wwPDB X-ray Structure Validation Summary Report (i)

Aug 6, 2023 – 03:51 PM EDT

PDB ID	:	1JZ5
Title	:	E. COLI (lacZ) BETA-GALACTOSIDASE IN COMPLEX WITH D-GALCT OPYRANOSYL-1-ON
Authors	:	Juers, D.H.; Matthews, B.W.
Deposited on	:	2001-09-13
Resolution	:	1.80 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

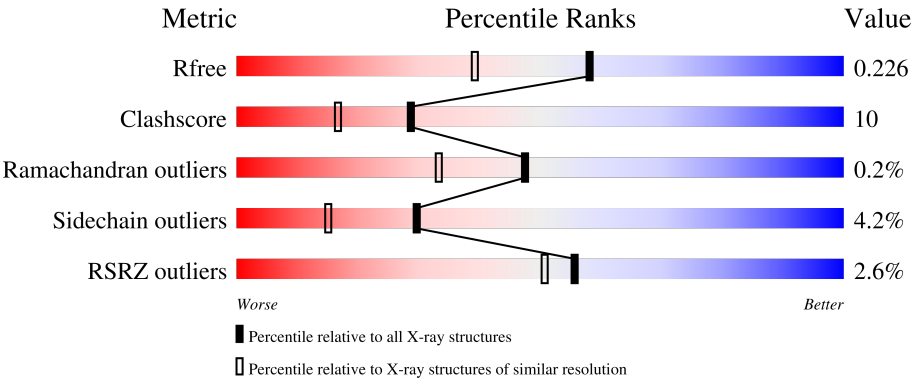
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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 >=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 <=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	2% 68% 25% 5% •
1	B	1023	2% 68% 26% 5% •
1	C	1023	3% 70% 24% 5% ••
1	D	1023	3% 70% 24% 5% •

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
5	DMS	B	8414	-	-	X	-
5	DMS	B	8504	-	-	X	-
5	DMS	D	8419	-	-	X	-
5	DMS	D	8508	-	-	X	-
5	DMS	D	8701	-	X	-	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 36964 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	98	11	0
			8205	5185	1455	1527	38			
1	B	1011	Total	C	N	O	S	64	11	0
			8205	5185	1455	1527	38			
1	C	1011	Total	C	N	O	S	49	11	0
			8205	5185	1455	1527	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:

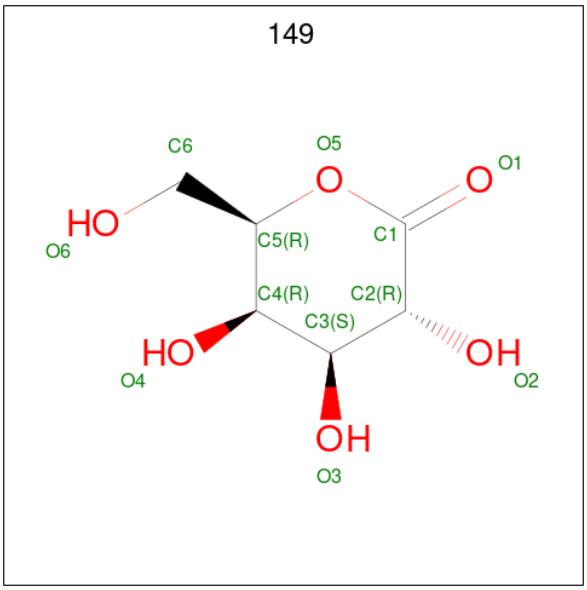
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	THR	cloning artifact	? P00722
A	2	SER	MET	cloning artifact	? P00722
A	3	HIS	ILE	cloning artifact	? P00722
A	4	MET	THR	cloning artifact	? P00722
A	5	LEU	ASP	cloning artifact	? P00722
A	6	GLU	SER	cloning artifact	? P00722
A	7	ASP	LEU	cloning artifact	? 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 D-galactonolactone (three-letter code: 149) (formula: $C_6H_{10}O_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 12 6 6	0	0
2	B	1	Total C O 12 6 6	0	0
2	C	1	Total C O 12 6 6	0	0
2	D	1	Total C O 12 6 6	0	0

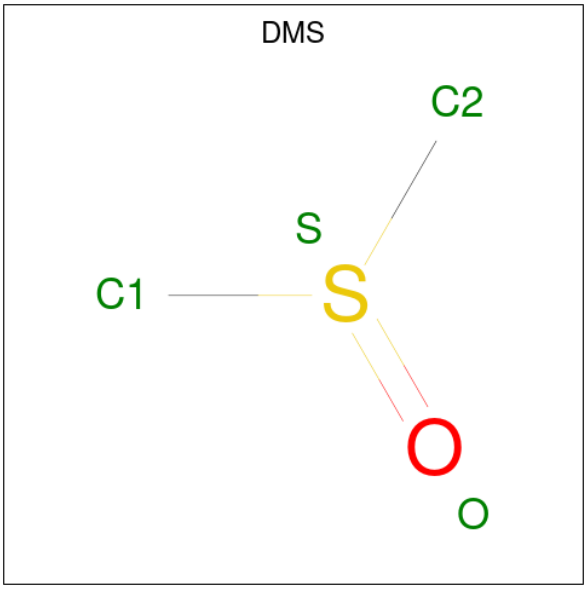
- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 5 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C_2H_6OS).



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 C O S 4 2 1 1	0	0
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
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
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
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
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
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
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
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
5	A	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	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	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	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	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	C	1	Total C O S 4 2 1 1	0	0
5	D	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	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	1	Total C O S 4 2 1 1	0	0
5	D	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	D	1	Total	C	O	S	0	0
			4	2	1	1		

- Molecule 6 is water.

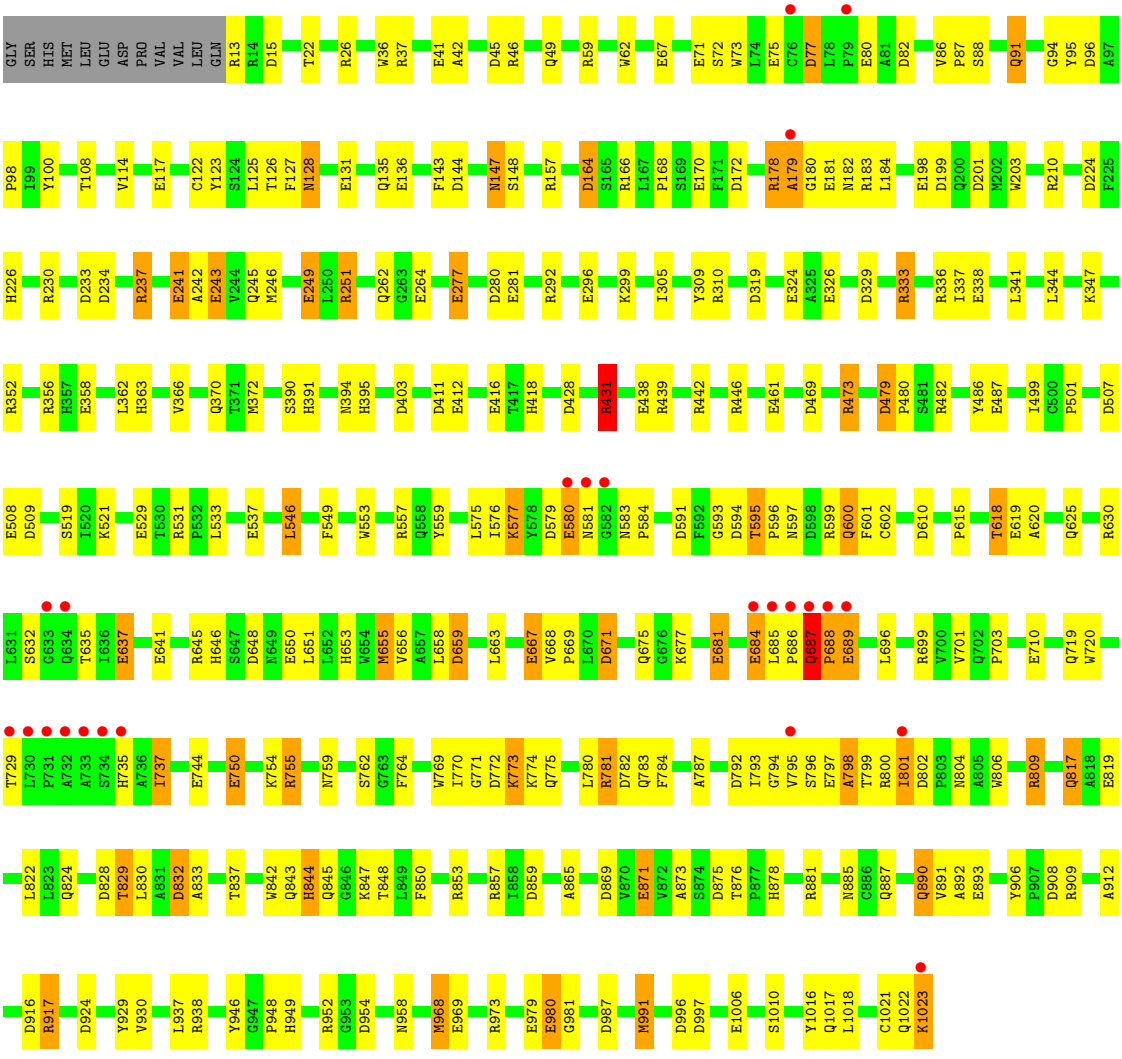
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	953	Total	O	0	0
			953	953		
6	B	937	Total	O	0	0
			937	937		
6	C	952	Total	O	0	0
			952	952		
6	D	959	Total	O	0	0
			959	959		

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.

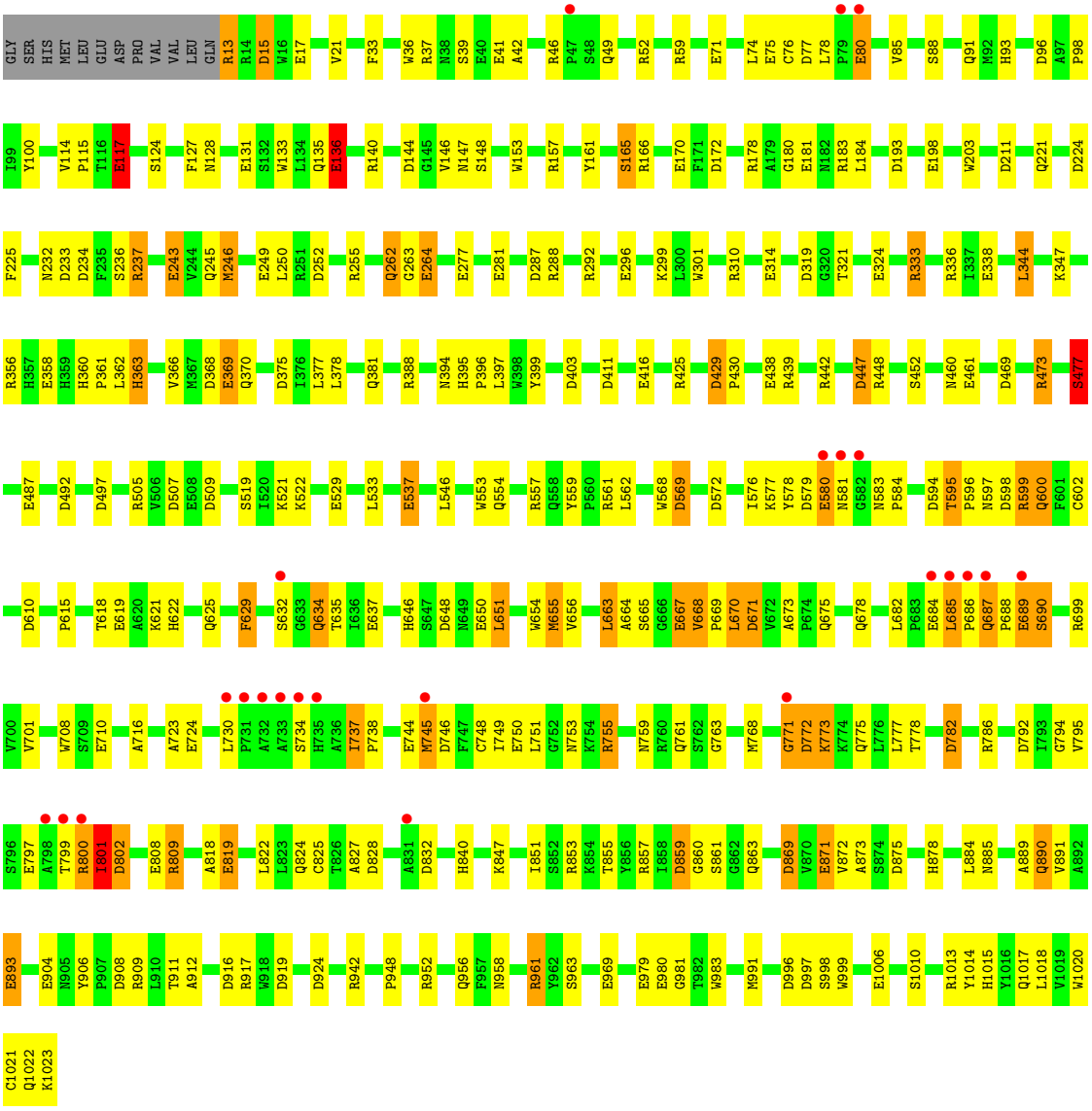
• Molecule 1: Beta-Galactosidase

Chain A:



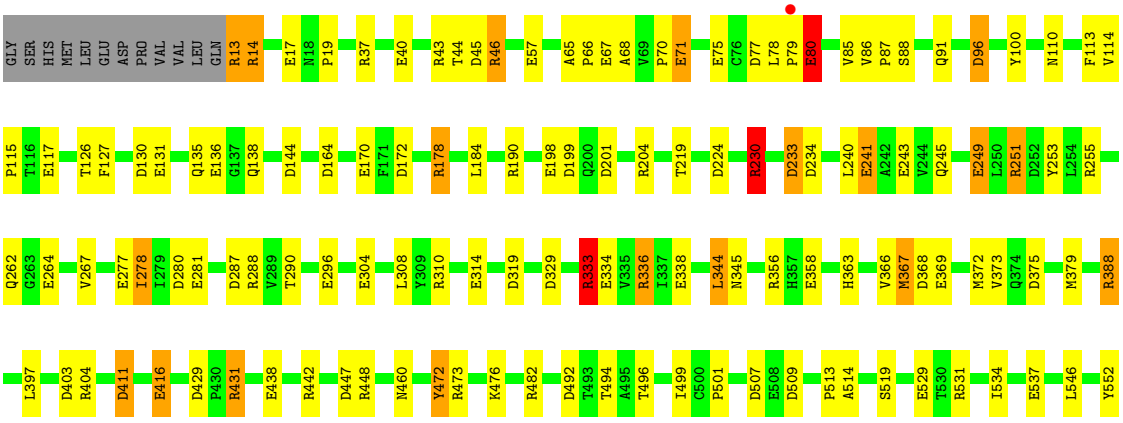
• Molecule 1: Beta-Galactosidase

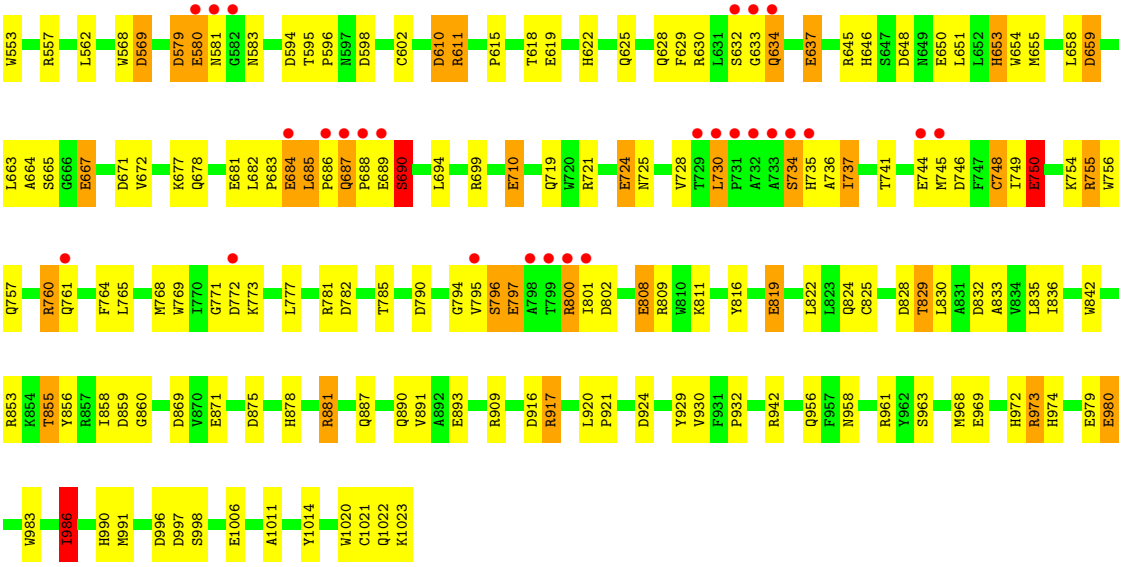
Chain B:



• Molecule 1: Beta-Galactosidase

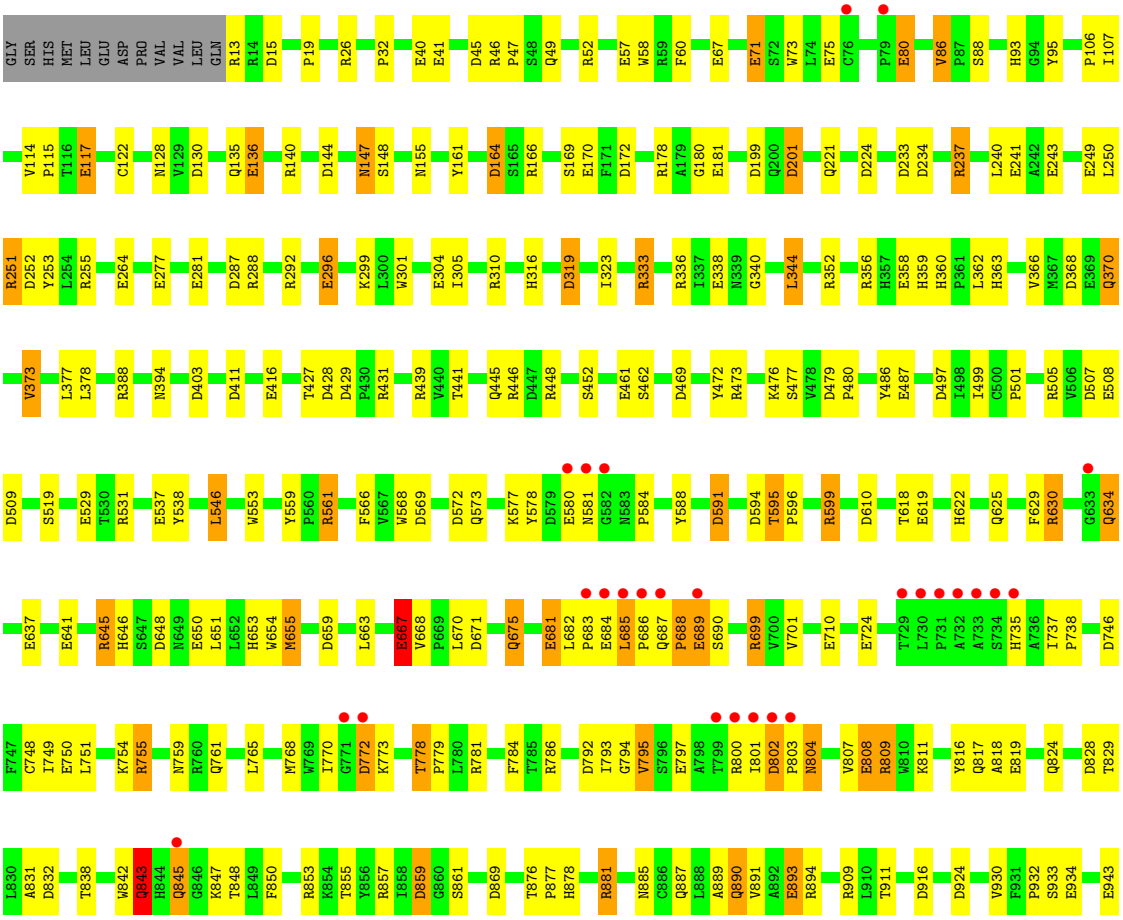
Chain C:

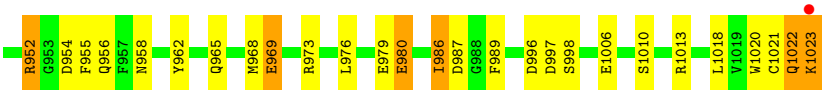




• Molecule 1: Beta-Galactosidase

Chain D:




R952
G953
D954
F955
Q956
F957
N958
Y962
Q965
M968
E969
R973
L976
E979
E980
I986
D987
G988
F989
D996
D997
S998
E1006
S1010
R1013
L1018
V1019
W1020
C1021
Q1022
K1023

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	149.69Å 168.02Å 201.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	17.00 – 1.80 17.00 – 1.80	Depositor EDS
% Data completeness (in resolution range)	90.4 (17.00-1.80) 90.0 (17.00-1.80)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.93 (at 1.80Å)	Xtriage
Refinement program	TNT	Depositor
R, R_{free}	0.160 , 0.231 0.163 , 0.226	Depositor DCC
R_{free} test set	6053 reflections (1.44%)	wwPDB-VP
Wilson B-factor (Å^2)	19.0	Xtriage
Anisotropy	0.257	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 106.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	36964	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 37.04 % 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 4.5593e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: MG, DMS, NA, 149

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	1.13	46/8448 (0.5%)	1.66	159/11526 (1.4%)
1	B	1.14	46/8448 (0.5%)	1.70	177/11526 (1.5%)
1	C	1.12	45/8448 (0.5%)	1.64	148/11526 (1.3%)
1	D	1.14	43/8367 (0.5%)	1.66	164/11415 (1.4%)
All	All	1.13	180/33711 (0.5%)	1.67	648/45993 (1.4%)

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	D	1	0

The worst 5 of 180 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	537	GLU	CD-OE2	11.40	1.38	1.25
1	A	681	GLU	CD-OE2	10.55	1.37	1.25
1	A	710	GLU	CD-OE2	10.27	1.36	1.25
1	D	893	GLU	CD-OE2	9.79	1.36	1.25
1	D	537	GLU	CD-OE2	9.66	1.36	1.25

The worst 5 of 648 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	699	ARG	NE-CZ-NH1	21.57	131.08	120.30
1	B	699	ARG	NE-CZ-NH2	-16.36	112.12	120.30
1	B	442	ARG	NE-CZ-NH2	-16.07	112.27	120.30
1	C	809	ARG	NE-CZ-NH1	15.09	127.85	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	442	ARG	NE-CZ-NH1	13.86	127.23	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	D	634	GLN	CA

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	8205	0	7791	158	0
1	B	8205	0	7791	160	0
1	C	8205	0	7791	159	0
1	D	8125	0	7716	142	0
2	A	12	0	9	0	0
2	B	12	0	9	1	0
2	C	12	0	9	0	0
2	D	12	0	9	0	0
3	A	3	0	0	0	0
3	B	3	0	0	0	0
3	C	2	0	0	0	0
3	D	3	0	0	0	0
4	A	4	0	0	0	0
4	B	4	0	0	0	0
4	C	4	0	0	0	0
4	D	4	0	0	0	0
5	A	80	0	120	4	0
5	B	88	0	132	10	0
5	C	88	0	132	7	0
5	D	92	0	138	18	0
6	A	953	0	0	22	0
6	B	937	0	0	25	1
6	C	952	0	0	15	0
6	D	959	0	0	23	1
All	All	36964	0	31647	623	1

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 10.

The worst 5 of 623 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:634:GLN:H	1:C:634:GLN:NE2	1.37	1.22
1:A:237:ARG:HH11	1:A:237:ARG:HB3	1.20	1.06
1:C:634:GLN:HE21	1:C:634:GLN:N	1.57	1.01
1:B:232:ASN:ND2	1:B:237:ARG:HG3	1.82	0.94
1:B:13:ARG:HG3	1:C:13:ARG:CZ	1.98	0.93

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:B:9538:HOH:O	6:D:9651:HOH:O[2_454]	2.09	0.11

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	1020/1023 (100%)	980 (96%)	35 (3%)	5 (0%)	29	15
1	B	1020/1023 (100%)	985 (97%)	34 (3%)	1 (0%)	51	36
1	C	1020/1023 (100%)	980 (96%)	38 (4%)	2 (0%)	47	33
1	D	1009/1023 (99%)	971 (96%)	35 (4%)	3 (0%)	41	27
All	All	4069/4092 (99%)	3916 (96%)	142 (4%)	11 (0%)	47	27

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	690	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	690	SER
1	D	688	PRO
1	A	688	PRO
1	A	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	A	873/875 (100%)	838 (96%)	35 (4%)	31	16
1	B	873/875 (100%)	833 (95%)	40 (5%)	27	13
1	C	873/875 (100%)	830 (95%)	43 (5%)	25	11
1	D	864/875 (99%)	828 (96%)	36 (4%)	30	15
All	All	3483/3500 (100%)	3329 (96%)	154 (4%)	30	14

5 of 154 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	956	GLN
1	D	801	ILE
1	D	71	GLU
1	D	599	ARG
1	D	956	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	878	HIS
1	D	878	HIS
1	D	135	GLN
1	D	624	GLN
1	D	977	HIS

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 118 ligands modelled in this entry, 27 are monoatomic - leaving 91 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	D	8410	-	3,3,3	1.71	0	3,3,3	0.51	0
5	DMS	C	8401	-	3,3,3	1.18	0	3,3,3	0.04	0
5	DMS	D	8705	-	3,3,3	0.89	0	3,3,3	0.19	0
5	DMS	A	8412	-	3,3,3	1.34	0	3,3,3	0.30	0
5	DMS	B	8425	4	3,3,3	2.75	1 (33%)	3,3,3	0.61	0
5	DMS	A	8408	-	3,3,3	0.38	0	3,3,3	0.44	0
5	DMS	D	8416	-	3,3,3	0.54	0	3,3,3	0.43	0
5	DMS	B	8402	-	3,3,3	1.36	1 (33%)	3,3,3	0.20	0
5	DMS	B	8404	-	3,3,3	0.78	0	3,3,3	0.07	0
5	DMS	D	8408	-	3,3,3	1.55	1 (33%)	3,3,3	0.38	0
5	DMS	D	8402	-	3,3,3	0.71	0	3,3,3	0.67	0
5	DMS	B	8421	-	3,3,3	0.73	0	3,3,3	0.05	0
5	DMS	D	8701	-	3,3,3	2.22	3 (100%)	3,3,3	0.63	0
5	DMS	B	8502	-	3,3,3	1.46	1 (33%)	3,3,3	0.87	0
5	DMS	D	8423	-	3,3,3	0.64	0	3,3,3	0.25	0
5	DMS	B	8405	-	3,3,3	1.35	0	3,3,3	0.41	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	DMS	A	8403	-	3,3,3	1.91	2 (66%)	3,3,3	1.02	0
5	DMS	C	8402	-	3,3,3	2.16	1 (33%)	3,3,3	0.47	0
5	DMS	B	8417	-	3,3,3	1.19	0	3,3,3	0.60	0
5	DMS	A	8414	-	3,3,3	1.31	0	3,3,3	0.30	0
5	DMS	C	8409	-	3,3,3	0.68	0	3,3,3	0.72	0
5	DMS	A	8401	-	3,3,3	0.80	0	3,3,3	0.48	0
5	DMS	D	8703	-	3,3,3	1.78	1 (33%)	3,3,3	0.26	0
5	DMS	D	8501	-	3,3,3	0.63	0	3,3,3	0.38	0
5	DMS	B	8412	-	3,3,3	1.25	1 (33%)	3,3,3	0.51	0
5	DMS	C	8407	-	3,3,3	1.62	1 (33%)	3,3,3	0.53	0
5	DMS	A	8602	-	3,3,3	0.43	0	3,3,3	0.49	0
5	DMS	D	8409	-	3,3,3	1.50	0	3,3,3	0.48	0
5	DMS	D	8411	-	3,3,3	0.72	0	3,3,3	0.37	0
5	DMS	B	8601	-	3,3,3	0.86	0	3,3,3	0.61	0
5	DMS	B	8408	-	3,3,3	0.29	0	3,3,3	0.49	0
5	DMS	A	8402	-	3,3,3	1.49	0	3,3,3	0.63	0
5	DMS	B	8423	-	3,3,3	0.82	0	3,3,3	0.32	0
5	DMS	C	8601	-	3,3,3	1.13	0	3,3,3	1.06	0
5	DMS	D	8403	-	3,3,3	1.75	1 (33%)	3,3,3	0.32	0
5	DMS	A	8416	-	3,3,3	0.83	0	3,3,3	0.87	0
5	DMS	B	8410	-	3,3,3	1.24	0	3,3,3	0.30	0
5	DMS	D	8419	-	3,3,3	1.76	1 (33%)	3,3,3	0.44	0
2	149	B	2001	4	12,12,12	2.19	1 (8%)	15,17,17	1.11	0
5	DMS	C	8425	4	3,3,3	0.96	0	3,3,3	0.52	0
5	DMS	C	8416	-	3,3,3	1.27	1 (33%)	3,3,3	0.34	0
5	DMS	A	8406	-	3,3,3	1.01	0	3,3,3	0.39	0
5	DMS	C	8415	-	3,3,3	1.81	1 (33%)	3,3,3	0.30	0
5	DMS	A	8421	-	3,3,3	0.91	0	3,3,3	0.83	0
5	DMS	A	8419	-	3,3,3	1.09	0	3,3,3	0.18	0
5	DMS	C	8403	-	3,3,3	1.63	0	3,3,3	0.56	0
2	149	D	2001	4	12,12,12	1.94	3 (25%)	15,17,17	1.80	4 (26%)
5	DMS	C	8421	-	3,3,3	0.66	0	3,3,3	0.13	0
5	DMS	B	8409	-	3,3,3	1.75	1 (33%)	3,3,3	1.13	0
5	DMS	A	8404	-	3,3,3	1.86	1 (33%)	3,3,3	0.45	0
5	DMS	C	8404	-	3,3,3	1.70	1 (33%)	3,3,3	0.60	0
5	DMS	B	8508	-	3,3,3	1.43	0	3,3,3	0.20	0
5	DMS	B	8401	-	3,3,3	0.88	0	3,3,3	0.55	0
5	DMS	A	8409	-	3,3,3	2.75	1 (33%)	3,3,3	0.22	0
5	DMS	D	8412	-	3,3,3	1.01	0	3,3,3	0.43	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	DMS	D	8414	-	3,3,3	0.67	0	3,3,3	0.31	0
5	DMS	B	8420	-	3,3,3	0.59	0	3,3,3	0.84	0
5	DMS	A	8405	-	3,3,3	0.59	0	3,3,3	0.63	0
5	DMS	C	8414	-	3,3,3	1.87	1 (33%)	3,3,3	0.66	0
5	DMS	D	8401	-	3,3,3	0.91	0	3,3,3	0.29	0
5	DMS	C	8423	-	3,3,3	0.94	0	3,3,3	0.38	0
5	DMS	C	8405	-	3,3,3	1.55	1 (33%)	3,3,3	0.51	0
2	149	C	2001	4	12,12,12	1.53	2 (16%)	15,17,17	1.80	3 (20%)
5	DMS	C	8417	-	3,3,3	0.72	0	3,3,3	0.30	0
5	DMS	C	8412	-	3,3,3	0.92	0	3,3,3	0.50	0
5	DMS	C	8411	-	3,3,3	1.13	0	3,3,3	0.10	0
5	DMS	C	8410	-	3,3,3	0.94	0	3,3,3	0.39	0
5	DMS	D	8406	-	3,3,3	0.70	0	3,3,3	0.72	0
5	DMS	A	8504	-	3,3,3	0.68	0	3,3,3	0.40	0
5	DMS	A	8425	4	3,3,3	1.56	1 (33%)	3,3,3	0.18	0
5	DMS	B	8403	-	3,3,3	1.69	1 (33%)	3,3,3	0.25	0
5	DMS	D	8417	-	3,3,3	0.81	0	3,3,3	0.14	0
5	DMS	B	8416	-	3,3,3	0.81	0	3,3,3	0.25	0
5	DMS	D	8404	-	3,3,3	1.02	0	3,3,3	0.19	0
5	DMS	C	8408	-	3,3,3	0.98	0	3,3,3	0.41	0
5	DMS	D	8503	-	3,3,3	0.51	0	3,3,3	0.29	0
5	DMS	A	8407	-	3,3,3	1.38	0	3,3,3	0.34	0
5	DMS	D	8421	-	3,3,3	0.79	0	3,3,3	0.42	0
5	DMS	C	8428	-	3,3,3	0.81	0	3,3,3	0.29	0
2	149	A	2001	4	12,12,12	2.03	2 (16%)	15,17,17	1.82	3 (20%)
5	DMS	D	8508	-	3,3,3	1.59	1 (33%)	3,3,3	0.11	0
5	DMS	B	8504	-	3,3,3	1.24	1 (33%)	3,3,3	0.33	0
5	DMS	A	8411	-	3,3,3	1.07	0	3,3,3	0.80	0
5	DMS	C	8420	-	3,3,3	1.23	1 (33%)	3,3,3	0.70	0
5	DMS	B	8406	-	3,3,3	0.78	0	3,3,3	0.32	0
5	DMS	A	8503	-	3,3,3	1.35	0	3,3,3	0.40	0
5	DMS	B	8411	-	3,3,3	0.29	0	3,3,3	0.58	0
5	DMS	B	8414	-	3,3,3	0.63	0	3,3,3	0.30	0
5	DMS	D	8405	-	3,3,3	0.38	0	3,3,3	1.24	0
5	DMS	A	8501	-	3,3,3	1.03	0	3,3,3	0.33	0
5	DMS	C	8501	-	3,3,3	0.85	0	3,3,3	0.22	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	149	A	2001	4	-	1/2/22/22	0/1/1/1
2	149	D	2001	4	-	2/2/22/22	0/1/1/1
2	149	C	2001	4	-	1/2/22/22	0/1/1/1
2	149	B	2001	4	-	1/2/22/22	0/1/1/1

The worst 5 of 36 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	2001	149	O5-C1	6.52	1.44	1.34
2	D	2001	149	O5-C1	5.15	1.42	1.34
2	A	2001	149	O5-C1	4.92	1.42	1.34
5	B	8425	DMS	O-S	4.67	1.81	1.50
5	A	8409	DMS	O-S	4.32	1.79	1.50

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	2001	149	C3-C4-C5	-4.77	101.73	110.24
2	C	2001	149	O4-C4-C5	-4.09	99.14	109.30
2	D	2001	149	C3-C4-C5	-3.85	103.38	110.24
2	D	2001	149	O4-C4-C5	3.27	117.43	109.30
2	C	2001	149	O5-C5-C4	2.78	114.67	109.73

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	2001	149	O5-C5-C6-O6
2	C	2001	149	O5-C5-C6-O6
2	A	2001	149	O5-C5-C6-O6
2	B	2001	149	O5-C5-C6-O6
2	D	2001	149	C4-C5-C6-O6

There are no ring outliers.

18 monomers are involved in 40 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	8402	DMS	1	0
5	D	8423	DMS	3	0
5	C	8402	DMS	3	0
5	D	8703	DMS	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	8419	DMS	5	0
2	B	2001	149	1	0
5	D	8412	DMS	2	0
5	B	8420	DMS	1	0
5	A	8405	DMS	1	0
5	C	8417	DMS	1	0
5	C	8410	DMS	1	0
5	A	8425	DMS	3	0
5	D	8404	DMS	2	0
5	D	8503	DMS	1	0
5	D	8508	DMS	4	0
5	B	8504	DMS	4	0
5	C	8420	DMS	2	0
5	B	8414	DMS	4	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^{th} 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^2)	Q<0.9
1	A	1009/1023 (98%)	-0.44	24 (2%)	59	54	11, 23, 51, 90	9 (0%)
1	B	1011/1023 (98%)	-0.41	24 (2%)	59	54	13, 23, 49, 90	8 (0%)
1	C	1011/1023 (98%)	-0.43	28 (2%)	53	47	14, 22, 52, 89	6 (0%)
1	D	1011/1023 (98%)	-0.42	28 (2%)	53	47	13, 23, 52, 94	0
All	All	4042/4092 (98%)	-0.42	104 (2%)	56	51	11, 23, 52, 94	23 (0%)

The worst 5 of 104 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	735	HIS	7.9
1	C	732	ALA	7.5
1	A	686	PRO	7.4
1	D	732	ALA	7.2
1	D	735	HIS	7.1

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} 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^2)	Q<0.9
5	DMS	A	8407	4/4	0.79	0.18	32,46,85,100	0
3	MG	B	3003	1/1	0.80	0.14	53,53,53,53	0
5	DMS	C	8416	4/4	0.85	0.24	47,66,73,96	0
5	DMS	D	8703	4/4	0.88	0.23	28,62,66,100	0
5	DMS	C	8407	4/4	0.89	0.15	29,37,51,72	0
5	DMS	B	8508	4/4	0.90	0.17	41,54,84,84	0
5	DMS	D	8423	4/4	0.92	0.15	48,53,78,100	0
5	DMS	B	8420	4/4	0.92	0.11	35,42,50,100	0
5	DMS	D	8503	4/4	0.93	0.23	36,66,84,90	0
5	DMS	D	8417	4/4	0.93	0.14	30,32,35,100	0
5	DMS	A	8404	4/4	0.94	0.10	26,32,43,45	0
5	DMS	C	8420	4/4	0.94	0.14	48,49,57,96	0
5	DMS	A	8409	4/4	0.95	0.12	30,34,43,51	0
5	DMS	A	8425	4/4	0.95	0.17	34,43,49,89	0
5	DMS	A	8503	4/4	0.95	0.22	35,51,100,100	0
5	DMS	C	8417	4/4	0.95	0.10	33,36,57,63	0
5	DMS	B	8410	4/4	0.95	0.11	43,50,63,100	0
5	DMS	C	8423	4/4	0.95	0.13	35,64,64,100	0
5	DMS	C	8425	4/4	0.95	0.19	37,57,68,100	0
5	DMS	D	8410	4/4	0.95	0.17	47,58,100,100	0
5	DMS	A	8406	4/4	0.95	0.10	26,37,56,62	0
5	DMS	D	8419	4/4	0.95	0.11	42,45,63,78	0
5	DMS	B	8423	4/4	0.95	0.18	42,61,69,100	0
5	DMS	B	8425	4/4	0.95	0.14	24,31,42,49	0
5	DMS	B	8502	4/4	0.95	0.12	36,42,48,51	0
5	DMS	B	8417	4/4	0.96	0.11	31,32,68,100	0
5	DMS	A	8419	4/4	0.96	0.11	40,49,58,79	0
5	DMS	A	8421	4/4	0.96	0.26	61,66,87,100	0
3	MG	D	3005	1/1	0.96	0.08	38,38,38,38	0
4	NA	D	3104	1/1	0.96	0.06	37,37,37,37	0
5	DMS	B	8409	4/4	0.96	0.13	29,33,37,43	0
5	DMS	D	8414	4/4	0.96	0.15	36,58,70,100	0
5	DMS	D	8416	4/4	0.96	0.17	32,42,47,79	0
5	DMS	B	8601	4/4	0.96	0.12	40,47,61,100	0
5	DMS	A	8414	4/4	0.96	0.10	35,40,65,100	0
5	DMS	C	8408	4/4	0.96	0.08	26,37,46,52	0
5	DMS	C	8409	4/4	0.96	0.14	30,40,52,64	0
5	DMS	C	8415	4/4	0.96	0.13	27,32,55,59	0
5	DMS	D	8705	4/4	0.96	0.13	33,46,55,58	0
5	DMS	B	8504	4/4	0.97	0.07	26,46,64,82	0
5	DMS	A	8412	4/4	0.97	0.09	39,43,43,100	0
5	DMS	D	8404	4/4	0.97	0.10	39,39,54,63	0
5	DMS	D	8406	4/4	0.97	0.13	29,34,37,57	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	DMS	D	8409	4/4	0.97	0.10	31,31,42,47	0
4	NA	D	3103	1/1	0.97	0.10	38,38,38,38	0
5	DMS	B	8414	4/4	0.97	0.20	39,55,61,100	0
5	DMS	A	8501	4/4	0.97	0.08	23,33,41,43	0
3	MG	A	3005	1/1	0.97	0.06	45,45,45,45	0
5	DMS	C	8410	4/4	0.97	0.12	39,62,82,100	0
5	DMS	B	8421	4/4	0.97	0.11	44,57,59,87	0
5	DMS	D	8501	4/4	0.97	0.07	40,45,46,55	0
5	DMS	A	8602	4/4	0.97	0.24	50,51,100,100	0
5	DMS	D	8508	4/4	0.97	0.11	47,58,68,100	0
5	DMS	B	8404	4/4	0.97	0.06	26,28,36,40	0
5	DMS	B	8408	4/4	0.97	0.14	35,38,40,45	0
5	DMS	B	8412	4/4	0.98	0.06	29,30,35,38	0
4	NA	C	3104	1/1	0.98	0.07	31,31,31,31	0
5	DMS	B	8416	4/4	0.98	0.14	38,42,47,50	0
5	DMS	C	8421	4/4	0.98	0.14	41,65,100,100	0
5	DMS	A	8416	4/4	0.98	0.26	23,69,70,100	0
2	149	D	2001	12/12	0.98	0.06	12,16,21,25	0
5	DMS	C	8428	4/4	0.98	0.26	10,19,20,29	4
5	DMS	C	8501	4/4	0.98	0.08	26,39,40,53	0
5	DMS	C	8601	4/4	0.98	0.13	37,45,50,57	0
5	DMS	D	8402	4/4	0.98	0.06	18,37,39,40	0
5	DMS	D	8403	4/4	0.98	0.11	26,38,40,52	0
2	149	A	2001	12/12	0.98	0.07	16,19,26,28	0
5	DMS	A	8403	4/4	0.98	0.08	24,28,34,35	0
5	DMS	D	8408	4/4	0.98	0.11	30,37,40,44	0
3	MG	B	3002	1/1	0.98	0.06	22,22,22,22	0
2	149	B	2001	12/12	0.98	0.08	16,20,24,33	0
5	DMS	D	8412	4/4	0.98	0.08	35,37,41,100	0
2	149	C	2001	12/12	0.98	0.07	17,20,26,26	0
5	DMS	B	8402	4/4	0.98	0.08	17,26,33,37	0
5	DMS	A	8408	4/4	0.98	0.09	26,39,45,100	0
5	DMS	C	8402	4/4	0.98	0.07	17,30,34,43	0
5	DMS	D	8421	4/4	0.98	0.17	44,49,60,100	0
5	DMS	C	8404	4/4	0.98	0.06	22,25,25,34	0
5	DMS	B	8405	4/4	0.98	0.10	35,35,36,41	0
5	DMS	B	8406	4/4	0.98	0.10	40,43,56,59	0
4	NA	A	3104	1/1	0.98	0.09	30,30,30,30	0
5	DMS	D	8701	4/4	0.98	0.08	21,25,32,50	0
5	DMS	A	8411	4/4	0.98	0.06	25,34,38,59	0
4	NA	B	3104	1/1	0.98	0.14	33,33,33,33	0
4	NA	D	3101	1/1	0.99	0.05	21,21,21,21	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	NA	D	3102	1/1	0.99	0.05	19,19,19,19	0
3	MG	D	3001	1/1	0.99	0.04	23,23,23,23	0
3	MG	B	3001	1/1	0.99	0.04	18,18,18,18	0
5	DMS	A	8401	4/4	0.99	0.10	16,16,18,22	0
5	DMS	A	8402	4/4	0.99	0.06	22,36,36,52	0
5	DMS	D	8401	4/4	0.99	0.06	15,21,22,23	0
5	DMS	A	8504	4/4	0.99	0.10	32,45,52,100	0
4	NA	A	3101	1/1	0.99	0.05	22,22,22,22	0
5	DMS	B	8401	4/4	0.99	0.08	17,18,25,33	0
5	DMS	D	8405	4/4	0.99	0.07	26,30,41,43	0
4	NA	A	3103	1/1	0.99	0.05	34,34,34,34	0
5	DMS	C	8401	4/4	0.99	0.06	15,22,22,24	0
5	DMS	B	8403	4/4	0.99	0.08	25,29,35,37	0
5	DMS	C	8403	4/4	0.99	0.12	25,26,27,33	0
5	DMS	D	8411	4/4	0.99	0.07	32,32,37,97	0
5	DMS	A	8405	4/4	0.99	0.07	26,30,34,46	0
5	DMS	C	8405	4/4	0.99	0.07	26,27,34,35	0
3	MG	A	3002	1/1	0.99	0.04	22,22,22,22	0
4	NA	B	3101	1/1	0.99	0.06	21,21,21,21	0
4	NA	B	3103	1/1	0.99	0.05	29,29,29,29	0
3	MG	A	3001	1/1	0.99	0.04	23,23,23,23	0
5	DMS	C	8411	4/4	0.99	0.12	33,35,36,91	0
5	DMS	C	8412	4/4	0.99	0.08	31,35,41,100	0
5	DMS	C	8414	4/4	0.99	0.09	28,31,53,70	0
4	NA	C	3102	1/1	0.99	0.04	19,19,19,19	0
5	DMS	B	8411	4/4	0.99	0.06	37,37,37,79	0
4	NA	C	3103	1/1	0.99	0.09	36,36,36,36	0
3	MG	C	3002	1/1	0.99	0.05	18,18,18,18	0
4	NA	C	3101	1/1	1.00	0.06	19,19,19,19	0
4	NA	A	3102	1/1	1.00	0.04	17,17,17,17	0
4	NA	B	3102	1/1	1.00	0.04	20,20,20,20	0
3	MG	C	3001	1/1	1.00	0.03	18,18,18,18	0
3	MG	D	3002	1/1	1.00	0.04	20,20,20,20	0

6.5 Other polymers (i)

There are no such residues in this entry.