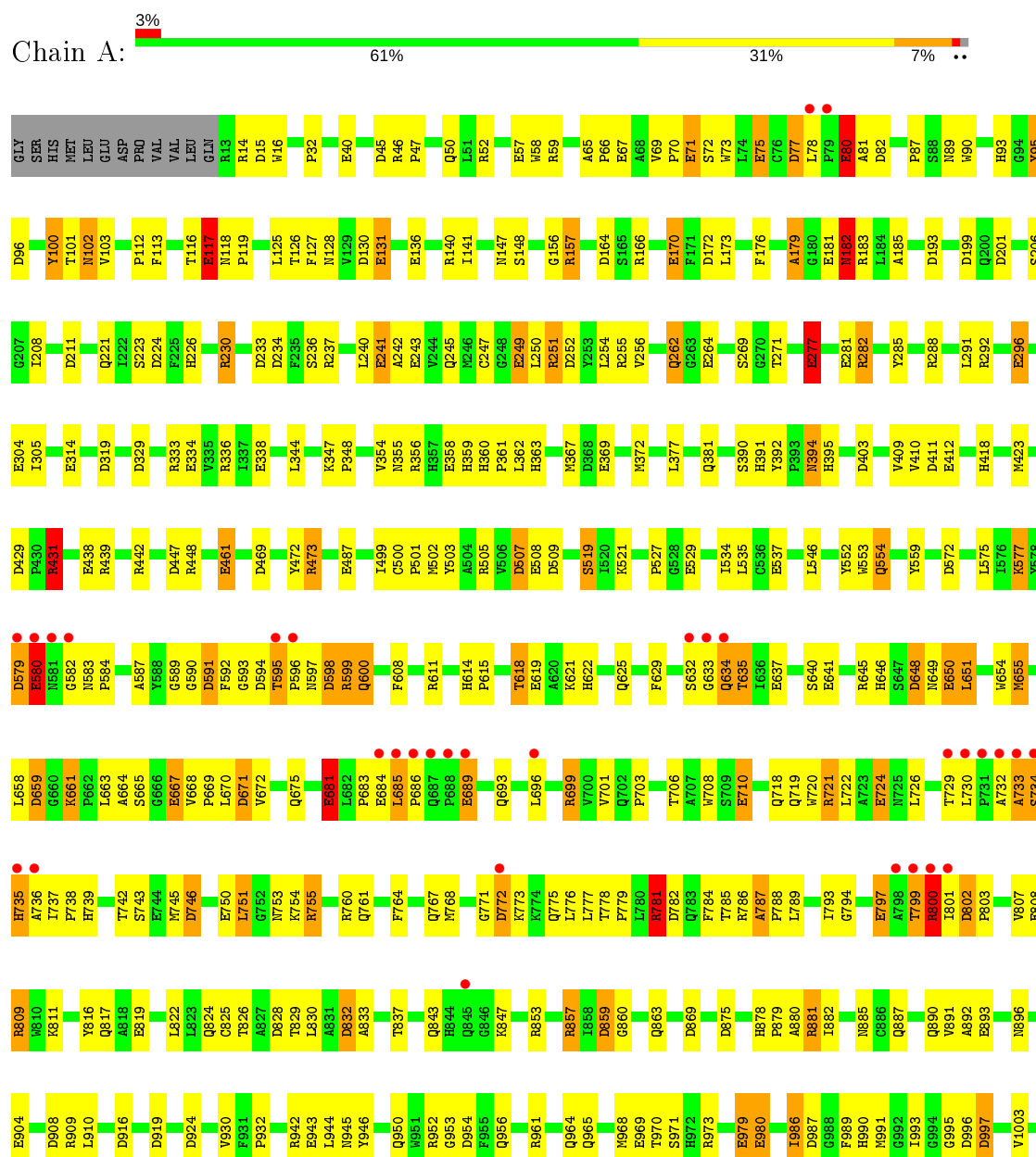
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	775	Total 775	O 775	0	0
6	B	747	Total 747	O 747	0	0
6	C	719	Total 719	O 719	0	0
6	D	764	Total 764	O 764	0	0

### 3 Residue-property plots

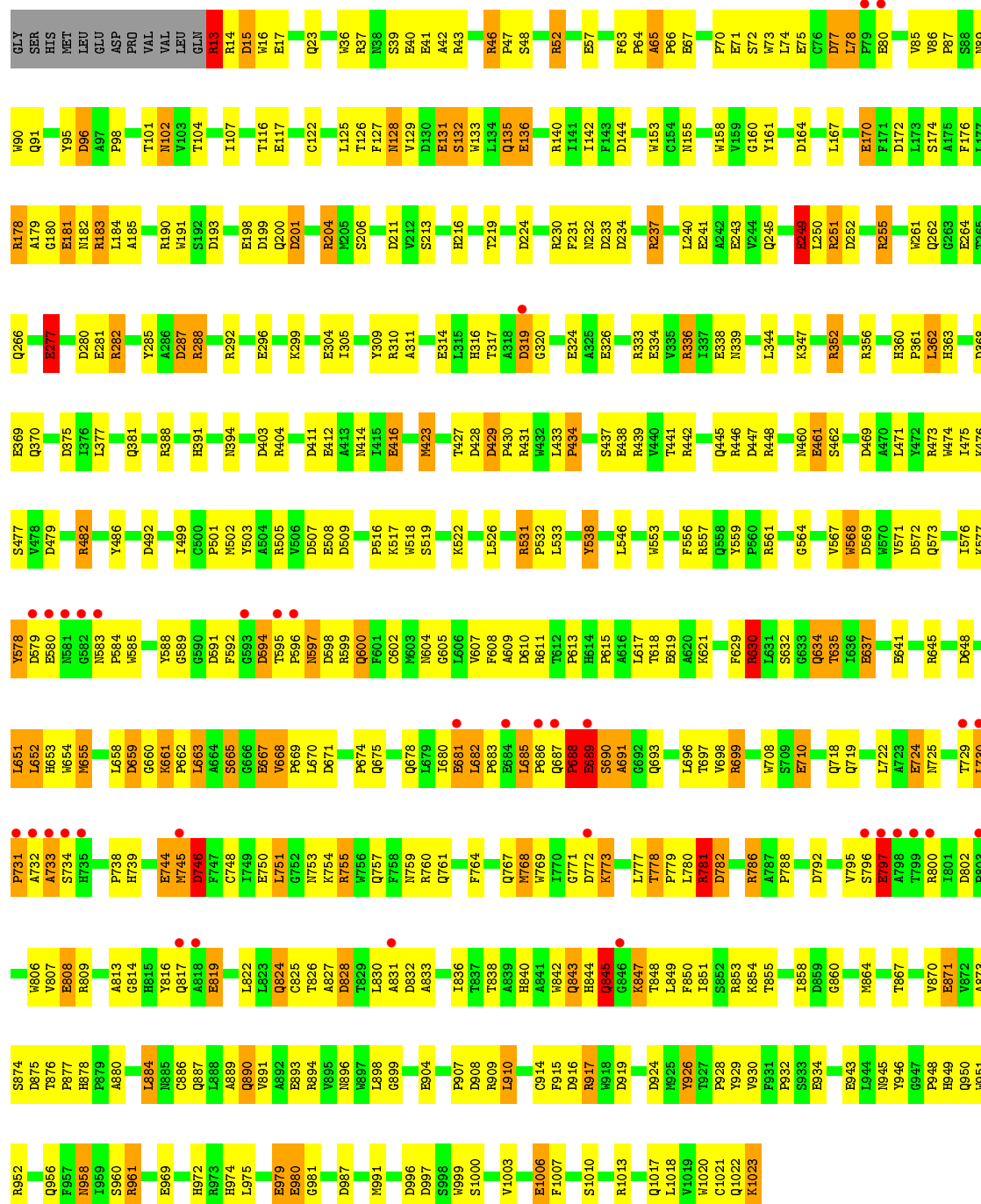
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are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

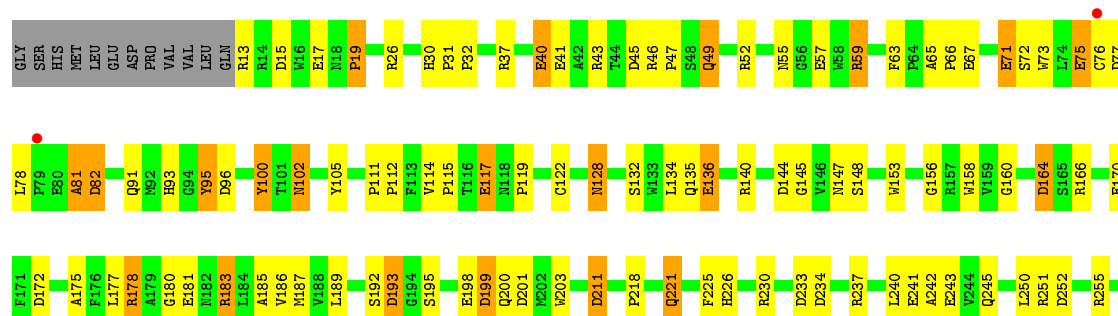


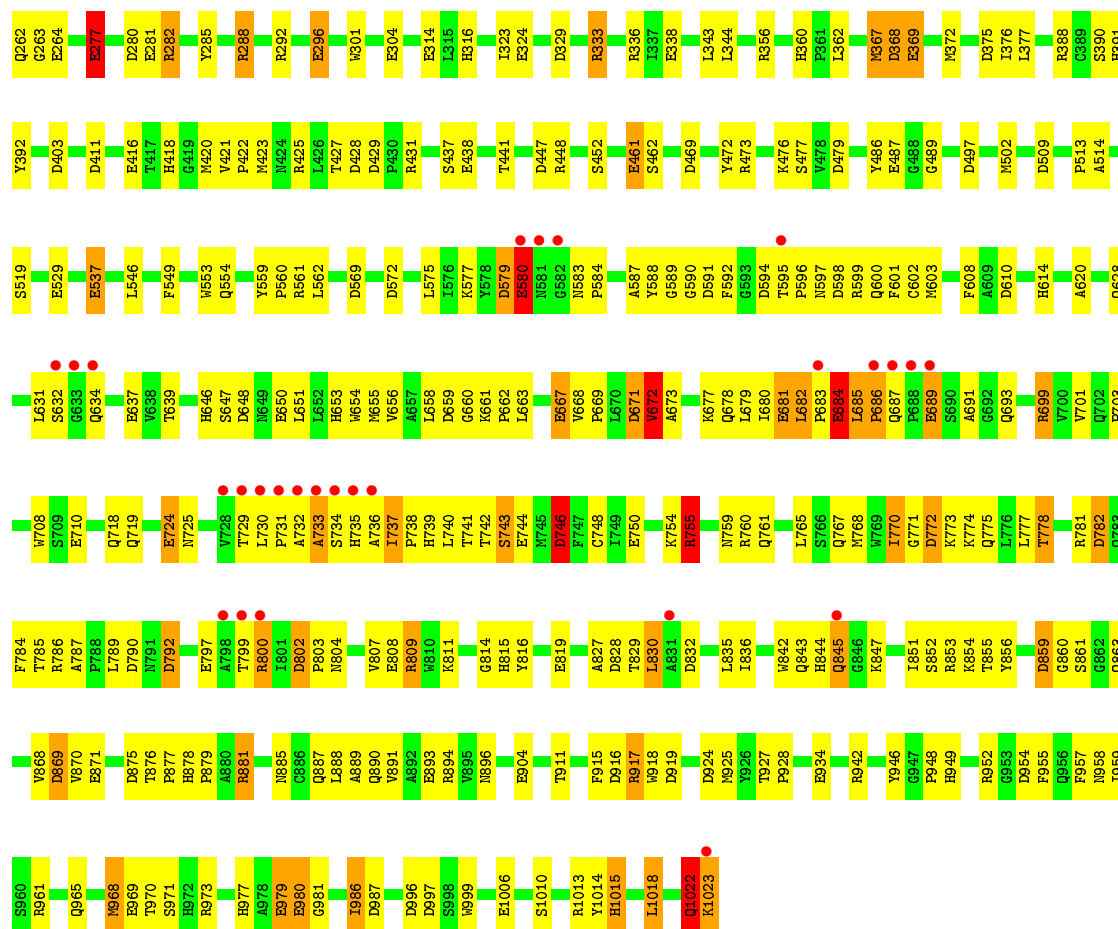
• Molecule 1: Beta-Galactosidase





Chain D: 





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	149.69Å 168.40Å 200.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	17.00 – 2.10 17.07 – 2.10	Depositor EDS
% Data completeness (in resolution range)	86.2 (17.00-2.10) 84.6 (17.07-2.10)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.37 (at 2.11Å)	Xtrriage
Refinement program	TNT	Depositor
R, $R_{free}$	0.163 , 0.270 0.162 , 0.258	Depositor DCC
$R_{free}$ test set	3547 reflections (1.43%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.9	Xtrriage
Anisotropy	0.743	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 118.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	35881	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 38.93 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 3.4238e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: NA, GTZ, DMS, MG

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	1.15	49/8367 (0.6%)	1.60	139/11415 (1.2%)
1	B	1.14	48/8367 (0.6%)	1.66	169/11415 (1.5%)
1	C	1.15	56/8367 (0.7%)	1.60	144/11415 (1.3%)
1	D	1.15	48/8367 (0.6%)	1.62	150/11415 (1.3%)
All	All	1.15	201/33468 (0.6%)	1.62	602/45660 (1.3%)

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	B	1	0
1	D	1	0
All	All	2	0

The worst 5 of 201 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	170	GLU	CD-OE2	10.20	1.36	1.25
1	B	461	GLU	CD-OE2	9.77	1.36	1.25
1	A	136	GLU	CD-OE2	9.02	1.35	1.25
1	A	249	GLU	CD-OE2	8.92	1.35	1.25
1	C	1006	GLU	CD-OE2	8.85	1.35	1.25

The worst 5 of 602 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	687	GLN	C-N-CD	-17.57	81.95	120.60
1	D	336	ARG	NE-CZ-NH1	14.18	127.39	120.30

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	95	TYR	CB-CG-CD2	-12.66	113.41	121.00
1	C	961	ARG	NE-CZ-NH1	11.97	126.29	120.30
1	B	144	ASP	CB-CG-OD2	-11.94	107.56	118.30

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	B	855	THR	CB
1	D	128	ASN	CA

There are no planarity outliers.

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	8125	0	7716	279	0
1	B	8125	0	7716	350	0
1	C	8125	0	7716	311	0
1	D	8125	0	7716	282	0
2	A	14	0	10	2	0
2	B	14	0	9	1	0
2	C	14	0	9	0	0
2	D	14	0	9	3	0
3	A	4	0	0	0	0
3	B	3	0	0	0	0
3	C	3	0	0	0	0
3	D	3	0	0	0	0
4	A	3	0	0	0	0
4	B	3	0	0	0	0
4	C	2	0	0	0	0
4	D	3	0	0	0	0
5	A	76	0	114	11	0
5	B	68	0	102	13	0
5	C	80	0	120	15	0
5	D	72	0	108	10	0
6	A	775	0	0	30	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	B	747	0	0	29	0
6	C	719	0	0	26	0
6	D	764	0	0	28	0
All	All	35881	0	31345	1198	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 19.

The worst 5 of 1198 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:804:ASN:HD22	1:D:809:ARG:NH2	1.40	1.17
1:B:685:LEU:HD23	1:B:686:PRO:HD2	1.17	1.16
1:C:800:ARG:HA	1:C:800:ARG:HE	1.08	1.12
1:C:761:GLN:HE21	1:C:761:GLN:N	1.48	1.11
1:B:251:ARG:HD2	5:B:8416:DMS:H22	1.36	1.07

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	1009/1023 (99%)	956 (95%)	52 (5%)	1 (0%)	51	54
1	B	1009/1023 (99%)	947 (94%)	55 (6%)	7 (1%)	22	18
1	C	1009/1023 (99%)	954 (94%)	50 (5%)	5 (0%)	29	26
1	D	1009/1023 (99%)	963 (95%)	42 (4%)	4 (0%)	34	32
All	All	4036/4092 (99%)	3820 (95%)	199 (5%)	17 (0%)	34	32

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	733	ALA
1	B	688	PRO
1	B	689	GLU
1	B	731	PRO
1	B	733	ALA

### 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	864/875 (99%)	808 (94%)	56 (6%)	17	14
1	B	864/875 (99%)	797 (92%)	67 (8%)	12	9
1	C	864/875 (99%)	812 (94%)	52 (6%)	19	16
1	D	864/875 (99%)	821 (95%)	43 (5%)	24	23
All	All	3456/3500 (99%)	3238 (94%)	218 (6%)	18	15

5 of 218 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	746	ASP
1	B	1023	LYS
1	D	737	ILE
1	B	757	GLN
1	B	828	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 68 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	718	GLN
1	C	102	ASN
1	D	761	GLN
1	B	757	GLN
1	B	844	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 102 ligands modelled in this entry, 24 are monoatomic - leaving 78 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$
5	DMS	B	8402	-	3,3,3	1.05	0	3,3,3	0.08	0
5	DMS	C	8402	-	3,3,3	1.46	0	3,3,3	0.33	0
5	DMS	A	8407	-	3,3,3	1.21	1 (33%)	3,3,3	0.25	0
2	GTZ	D	2001	4	13,15,15	1.45	2 (15%)	12,22,22	2.00	3 (25%)
5	DMS	A	8404	-	3,3,3	1.36	0	3,3,3	0.29	0
5	DMS	A	8421	-	3,3,3	0.97	0	3,3,3	0.44	0
5	DMS	B	8405	-	3,3,3	1.78	1 (33%)	3,3,3	0.22	0
5	DMS	D	8401	-	3,3,3	1.06	0	3,3,3	0.29	0
5	DMS	C	8409	-	3,3,3	0.62	0	3,3,3	0.41	0
5	DMS	D	8405	-	3,3,3	1.34	1 (33%)	3,3,3	0.79	0
5	DMS	C	8412	-	3,3,3	1.19	0	3,3,3	0.66	0
5	DMS	C	8421	-	3,3,3	0.62	0	3,3,3	0.53	0
5	DMS	C	8401	-	3,3,3	1.66	1 (33%)	3,3,3	0.99	0
5	DMS	D	8404	-	3,3,3	0.58	0	3,3,3	0.22	0
5	DMS	C	8423	-	3,3,3	0.66	0	3,3,3	0.19	0
5	DMS	C	8407	-	3,3,3	0.94	0	3,3,3	0.31	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	DMS	B	8407	-	3,3,3	1.44	1 (33%)	3,3,3	0.68	0
5	DMS	D	8407	-	3,3,3	0.83	0	3,3,3	0.08	0
5	DMS	C	8403	-	3,3,3	1.27	0	3,3,3	1.03	0
5	DMS	C	8417	-	3,3,3	0.73	0	3,3,3	0.15	0
5	DMS	A	8425	3	3,3,3	1.63	1 (33%)	3,3,3	0.08	0
5	DMS	A	8410	-	3,3,3	0.43	0	3,3,3	0.51	0
5	DMS	C	8425	3	3,3,3	0.83	0	3,3,3	0.06	0
5	DMS	C	8501	-	3,3,3	1.00	0	3,3,3	0.50	0
5	DMS	A	8405	-	3,3,3	2.10	1 (33%)	3,3,3	0.17	0
5	DMS	B	8401	-	3,3,3	0.55	0	3,3,3	0.97	0
5	DMS	A	8412	-	3,3,3	0.95	0	3,3,3	0.69	0
5	DMS	C	8405	-	3,3,3	2.05	1 (33%)	3,3,3	0.36	0
5	DMS	D	8413	-	3,3,3	0.98	0	3,3,3	0.19	0
5	DMS	D	8409	-	3,3,3	1.15	0	3,3,3	0.58	0
5	DMS	D	8705	-	3,3,3	1.00	0	3,3,3	0.23	0
5	DMS	A	8411	-	3,3,3	0.61	0	3,3,3	0.22	0
5	DMS	D	8408	-	3,3,3	1.44	1 (33%)	3,3,3	0.60	0
5	DMS	B	8425	3	3,3,3	1.08	0	3,3,3	0.32	0
5	DMS	A	8420	-	3,3,3	0.77	0	3,3,3	0.67	0
5	DMS	D	8703	-	3,3,3	1.50	0	3,3,3	0.30	0
5	DMS	B	8403	-	3,3,3	1.55	1 (33%)	3,3,3	0.61	0
5	DMS	B	8404	-	3,3,3	0.36	0	3,3,3	0.74	0
5	DMS	A	8414	-	3,3,3	1.12	0	3,3,3	0.55	0
5	DMS	A	8417	-	3,3,3	0.99	0	3,3,3	0.26	0
5	DMS	B	8412	-	3,3,3	1.04	0	3,3,3	0.69	0
5	DMS	A	8504	-	3,3,3	0.42	0	3,3,3	0.24	0
5	DMS	D	8402	-	3,3,3	1.08	0	3,3,3	1.18	1 (33%)
5	DMS	B	8416	-	3,3,3	0.77	0	3,3,3	0.36	0
5	DMS	D	8411	-	3,3,3	1.27	0	3,3,3	0.13	0
2	GTZ	A	2001	4	13,15,15	0.90	0	12,22,22	1.21	1 (8%)
5	DMS	C	8415	-	3,3,3	2.17	2 (66%)	3,3,3	0.40	0
5	DMS	A	8501	-	3,3,3	1.31	1 (33%)	3,3,3	0.43	0
5	DMS	D	8701	-	3,3,3	0.72	0	3,3,3	0.69	0
5	DMS	A	8402	-	3,3,3	0.90	0	3,3,3	0.92	0
2	GTZ	C	2001	4	13,15,15	0.89	1 (7%)	12,22,22	1.17	0
5	DMS	B	8601	-	3,3,3	0.82	0	3,3,3	0.46	0
5	DMS	D	8506	-	3,3,3	2.06	1 (33%)	3,3,3	0.42	0
5	DMS	C	8414	-	3,3,3	1.09	0	3,3,3	0.88	0
5	DMS	B	8414	-	3,3,3	0.59	0	3,3,3	0.54	0
5	DMS	D	8412	-	3,3,3	1.45	0	3,3,3	1.28	1 (33%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	DMS	D	8414	-	3,3,3	0.33	0	3,3,3	0.72	0
5	DMS	C	8601	-	3,3,3	1.17	0	3,3,3	0.84	0
5	DMS	A	8408	-	3,3,3	1.06	0	3,3,3	0.13	0
5	DMS	C	8404	-	3,3,3	2.12	2 (66%)	3,3,3	0.47	0
5	DMS	B	8423	-	3,3,3	0.83	0	3,3,3	0.16	0
5	DMS	D	8406	-	3,3,3	0.78	0	3,3,3	0.36	0
5	DMS	B	8408	-	3,3,3	1.01	0	3,3,3	0.66	0
5	DMS	B	8411	-	3,3,3	1.34	0	3,3,3	0.69	0
5	DMS	B	8409	-	3,3,3	0.80	0	3,3,3	1.09	0
2	GTZ	B	2001	4	13,15,15	0.94	1 (7%)	12,22,22	1.42	1 (8%)
5	DMS	D	8501	-	3,3,3	1.08	0	3,3,3	0.27	0
5	DMS	B	8417	-	3,3,3	1.48	1 (33%)	3,3,3	0.14	0
5	DMS	A	8401	-	3,3,3	0.91	0	3,3,3	0.43	0
5	DMS	C	8504	-	3,3,3	0.44	0	3,3,3	0.22	0
5	DMS	D	8403	-	3,3,3	1.66	1 (33%)	3,3,3	0.22	0
5	DMS	A	8406	-	3,3,3	0.80	0	3,3,3	0.25	0
5	DMS	A	8403	-	3,3,3	1.69	1 (33%)	3,3,3	1.12	0
5	DMS	B	8421	-	3,3,3	0.79	0	3,3,3	0.34	0
5	DMS	C	8419	-	3,3,3	0.42	0	3,3,3	0.12	0
5	DMS	C	8602	-	3,3,3	0.74	0	3,3,3	0.44	0
5	DMS	C	8411	-	3,3,3	1.60	0	3,3,3	0.96	0
5	DMS	A	8409	-	3,3,3	2.96	1 (33%)	3,3,3	0.35	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	GTZ	A	2001	4	-	1/2/22/22	0/1/2/2
2	GTZ	C	2001	4	-	1/2/22/22	0/1/2/2
2	GTZ	B	2001	4	-	1/2/22/22	0/1/2/2
2	GTZ	D	2001	4	-	1/2/22/22	0/1/2/2

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	8409	DMS	O-S	4.96	1.83	1.50
2	D	2001	GTZ	N1-NB	3.76	1.40	1.34
5	A	8405	DMS	O-S	3.41	1.73	1.50
5	D	8506	DMS	C2-S	3.04	1.98	1.75

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	8415	DMS	C1-S	2.85	1.97	1.75

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	2001	GTZ	O3-C3-C2	-4.19	99.63	108.98
2	D	2001	GTZ	O2-C2-C3	3.40	115.69	108.55
2	D	2001	GTZ	C3-C4-C5	-2.68	106.84	111.37
2	B	2001	GTZ	O4-C4-C3	-2.63	104.26	110.35
2	A	2001	GTZ	C3-C4-C5	-2.56	107.04	111.37

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	2001	GTZ	N5-C5-C6-O6
2	C	2001	GTZ	N5-C5-C6-O6
2	B	2001	GTZ	N5-C5-C6-O6
2	A	2001	GTZ	N5-C5-C6-O6

There are no ring outliers.

21 monomers are involved in 55 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	2001	GTZ	3	0
5	A	8404	DMS	5	0
5	D	8404	DMS	6	0
5	C	8423	DMS	4	0
5	C	8417	DMS	1	0
5	A	8425	DMS	2	0
5	A	8410	DMS	1	0
5	A	8420	DMS	1	0
5	B	8416	DMS	8	0
2	A	2001	GTZ	2	0
5	A	8402	DMS	1	0
5	B	8601	DMS	1	0
5	C	8414	DMS	3	0
5	C	8601	DMS	1	0
5	B	8408	DMS	1	0
2	B	2001	GTZ	1	0
5	B	8417	DMS	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	8504	DMS	5	0
5	D	8403	DMS	4	0
5	A	8406	DMS	1	0
5	C	8419	DMS	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	1011/1023 (98%)	-0.35	32 (3%)	47	54	16, 29, 58, 100	0
1	B	1011/1023 (98%)	-0.23	35 (3%)	44	50	18, 32, 62, 100	0
1	C	1011/1023 (98%)	-0.28	30 (2%)	50	56	16, 31, 62, 100	0
1	D	1011/1023 (98%)	-0.33	29 (2%)	51	57	16, 29, 60, 100	0
All	All	4044/4092 (98%)	-0.30	126 (3%)	49	55	16, 30, 61, 100	0

The worst 5 of 126 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	732	ALA	9.9
1	A	799	THR	7.8
1	C	732	ALA	7.3
1	C	731	PRO	7.2
1	B	733	ALA	7.0

### 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 carbohydrates in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. 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
5	DMS	B	8421	4/4	0.89	0.19	30,50,73,100	0
5	DMS	A	8417	4/4	0.90	0.34	34,100,100,100	0
5	DMS	B	8416	4/4	0.91	0.25	55,62,100,100	0
5	DMS	D	8413	4/4	0.91	0.31	82,100,100,100	0
5	DMS	A	8409	4/4	0.92	0.15	38,43,63,69	0
5	DMS	C	8415	4/4	0.94	0.10	35,39,40,90	0
5	DMS	D	8703	4/4	0.94	0.16	32,58,79,100	0
5	DMS	B	8601	4/4	0.94	0.14	39,46,100,100	0
5	DMS	A	8410	4/4	0.94	0.10	51,59,81,100	0
3	MG	A	3005	1/1	0.94	0.10	49,49,49,49	0
2	GTZ	B	2001	14/14	0.94	0.10	18,30,37,38	0
2	GTZ	A	2001	14/14	0.95	0.09	21,30,50,56	0
5	DMS	B	8414	4/4	0.95	0.13	54,57,64,83	0
2	GTZ	C	2001	14/14	0.95	0.09	18,27,32,41	0
5	DMS	C	8425	4/4	0.95	0.15	38,40,45,55	0
5	DMS	A	8407	4/4	0.95	0.12	32,53,72,100	0
5	DMS	B	8407	4/4	0.95	0.10	28,31,59,100	0
5	DMS	C	8409	4/4	0.95	0.13	34,42,66,66	0
5	DMS	C	8414	4/4	0.95	0.14	26,70,72,100	0
5	DMS	C	8419	4/4	0.95	0.22	51,90,100,100	0
5	DMS	D	8407	4/4	0.95	0.12	45,65,73,80	0
5	DMS	A	8425	4/4	0.96	0.11	34,36,51,55	0
5	DMS	C	8504	4/4	0.96	0.12	43,73,100,100	0
5	DMS	C	8601	4/4	0.96	0.12	43,48,53,54	0
5	DMS	C	8407	4/4	0.96	0.13	42,64,100,100	0
5	DMS	A	8421	4/4	0.96	0.30	60,63,93,100	0
5	DMS	D	8506	4/4	0.96	0.10	47,60,62,100	0
5	DMS	D	8414	4/4	0.96	0.15	31,44,67,100	0
5	DMS	D	8412	4/4	0.96	0.10	20,40,45,59	0
5	DMS	C	8423	4/4	0.97	0.12	65,71,86,100	0
5	DMS	A	8411	4/4	0.97	0.11	45,50,77,100	0
4	NA	C	3101	1/1	0.97	0.05	28,28,28,28	0
3	MG	B	3002	1/1	0.97	0.09	27,27,27,27	0
3	MG	B	3001	1/1	0.97	0.08	31,31,31,31	0
5	DMS	C	8417	4/4	0.97	0.07	29,51,57,58	0
5	DMS	B	8423	4/4	0.97	0.10	50,60,68,100	0
5	DMS	A	8420	4/4	0.97	0.09	45,54,56,58	0
5	DMS	A	8404	4/4	0.97	0.09	28,36,46,57	0
2	GTZ	D	2001	14/14	0.97	0.07	13,23,36,46	0
5	DMS	A	8501	4/4	0.97	0.10	31,40,43,51	0
5	DMS	A	8414	4/4	0.97	0.11	56,58,63,100	0
5	DMS	C	8421	4/4	0.97	0.18	48,49,61,61	0
5	DMS	B	8409	4/4	0.97	0.13	35,37,43,66	0

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	DMS	B	8408	4/4	0.97	0.21	22,43,53,84	0
4	NA	D	3102	1/1	0.97	0.05	25,25,25,25	0
5	DMS	A	8406	4/4	0.97	0.12	28,55,72,77	0
5	DMS	A	8405	4/4	0.98	0.07	25,39,43,49	0
3	MG	A	3105	1/1	0.98	0.11	33,33,33,33	0
5	DMS	B	8404	4/4	0.98	0.06	33,35,51,63	0
3	MG	C	3002	1/1	0.98	0.06	27,27,27,27	0
5	DMS	D	8406	4/4	0.98	0.09	29,29,34,34	0
5	DMS	B	8425	4/4	0.98	0.15	49,53,54,100	0
5	DMS	B	8405	4/4	0.98	0.09	31,38,44,64	0
4	NA	A	3101	1/1	0.98	0.05	29,29,29,29	0
5	DMS	C	8412	4/4	0.98	0.09	35,44,48,66	0
3	MG	D	3005	1/1	0.98	0.02	39,39,39,39	0
4	NA	B	3103	1/1	0.98	0.06	41,41,41,41	0
5	DMS	C	8411	4/4	0.98	0.14	31,44,51,100	0
5	DMS	B	8417	4/4	0.98	0.09	40,42,65,92	0
3	MG	B	3105	1/1	0.98	0.06	36,36,36,36	0
5	DMS	D	8408	4/4	0.98	0.15	34,39,50,58	0
5	DMS	D	8404	4/4	0.98	0.09	32,62,64,77	0
3	MG	A	3002	1/1	0.98	0.03	28,28,28,28	0
4	NA	B	3101	1/1	0.98	0.05	35,35,35,35	0
5	DMS	D	8409	4/4	0.98	0.11	33,33,42,43	0
5	DMS	D	8705	4/4	0.98	0.11	43,50,68,69	0
5	DMS	C	8402	4/4	0.98	0.09	19,33,40,45	0
5	DMS	A	8408	4/4	0.98	0.07	27,56,63,80	0
5	DMS	B	8412	4/4	0.98	0.07	37,41,45,55	0
5	DMS	D	8701	4/4	0.98	0.08	22,26,33,100	0
5	DMS	C	8501	4/4	0.98	0.10	31,35,42,75	0
5	DMS	C	8403	4/4	0.98	0.19	25,28,38,39	0
5	DMS	C	8602	4/4	0.98	0.10	55,63,64,100	0
5	DMS	D	8403	4/4	0.98	0.11	20,40,42,48	0
4	NA	D	3103	1/1	0.98	0.07	45,45,45,45	0
3	MG	C	3001	1/1	0.99	0.04	25,25,25,25	0
5	DMS	D	8501	4/4	0.99	0.07	37,38,47,52	0
4	NA	A	3102	1/1	0.99	0.07	24,24,24,24	0
5	DMS	B	8411	4/4	0.99	0.07	36,38,42,100	0
3	MG	C	3105	1/1	0.99	0.22	41,41,41,41	0
4	NA	B	3102	1/1	0.99	0.05	24,24,24,24	0
5	DMS	D	8405	4/4	0.99	0.08	29,37,39,67	0
5	DMS	C	8401	4/4	0.99	0.07	25,27,29,31	0
3	MG	D	3002	1/1	0.99	0.09	27,27,27,27	0
5	DMS	A	8401	4/4	0.99	0.14	16,22,28,29	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	DMS	A	8504	4/4	0.99	0.13	23,50,53,92	0
3	MG	D	3001	1/1	0.99	0.04	27,27,27,27	0
4	NA	C	3102	1/1	0.99	0.05	22,22,22,22	0
5	DMS	C	8405	4/4	0.99	0.09	30,37,39,40	0
3	MG	A	3001	1/1	0.99	0.05	22,22,22,22	0
5	DMS	D	8411	4/4	0.99	0.07	29,41,62,100	0
4	NA	D	3101	1/1	0.99	0.08	28,28,28,28	0
5	DMS	B	8403	4/4	0.99	0.08	28,35,38,42	0
5	DMS	C	8404	4/4	0.99	0.07	31,31,34,41	0
5	DMS	D	8402	4/4	0.99	0.07	21,30,39,39	0
4	NA	A	3103	1/1	0.99	0.08	37,37,37,37	0
5	DMS	A	8402	4/4	0.99	0.06	21,35,46,52	0
5	DMS	B	8402	4/4	0.99	0.07	21,27,35,38	0
5	DMS	B	8401	4/4	0.99	0.12	29,31,38,46	0
5	DMS	A	8412	4/4	0.99	0.08	30,45,46,100	0
5	DMS	D	8401	4/4	0.99	0.09	21,25,27,28	0
5	DMS	A	8403	4/4	0.99	0.10	29,35,42,56	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.